JYU DISSERTATIONS 328

Nguyen Tran Thuan

Weighted BMO, Riemann-Liouville Type Operators, and Approximation of Stochastic Integrals in Models with Jumps

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF MATHEMATICS AND SCIENCE

JYU DISSERTATIONS 328

Nguyen Tran Thuan

Weighted BMO, Riemann-Liouville Type Operators, and Approximation of Stochastic Integrals in Models with Jumps

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella julkisesti tarkastettavaksi joulukuun 9. päivänä 2020 kello 12.

Academic dissertation to be publicly discussed, by permission of the Faculty of Mathematics and Science of the University of Jyväskylä, on December 9, 2020, at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2020

Editors
Stefan Geiss
Department of Mathematics and Statistics, University of Jyväskylä
Päivi Vuorio
Open Science Centre, University of Jyväskylä

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-8442-7

ISBN 978-951-39-8442-7 (PDF)
URN:ISBN:978-951-39-8442-7
ISSN 2489-9003

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Stefan Geiss, for being a wonderful supervisor that I could imagine. Thank you for inspiring me to stochastic analysis, for teaching me with your admirable passion and tremendous patience, for energizing me whenever I feel exhausted with mathematical depression, and for encouraging me in every small achievement I have got. You showed a great trust in me and made me confident when I had started the doctoral study with a different background and very little knowledge about the research topic. This thesis would not have been possible without your conscientious guidance and support.

I am very grateful to my co-supervisor, Christel Geiss, for all fruitful discussions both inside and outside mathematics. Thank you for, as always, reading carefully and checking critically my first drafts with many helpful suggestions which lead to a significant improvement of results in this thesis. Together with Stefan Geiss, you have offered me a lot of help during my living here.

I wish to thank René Schilling, who has agreed to be the opponent at the public examination of my doctoral thesis. I also wish to thank the pre-examiners, Friedrich Hubalek and Thorsten Rheinländer, for their carefully reading and giving valuable comments in an earlier version of the thesis.

I wish to thank University of Jyväskylä, especially Department of Mathematics and Statistics, for giving an excellent working place and a friendly living environment during my doctoral study. Special thanks go to my colleagues at the department and my friends who supported me and shared with me many exciting moments in these few years. I want to thank Eija Laukkarinen for her kind help to translate the abstract of this thesis to Finnish.

I am thankful to my master's thesis supervisor, Nguyễn Văn Quảng, for guiding me the first steps in doing mathematics research. Many thanks go to Vinh University, Vietnam, and to my colleagues in Department of Mathematics for their encouragement.

Most importantly, I want to thank my dear family and my parents for their constant love and unconditional support in every step of my life. Especially, this thesis is dedicated to my beloved wife, Đinh Thanh Giang, and to my lovely daughter, Nguyễn Việt Linh. Thank you so much for always standing by me.

Jyväskylä, November 2020

Nguyễn Trần Thuận

Abstract

This thesis investigates the interplay between weighted bounded mean oscillation (BMO), Riemann–Liouville type operators applied to càdlàg processes, real interpolation, gradient type estimates for functionals on the Lévy–Itô space, and approximation for stochastic integrals with jumps.

There are two main parts included in this thesis. The first part discusses the connections between the approximation problem in L_2 or in weighted BMO, Riemann–Liouville type operators, and the real interpolation theory in a general framework (Chapter 3).

The second part provides various applications of results in the first part to several models: diffusions in the Brownian setting (Section 3.5) and certain jump models (Chapter 4) for which the (exponential) Lévy settings are typical examples (Chapter 6 and Chapter 7). Especially, for the models with jumps we propose a new approximation scheme based on an adjustment of the Riemann approximation of stochastic integrals so that one can effectively exploit the features of weighted BMO.

In our context, making a bridge from the first to the second part requires gradient type estimates for a semigroup acting on Hölder functions in both the Brownian setting (Section 3.5) and the (exponential) Lévy setting (Chapter 5). In the latter case, we consider a kind of gradient processes appearing naturally from the Malliavin derivative of functionals of the Lévy process, and we show how the gradient behaves in time depending on the "direction" one tests.

Tiivistelmä

Painotettu rajoitettu keskiheilahtelu, Riemann–Liouville-tyyppiset operaattorit ja stokastisten integraalien approksimointi malleissa, joissa on hyppyjä

Väitöskirjassa yhdistyvät painotettu rajoitettu keskiheilahtelu, càdlàg-prosesseihin sovelletut Riemann–Liouville-tyyppiset operaattorit, reaalinen interpolointi, Lévy–Itô-avaruuden funktionaalien gradienttityyppiset estimaatit sekä hyppyprosesseihin perustuvien stokastisten integraalien approksimointi. Tutkimuksen kohteena on näiden keskinäinen vuorovaikutus.

Väitöskirjassa on kaksi keskeistä osaa. Ensimmäinen osa käsittelee yhteyksiä L_2-mielessä tai painotetun rajoitetun keskiheilahtelun mielessä approksimoinnin, Riemann–Liouville tyyppisten operaattoreiden ja yleisen viitekehyksen reaalisen interpoloinnin välillä (Luku 3). Toinen osa käsittää erilaisia sovelluksia ensimmäisen osan tuloksille useissa malleissa: Brownin liikkeeseen perustuvat diffuusiot (Luku 3.5) ja tietyt hyppyprosessit (Luku 4), joista (eksponentiaaliset) Lévy-prosessit ovat tyypillisiä esimerkkejä (Luvut 6 ja 7). Erityisesti hyppyjä sisältäville malleille esitämme uuden approksimointiskeeman, joka perustuu stokastisten integraalien Riemann-approksimointiin siten, että painotetun rajoitetun keskiheilahtelun piirteitä voi hyödyntää tehokkaasti.

Tässä kontekstissa ensimmäisen ja toisen osan yhdistäminen vaatii gradienttityyppisiä estimaatteja eräälle puoliryhmälle Hölder-funktioilla sekä Brownisessa tapauksessa (Luku 3.5) että (eksponentiaalisen) Lévy-prosessin tapauksessa (Luku 5). Jälkimmäisessä käytämme Lévy-prosessin funktionaalin Malliavin-derivaatasta luonnollisesti muodostuvaa gradienttiprosessin kaltaista prosessia, ja näytämme miten gradientti muuttuu ajan suhteen riippuen testattavaksi valitusta “suunnasta”.

Contents

List of symbols

General notations

$\mathbb{R}_0$	$\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$
$\mathbb{R}_+$	$\mathbb{R}_+ := (0,\infty)$
$\inf\emptyset$	$\inf\emptyset := \infty$
0^0	$0^0 := 1$
$a \vee b$	maximum of a and b
$a \wedge b$	minimum of a and b
$\mathbb{1}_A$	indicator function of the set A
$A \succeq_c B$	$cA \geqslant B$
$A \preceq_c B$	$A \leqslant cB$
$A \sim_c B$	$\frac{1}{c}A \leqslant B \leqslant cA$
càdlàg	finite left-limits and right-continuous
sign	the sign function

Probability/Measure theory

$\mathcal{B}(\mathbb{R}^d)$	Borel σ-algebra on $\mathbb{R}^d$
λ	Lebesgue measure on $\mathbb{R}$
δ_x	Dirac measure at x
$\mathrm{supp}(\mu)$	support of the measure μ
$\lvert\mu\rvert$	variation of the (finite signed) measure μ
$\lVert\mu\rVert_{\mathrm{TV}}$	total variation of the (finite signed) measure μ
$\nu \ll \mu$	ν is absolutely continuous w.r.t. μ
$\mathbb{P}_X$	push-forward measure of the measure $\mathbb{P}$ w.r.t. X
$\mathbb{E}_{\mathcal{G}}[X]$	conditional expectation of X given a σ-algebra $\mathcal{G}$

Time-net

$\mathcal{T}_{\mathrm{det}}$	family of deterministic time-nets (on $[0,T]$)
$\lVert\tau\rVert_\theta$	mesh size of $\tau \in \mathcal{T}_{\mathrm{det}}$ w.r.t. $\theta \in (0,1]$

Spaces of functions

$B_b(\mathbb{R})$	Borel bounded functions $f:\mathbb{R}\to\mathbb{R}$
$C^n(\mathbb{R})$	n-times continuously differentiable functions
$C^\infty(\mathbb{R})$	$C^\infty(\mathbb{R}) := \cap_{n=1}^\infty C^n(\mathbb{R})$
$C_b^\infty(\mathbb{R})$	functions $f \in B_b(\mathbb{R}) \cap C^\infty(\mathbb{R})$ with derivatives $f^{(k)} \in B_b(\mathbb{R})$, $k \geqslant 1$
$C_b^0(\mathbb{R})$	bounded continuous functions vanishing at zero
$\mathrm{Höl}_\eta(\mathbb{R})$	Hölder continuous functions with the exponent $\eta \in (0,1]$

Stochastic processes

T	$T \in (0,\infty)$ finite time horizon
$\mathbb{I}$	$\mathbb{I} = [0,T]$ or $\mathbb{I} = [0,T)$
$\mathbb{F}$	$\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ filtration
$\mathrm{CL}(\mathbb{I})$	càdlàg on $\mathbb{I}$ and $\mathbb{F}$-adapted processes
$\mathrm{CL}_0(\mathbb{I})$	$X \in \mathrm{CL}_0(\mathbb{I}) \Leftrightarrow X \in \mathrm{CL}(\mathbb{I})$ and $X_0 = 0$ a.s.
$\mathrm{CL}^+(\mathbb{I})$	$X \in \mathrm{CL}^+(\mathbb{I}) \Leftrightarrow X \in \mathrm{CL}(\mathbb{I})$ and $X \geqslant 0$
$\mathcal{S}_t$	collection of all stopping times $\rho:\Omega\to[0,t]$
$\langle M, M\rangle$	predictable quadratic variation of M (under $\mathbb{P}$)
$\mathbb{P}^*$	minimal martingale measure

Abbreviations

BMO	Bounded mean oscillation
FS	Föllmer–Schweizer
GKW	Galtchouk–Kunita–Watanabe
MVH	mean-variance hedging
LRM	local risk-minimizing

CHAPTER 1

Introduction

Assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ with finite time horizon $T > 0$. There are various applications in which stochastic processes $\varphi = (\varphi_t)_{t\in[0,T)}$ appear that have a singularity when $t \uparrow T$, for example in L_p for some $p \in [1, \infty]$. Examples are gradient processes obtained from (semi-linear) parabolic backward PDEs within the Feynman–Kac theory, where these processes appear as integrands in stochastic integral representations (see Section 3.5) or in backward stochastic differential equations as gradient processes. The same type of processes appear also as gradient processes originating from convolution semi-groups based on Lévy processes and that are used, for example, in Galtchouk–Kunita–Watanabe projections (see Chapters 5 and 6).

If one analyses these examples, then one realizes the following:

- **Self-similarity**: There is a Markovian structure behind that generates a self-similarity in the sense that, given $a \in (0, T)$ and $B \in \mathcal{F}_a$ of positive measure, then $(\varphi_t)_{t\in[a,T)}$ restricted to B has similar properties as $(\varphi_t)_{t\in[0,T)}$. If one is interested in good distributional estimates of $(\varphi_t)_{t\in[0,T)}$ or functionals of it, then it is useful to consider the BMO-setting: the particular feature of BMO-estimates is that one uses conditional L_2-estimates, where one might exploit conditional orthogonality, in order to deduce L_p-estimates for $p > 2$ or exponential estimates by John–Nirenberg type theorems.
- **Polynomial blow-up**: In the problems mentioned above the size of the singularity of φ (or, again, a functional of it) increases polynomially in time with a rate $(T - t)^{-\alpha}$ for some $\alpha > 0$. In particular, this often occurs in the presence of Hölder functionals as terminal conditions in backward problems.

The above observations lead to an interplay between *Riemann–Liouville (type) operators*, *BMO*, and the *real interpolation method*. These components interact as follows: We realized that the Riemann–Liouville operators allow for a transformation of a stochastic process with a certain singularity when $t \uparrow T$ into a stochastic process without this singularity (but without loosing any information about the process one is starting from). In particular, this is of interest for martingales. By the obtained formulas this opens a link to real interpolation theory, which has a natural explanation as we interpolate with a two-parametric scale between, for example, martingales without singularity and martingales with a singularity. As a consequence of the self-similarity of the singular process one is starting from, it is natural to think that the Riemann–Liouville operator turns this process into a BMO-process by removing the singularity but keeping the self-similarity. Therefore, we consider the stochastic processes transformed by the Riemann–Liouville type operator in the BMO-setting. One starting point to investigate the connections between Riemann–Liouville operators, BMO, and real interpolation is an approximation problem for stochastic integrals, so that we will deal with four objects that interact with each other.

In the second part of the thesis, we give applications of the first part to the discrete-time approximation problem for stochastic integrals in both Brownian setting and models with jumps. 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 start with the well-known Black–Scholes model. Assume that the discounted price process of a risky asset is modelled by the geometric Brownian motion $S_t = \mathrm{e}^{W_t - \frac{t}{2}}$, where $W = (W_t)_{t\in[0,T]}$ is a standard Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$. Here, $T > 0$ is a fixed finite time horizon and the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is assumed to satisfy the usual conditions (right continuity and completeness). For a Borel function $g\colon (0,\infty) \to \mathbb{R}$ with $g(S_T) \in L_2(\mathbb{P})$, one has the representation

$$g(S_T) = \mathbb{E}g(S_T) + \int_0^T \partial_y G(t, S_t)\mathrm{d}S_t, \tag{1.1}$$

where $G(t,y) := \mathbb{E}g(yS_{T-t})$ is the option price function and $(\partial_y G(t,S_t))_{t\in[0,T)}$ is the so-called delta-hedging strategy of the payoff $g(S_T)$. In mathematical finance, the stochastic integral in (1.1) can be interpreted as the theoretical hedging portfolio which is readjusted continuously in time. However, in practice this task is impossible because one can only rebalance the hedging portfolio finitely many times. This fact leads to a substitution of the stochastic integral by a discretised version which causes the discretisation error.

Let us recall some known results regarding the error caused from the Riemann approximation of the stochastic integral. For a deterministic time-net $\tau = (t_i)_{i=0}^n$, $0 = t_0 < t_1 \cdots < t_n = T$, we define the error process $E(g;\tau) = (E_t(g;\tau))_{t\in[0,T]}$ by

$$E_t(g;\tau) := \int_0^t \partial_y G(u, S_u)\mathrm{d}S_u - \sum_{i=1}^n \partial_y G(t_{i-1}, S_{t_{i-1}})(S_{t_i\wedge t} - S_{t_{i-1}\wedge t}). \tag{1.2}$$

For $\theta \in (0,1]$, we define the adapted time-nets $\tau_n^\theta = (t_{i,n}^\theta)_{i=0}^n$ by setting

$$t_{i,n}^\theta := T(1 - \sqrt[\theta]{1 - i/n}).$$

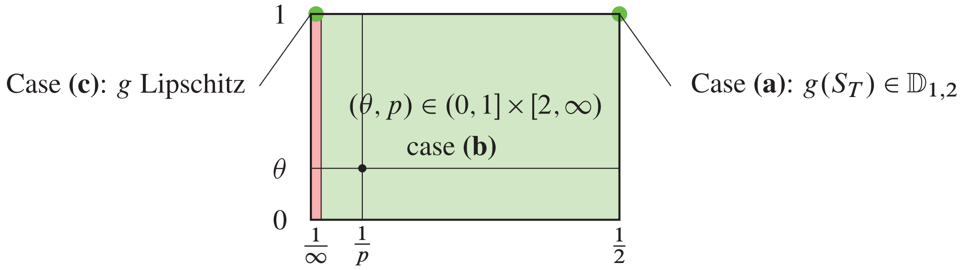
Then we have the following statements (among others), where $\theta \in (0,1]$ and $p \in [2,\infty)$:

Table 1.1:

	approximation rate	equivalent condition
(a)	$\sup_{n\geqslant 1} \sqrt{n}\, \big\| E_T(g;\tau_n^1)\big\|_{L_2(\mathbb{P})} < \infty$	$g(S_T) \in \mathbb{D}_{1,2}$
(b)	$\sup_{n\geqslant 1} \sqrt{n} \| E_T(g;\tau_n^\theta)\|_{L_p(\mathbb{P})} < \infty$	$\mathbb{E}\big\| \int_0^T (T-t)^{1-\theta} \big\|\partial_t G(t,S_t)\big\|^2 \mathrm{d}t\big\|^{\frac{p}{2}} < \infty$
(c)	$\sup_{n\geqslant 1} \sqrt{n} \| E(g;\tau_n^1)\|_{\mathrm{BMO}_2^S([0,T])} < \infty$	g is (equivalent to) a Lipschitz function

The case **(a)** was considered by C. Geiss and S. Geiss in [21] where $\mathbb{D}_{1,2}$ is the Malliavin–Sobolev space of differentiable random variables in the Malliavin sense. Several results in the L_2-setting were also obtained by Zhang [62], Gobet and Temam [31]. The case **(c)** was examined by S. Geiss [25] where the space $\mathrm{BMO}_2^S([0,T])$ is given in Section 2.2. The case **(b)** was studied by S. Geiss and Toivola [27] where the parameter θ stands for the fractional smoothness in the sense of fractional order Malliavin–Sobolev spaces obtained by real interpolation. The non-uniform time-nets τ_n^θ allow to achieve the optimal rate $\frac{1}{\sqrt{n}}$ by compensating the lack of smoothness when $g(S_T) \notin \mathbb{D}_{1,2}$.

One can visualize the cases **(a)**, **(b)**, and **(c)** as follows, where the known parts are in green and the unknown parts are in red:



The case $(\theta, p) \in (0,1) \times \{\infty\}$ was open. Here, in the limiting case $p = \infty$ we choose the (weighted) BMO spaces rather than the L_∞ spaces because it is in a line with scenarios in real analysis. Namely, we are going to investigate the case (θ, ∞) where the parameter $\theta \in (0,1)$ describes the fractional smoothness and ∞ means the (weighted) BMO spaces.

For the error process given in (1.2), using conditional Itô's isometry yields that for any $a \in [0, T)$, a.s.,

$$\begin{aligned}&\mathbb{E}_{\mathcal{F}_a}\left[|E_T(g;\tau) - E_a(g;\tau)|^2\right]\\ &\quad= \mathbb{E}_{\mathcal{F}_a}\left[\int_a^T \left|\partial_y G(u, S_u) - \sum_{i=1}^n \partial_y G(t_{i-1}, S_{t_{i-1}}) \mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 S_u^2 \mathrm{d}u\right]. \qquad (1.3)\end{aligned}$$

The quantity on the left-hand side of (1.3) appears in the definition of weighted BMO-norms of $E(g;\tau)$ (see Section 2.2), and the equality (1.3) suggests that one can reduce the original probabilistic problem to a “more deterministic” setting where the corresponding quadratic variation is employed. Therefore, in Chapter 3 we focus on investigating the approximation problem for the quadratic variation of the original error process.

This thesis contains original works of three preprints [29, 60, 61], where the author of this thesis has actively taken part in the research of the joint preprint [29]. Chapter 3 is written based on [29], Chapter 5 is based on [29, 60], Chapters 4 and 6 are based on [60], Chapter 7 is based on [61].

CHAPTER 2

Preliminaries

This section provides notations and summarizes some facts about weighted BMO spaces, Riemann–Liouville type operators, interpolation spaces, and time-nets.

2.1. Notations

General notations and conventions. Denote $\mathbb{R}_+ := (0, \infty)$ and $\mathbb{R}_0 := \mathbb{R} \backslash \{0\}$. For $a, b \in \mathbb{R}$, we set $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. In particular, $a^+ := a \vee 0$, $a^- := (-a) \vee 0$. For $A, B \geqslant 0$ and $c \geqslant 1$, the notation $A \sim_c B$ stands for $\frac{1}{c} A \leqslant B \leqslant c A$. The corresponding one-sided inequalities are abbreviated by $A \succeq_c B$ and $A \preceq_c B$.

The sign function is defined by setting $\operatorname{sign}(x) := 1$ for $x \geqslant 0$ and $\operatorname{sign}(x) := -1$ for $x < 0$.

For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable map $X : \Omega \to \mathbb{R}^d$, where $\mathbb{R}^d$ is equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$, the law of X is denoted by $\mathbb{P}_X$. If X is integrable (non-negative), then the (generalized) conditional expectation of X given a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is denoted by $\mathbb{E}_{\mathcal{G}}[X]$. We also agree on the notation $L_p(\mathbb{P}) := L_p(\Omega, \mathcal{F}, \mathbb{P})$.

We set $0^0 := 1$ and $\inf \emptyset := \infty$.

Notations about measures.

- The Lebesgue measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is denoted by λ.
- Given a finite signed measure μ on $\mathcal{B}(\mathbb{R})$, we denote by $|\mu| := \mu^+ + \mu^-$ its *variation* , where μ^+ and μ^- are the positive and negative variations of μ respectively (see, e.g., [50]). The *total variation* of μ is denoted by $\|\mu\|_{\mathrm{TV}} := |\mu|(\mathbb{R})$.
- For two measures μ and ν on a measurable space $(\Omega, \mathcal{F})$, we write $\nu \ll \mu$ if ν is absolutely continuous with respect to μ.
- For a set $A \in \mathcal{F}$ with $\mu(A) \in (0, \infty)$, we let μ_A be the normalized restriction of μ to the trace σ-algebra $\mathcal{F}|_A$.

Let μ be a measure on $\mathcal{B}(\mathbb{R}^d)$, then the *support* of μ is the closed set defined by

$$\operatorname{supp}(\mu) := \{x \in \mathbb{R}^d : \mu(U_\varepsilon(x)) > 0 \text{ for all } \varepsilon > 0\},$$

where $U_\varepsilon(x)$ is the open Euclidean ball centered at x with radius $\varepsilon > 0$.

Given a random variable $X : \Omega \to \mathbb{R}^d$, we let $\operatorname{supp}(X) := \operatorname{supp}(\mathbb{P}_X)$. One knows that $\mathbb{P}(\{X \in \operatorname{supp}(X)\}) = 1$, and that for independent random variables $X : \Omega \to \mathbb{R}^m$ and $Y : \Omega \to \mathbb{R}^n$ it holds $\operatorname{supp}((X, Y)) = \operatorname{supp}(X) \times \operatorname{supp}(Y)$.

Notations about stochastic processes. Let $T > 0$ be a fixed finite time horizon, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be 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. The conditions imposed on $\mathbb{F}$ allow us to assume that every martingale adapted to this filtration is *càdlàg* (right continuous with left limits). We use the following notations and conventions where

$$\mathbb{I} = [0, T] \quad \text{or} \quad \mathbb{I} = [0, T).$$

- For processes $X = (X_t)_{t\in\mathbb{I}}$ and $Y = (Y_t)_{t\in\mathbb{I}}$, we write $X = Y$ to indicate that $X_t = Y_t$ for all $t \in \mathbb{I}$ a.s., and similarly when the relation "=" is replaced by some other standard relations such as "$\leqslant$", "$\geqslant$", etc.
- For a càdlàg process $X = (X_t)_{t\in\mathbb{I}}$, the process $X_- = (X_{t-})_{t\in\mathbb{I}}$ is defined by setting $X_{0-} := X_0$ and $X_{t-} := \lim_{0<s\uparrow t} X_s$ for $t \in \mathbb{I}\backslash\{0\}$. We set $\Delta X := X - X_-$.
- $\mathrm{CL}(\mathbb{I})$ denotes the family of all càdlàg on $\mathbb{I}$ and $\mathbb{F}$-adapted processes.
- $\mathrm{CL}_0(\mathbb{I})$ (resp. $\mathrm{CL}^+(\mathbb{I})$) consists of all $X \in \mathrm{CL}(\mathbb{I})$ with $X_0 = 0$ a.s. (resp. $X \geqslant 0$).
- $\mathcal{P}$ is the predictable σ-algebra[1] on $\Omega \times [0,T]$ and $\widetilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{B}(\mathbb{R})$.

We recall some notions regarding semimartingales on the finite time interval $[0,T]$.

- A process $M \in \mathrm{CL}([0,T])$ is called a local (resp. locally square integrable) martingale if there is a sequence of non-decreasing stopping times $(\rho_n)_{n\geqslant 1}$ taking values in $[0,T]$ such that $\mathbb{P}(\rho_n < T) \to 0$ as $n \to \infty$ and the stopped process $M^{\rho_n} = (M_{t\wedge\rho_n})_{t\in[0,T]}$ is a martingale (resp. square integrable martingale) for all $n \geqslant 1$. Let $\mathcal{M}_2^0(\mathbb{P})$ be the space of all square integrable $\mathbb{P}$-martingales $M = (M_t)_{t\in[0,T]}$ with $M_0 = 0$ a.s.
- A process $S \in \mathrm{CL}([0,T])$ is called a semimartingale if S can be written as a sum of a local martingale and a process of finite variation a.s. The *quadratic covariation* of two semimartingales S and R is denoted by $[S,R]$. The predictable $\mathbb{Q}$-compensator of $[S,R]$, if it exists, is denoted by $\langle S,R\rangle^{\mathbb{Q}}$, where $\mathbb{Q}$ is a probability measure. We will omit the reference measure if there is no risk of confusion.
- Let M, N be locally square integrable martingales under a probability measure $\mathbb{Q}$. Then M and N are said to be $\mathbb{Q}$-*orthogonal* if $[M,N]$ is a local martingale under $\mathbb{Q}$, or equivalently, $\langle M,N\rangle^{\mathbb{Q}} = 0$.

2.2. Weighted bounded mean oscillation (BMO) spaces

For $t > 0$, we denote by $\mathcal{S}_t$ the collection of all stopping times $\rho\colon \Omega \to [0,t]$. Let

$$\mathbb{I} = [0,T) \quad \text{or} \quad \mathbb{I} = [0,T].$$

Definition 2.2.1 ([25, 29], *Weighted* BMO *and weight regularity*)**.** Let $p \in (0,\infty)$. For $Y \in \mathrm{CL}_0(\mathbb{I})$ and $\Phi \in \mathrm{CL}^+(\mathbb{I})$, we define

$$\|Y\|_{\mathrm{BMO}_p^{\Phi}(\mathbb{I})} := \inf\left\{c \geqslant 0 : \mathbb{E}_{\mathcal{F}_\rho}\left[|Y_t - Y_{\rho-}|^p\right] \leqslant c^p \Phi_\rho^p \text{ a.s. } \forall \rho \in \mathcal{S}_t, \forall t \in \mathbb{I}\right\},$$
$$\|Y\|_{\mathrm{bmo}_p^{\Phi}(\mathbb{I})} := \inf\left\{c \geqslant 0 : \mathbb{E}_{\mathcal{F}_\rho}\left[|Y_t - Y_{\rho}|^p\right] \leqslant c^p \Phi_\rho^p \text{ a.s. } \forall \rho \in \mathcal{S}_t, \forall t \in \mathbb{I}\right\},$$
$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})} := \inf\left\{c \geqslant 1 : \mathbb{E}_{\mathcal{F}_\rho}\left[\sup\nolimits_{\rho\leqslant t\in\mathbb{I}} \Phi_t^p\right] \leqslant c^p \Phi_\rho^p \text{ a.s. } \forall \text{ stopping times } \rho\colon \Omega \to \mathbb{I}\right\}.$$

If $\|Y\|_{\Theta} < \infty$ (resp. $\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})} < \infty$), then we write $Y \in \Theta$ for $\Theta \in \{\mathrm{BMO}_p^{\Phi}(\mathbb{I}), \mathrm{bmo}_p^{\Phi}(\mathbb{I})\}$ (resp. $\Phi \in \mathcal{SM}_p(\mathbb{I})$). In the non-weighted case, i.e. $\Phi \equiv 1$, we drop Φ and simply use the notation $\mathrm{BMO}_p(\mathbb{I})$ and $\mathrm{bmo}_p(\mathbb{I})$.

The theory of classical non-weighted BMO- and bmo-martingales can be found in Dellacherie and Meyer [16, Ch.VII] or Protter [47, Ch.IV], and they were used later in different contexts (see, e.g., Choulli, Krawczyk and Stricker [11], Delbaen et al. [15]).

---

[1] $\mathcal{P}$ is the σ-algebra generated by $\{A \times \{0\} : A \in \mathcal{F}_0\} \cup \{A \times (s,t] : 0 \leqslant s < t \leqslant T, A \in \mathcal{F}_s\}$.

It is clear from the definition that if $Y \in \mathrm{CL}_0(\mathbb{I})$ has continuous paths, then $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} = \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$. When Y has jumps, then the relation between weighted BMO and weighted bmo is as follows (the proof is provided in [29, Propositions A.5]).

Proposition 2.2.2. *For $\Phi \in \mathrm{CL}^+(\mathbb{I})$, $Y \in \mathrm{CL}_0(\mathbb{I})$,*

$$|\Delta Y|_{\Phi,\mathbb{I}} := \inf\{c > 0 : |\Delta Y_t| \leqslant c\Phi_t \quad \textit{for all } t \in \mathbb{I} \ \textit{a.s.}\},$$

and for $p \in (0,\infty)$ the following assertions are true:

(1) $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \leqslant 2^{(\frac{1}{p}-1)^+}\big(\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} + |\Delta Y|_{\Phi,\mathbb{I}}\big)$.

(2) *If $\mathbb{E}|\sup_{s\in[0,t]}\Phi_s|^p < \infty$ for all $t \in \mathbb{I}$, then*

$$\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \quad \textit{and} \quad |\Delta Y|_{\Phi,\mathbb{I}} \leqslant 2^{\frac{1}{p}\vee 1}\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}.$$

As verified in [29, Propositions A.4 and A.1], the definitions of weighted bmo and $\mathcal{SM}_p$ can be simplified by using deterministic times instead of stopping times, which means

$$\begin{aligned}
&\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})} = \inf\{c \geqslant 1 : \mathbb{E}_{\mathcal{F}_a}\big[\sup\nolimits_{a\leqslant t\in\mathbb{I}}\Phi_t^p\big] \leqslant c^p\Phi_a^p \quad \text{a.s. for all } a \in \mathbb{I}\},\\
&\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} = \inf\{c \geqslant 0 : \mathbb{E}_{\mathcal{F}_a}\big[|Y_t - Y_a|^p\big] \leqslant c^p\Phi_a^p \quad \text{a.s. for all } a \in [0,t] \text{ and } t \in \mathbb{I}\}. \quad (2.2.1)
\end{aligned}$$

Definition 2.2.3 ([25], *Reverse Hölder inequality*)**.** Let $\mathbb{Q}$ be a probability measure equivalent 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 $c > 0$ such that U satisfies the following *reverse Hölder inequality*

$$\sqrt[s]{\mathbb{E}_{\mathcal{F}_\rho}[U^s]} \leqslant c\,\mathbb{E}_{\mathcal{F}_\rho}[U] \quad \text{a.s.}, \forall \rho \in \mathcal{S}_T,$$

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

We summarize from [29, Proposition A.6] and [60, Proposition 2.5] some features of weighted BMO which play a key role in our applications. Notice that these results are *not* valid in general for weighted bmo.

Proposition 2.2.4 (Features of weighted BMO)**.** *Let $p \in (0,\infty)$.*

(1) (L_p*-estimate*) *For $r \in (0,\infty)$, there exists a constant $c_1 = c_1(p,r) > 0$ such that*

$$\|\sup\nolimits_{t\in\mathbb{I}}|Y_t|\|_{L_p(\mathbb{P})} \leqslant c_1\|\sup\nolimits_{t\in\mathbb{I}}\Phi_t\|_{L_p(\mathbb{P})}\|Y\|_{\mathrm{BMO}_r^\Phi(\mathbb{I})}.$$

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

(3) (*Change of measure*) *Let $\mathbb{I} = [0,T]$. 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 $c_3 = c(s,p,\|\Phi\|_{\mathcal{SM}_p(\mathbb{Q})}) > 0$ such that*

$$\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{Q})} \leqslant c_3\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}.$$

Here, $\mathrm{BMO}_p^\Phi(\mathbb{Q})$ *and $\mathcal{SM}_p(\mathbb{Q})$ mean the* BMO_p^Φ*- and the $\mathcal{SM}_p$-condition formulated under $\mathbb{Q}$ respectively.*

The benefit of Proposition 2.2.4(2) is as follows: If $p \in [2,\infty)$ (this is usually the case in applications), then one can choose $r = 2$ so that $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \sim_{c_2} \|\cdot\|_{\mathrm{BMO}_2^\Phi(\mathbb{I})}$, and then we can still exploit some similar techniques as in the L_2-theory to deal with $\|\cdot\|_{\mathrm{BMO}_2^\Phi(\mathbb{I})}$. Combining this observation with item (1) yields the following estimate provided that $\Phi \in \mathcal{SM}_p(\mathbb{I})$, $p \in [2,\infty)$:

$$\|\sup\nolimits_{t\in\mathbb{I}}|Y_t|\|_{L_p(\mathbb{P})} \leqslant c_1c_2\|\sup\nolimits_{t\in\mathbb{I}}\Phi_t\|_{L_p(\mathbb{P})}\|Y\|_{\mathrm{BMO}_2^\Phi(\mathbb{I})}.$$

Item (3) gives a change of the underlying measure which might be of interest for further applications in mathematical finance.

2.3. Riemann–Liouville type operator

Riemann–Liouville operators are a central object and tool in fractional calculus. It is natural and useful to extend them to random frameworks. There are two principal approaches: Directly translating the approach from fractional calculus, that uses Volterra kernels, leads to the notion of fractional processes, in particular fractional martingales. In our setting one would take a càdlàg process K and would consider

$$t \mapsto \int_0^t (t-u)^{\alpha-1} K_u \mathrm{d}u.$$

This yields to an approach natural for pathwise fractional calculus of stochastic processes and is used, for example, for Gaussian processes by Hu, Nualart and Song [33]. For our purpose we use the different approach

$$t \mapsto \int_0^T (T-u)^{\alpha-1} K_{u\wedge t} \mathrm{d}u$$

to define $\mathcal{I}_t^\alpha K$ in Definition 2.3.1 below. The idea behind the operator $\mathcal{I}^\alpha$ is to remove or reduce singularities of a càdlàg process $(K_t)_{t\in[0,T)}$ when $t \uparrow T$. As we see in Theorem 3.1.1 below, this approach is the right one to handle fractional smoothness in the Malliavin sense and in the sense of interpolation theory. One basic difference to the Volterra-kernel approach is that, starting with a (sub-, super-) martingale φ, we again obtain a (sub-, super-) martingale $\mathcal{I}^\alpha\varphi$. This second approach was exploited by S. Geiss and Toivola [28, Definition 4.2] and [27, Section 4], Applebaum and Bañuelos [2], and relates to fractional integral transforms of martingales (see, for example, Arai, Nakai and Sadasue [3]).

Definition 2.3.1 (*Riemann–Liouville type operator*)**.** For $\alpha > 0$ and a càdlàg function $K\colon [0,T) \to \mathbb{R}$, we define $\mathcal{I}^\alpha K = (\mathcal{I}_t^\alpha K)_{t\in[0,T)}$ by setting

$$\mathcal{I}_t^\alpha K := \frac{\alpha}{T^\alpha} \int_0^T (T-u)^{\alpha-1} K_{u\wedge t} \mathrm{d}u.$$

Moreover, for $\alpha = 0$ we define $\mathcal{I}_t^0 K := K_t$.

There are two reasons for using the normalizing factor $\frac{\alpha}{T^\alpha}$ in front of the integral: first, we want to interpret K as the integrand with respect to a probability measure, and secondly, this factor allows us to obtain a semigroup structure of $(\mathcal{I}^\alpha)_{\alpha\geqslant 0}$.

We summarize from [29, Section 3] some properties of $\mathcal{I}^\alpha$:

(1) (*Semigroup*) $\mathcal{I}_t^\alpha(\mathcal{I}_\cdot^\beta K) = \mathcal{I}_t^{\alpha+\beta} K$ for $t \in [0,T)$, $\alpha, \beta \geqslant 0$.
(2) (*Inverse formula*) $K_t = \left(\frac{T-t}{T}\right)^{-\alpha} \mathcal{I}_t^\alpha K - \frac{\alpha}{T^{-\alpha}} \int_0^t (T-u)^{-\alpha-1} \mathcal{I}_u^\alpha K \mathrm{d}u$, $t \in [0,T)$, $\alpha \geqslant 0$.
(3) (*Martingale preservation*) If $(\varphi_t)_{t\in[0,T)}$ is a càdlàg martingale (super-, or sub-martingale), then $(\mathcal{I}_t^\alpha \varphi)_{t\in[0,T)}$ is a càdlàg martingale (super-, or sub-martingale).

The semigroup structure can be also understood from equation (2.3.1) below in the martingale setting.

Proposition 2.3.2. *For $\alpha > 0$, a càdlàg martingale $\varphi = (\varphi_t)_{t\in[0,T)} \subseteq L_2(\mathbb{P})$ and $0 \leqslant a < t < T$ one has, a.s.,*

$$\mathcal{I}_t^\alpha \varphi = \varphi_0 + \int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha \mathrm{d}\varphi_u, \tag{2.3.1}$$

$$\mathbb{E}_{\mathcal{F}_a}\Big[|\mathcal{I}_t^\alpha \varphi - \mathcal{I}_a^\alpha \varphi|^2\Big] = 2\alpha \mathbb{E}_{\mathcal{F}_a}\left[\int_a^T |\varphi_{u\wedge t} - \varphi_a|^2 \left(\frac{T-u}{T}\right)^{2\alpha-1} \frac{\mathrm{d}u}{T}\right], \tag{2.3.2}$$

$$\mathbb{E}_{\mathcal{F}_a}\Big[|\mathcal{I}_t^\alpha \varphi - \mathcal{I}_a^\alpha \varphi|^2\Big] + \left(\frac{T-a}{T}\right)^{2\alpha} |\varphi_a|^2 = 2\alpha \mathbb{E}_{\mathcal{F}_a}\left[\int_a^T |\varphi_{u\wedge t}|^2 \left(\frac{T-u}{T}\right)^{2\alpha-1} \frac{\mathrm{d}u}{T}\right]. \tag{2.3.3}$$

PROOF. See the proof of [29, Proposition 3.8]. □

2.4. Interpolation spaces

Let (E_0, E_1) be a couple of real Banach spaces such that E_0 and E_1 are continuously embedded into some topological Hausdorff space X. For $x \in E_0 + E_1 := \{x = x_0 + x_1 : x_i \in E_i\}$ and $v \in (0,\infty)$, we define the K-functional

$$K(v,x;E_0,E_1) := \inf\{\|x_0\|_{E_0} + v\|x_1\|_{E_1} : x = x_0 + x_1\}.$$

Given $(\theta, q) \in (0,1) \times [1,\infty]$, we let

$$(E_0,E_1)_{\theta,q} := \Big\{x \in E_0 + E_1 : \|x\|_{(E_0,E_1)_{\theta,q}} := \|v \mapsto v^{-\theta} K(v,x;E_0,E_1)\|_{L_q((0,\infty),\frac{\mathrm{d}v}{v})} < \infty\Big\}.$$

We obtain a family of Banach spaces $((E_0,E_1)_{\theta,q}, \|\cdot\|_{(E_0,E_1)_{\theta,q}})$ with the order

$$(E_0,E_1)_{\theta,q_0} \subseteq (E_0,E_1)_{\theta,q_1} \quad \text{for all } \theta \in (0,1) \text{ and } 1 \leqslant q_0 \leqslant q_1 \leqslant \infty.$$

Moreover, if $E_1 \subseteq E_0$ with $\|x\|_{E_0} \leqslant c\|x\|_{E_1}$ for some $c > 0$, then one has

$$(E_0,E_1)_{\theta_0,q_0} \subseteq (E_0,E_1)_{\theta_1,q_1} \quad \text{for all } 0 < \theta_1 < \theta_0 < 1 \text{ and } q_0, q_1 \in [1,\infty].$$

Given a linear operator $T : E_0 + E_1 \to F_0 + F_1$ with $T : E_i \to F_i$ for $i = 0,1$, we use that the real interpolation method is an exact interpolation functor, i.e.

$$\|T : (E_0,E_1)_{\theta,q} \to (F_0,F_1)_{\theta,q}\| \leqslant \|T : E_0 \to F_0\|^{1-\theta} \|T : E_1 \to F_1\|^\theta. \tag{2.4.1}$$

For more information about the real interpolation method, the reader is referred to Bergh and Löfström [7].

We now give two types of Banach spaces obtained by interpolation which will be used later. Given a real Banach space E and $(q,s) \in [1,\infty] \times \mathbb{R}$, we use the Banach spaces

$$\ell_q^s(E) := \{(x_k)_{k=0}^\infty : \|(x_k)_{k=0}^\infty\|_{\ell_q^s(E)} := \|(2^{ks}\|x_k\|_E)_{k=0}^\infty\|_{\ell_q} < \infty\}$$

and set $\ell_q(E) := \ell_q^0(E)$. Here, ℓ_q consists of all q-summable sequences of real numbers where the supremum is taken if $q = \infty$. For $q_0, q_1, q \in [1,\infty]$ and $s_0, s_1 \in \mathbb{R}$ with $s_0 \neq s_1$, and $\theta \in (0,1)$, [7, Theorem 5.6.1] implies that

$$(\ell_{q_0}^{s_0}(E), \ell_{q_1}^{s_1}(E))_{\theta,q} = \ell_q^s(E) \quad \text{where } s := (1-\theta)s_0 + \theta s_1, \tag{2.4.2}$$

and where the norms are equivalent up to a multiplicative constant.

We turn to Hölder spaces and their interpolation. For $\eta \in [0,1]$, we define

$$\mathrm{Höl}_\eta(\mathbb{R}) := \left\{ f:\mathbb{R} \to \mathbb{R} \text{ Borel } : |f|_{\mathrm{Höl}_\eta(\mathbb{R})} := \sup_{-\infty<x<y<\infty} \frac{|f(x)-f(y)|}{|x-y|^\eta} < \infty \right\},$$

$$\begin{aligned}\text{Höl}^0_\eta(\mathbb{R}) &:= \{f \in \text{Höl}_\eta(\mathbb{R}) : f(0) = 0\},\\ \text{Höl}^0_{\eta,q}(\mathbb{R}) &:= (C^0_b(\mathbb{R}), \text{Höl}^0_1(\mathbb{R}))_{\eta,q} \quad \text{for } (\eta, q) \in \times(0,1)\times[1,\infty],\end{aligned}$$

where $C^0_b(\mathbb{R})$ is the family of all bounded continuous functions vanishing at zero, which is a Banach space with the supremum norm. It follows from the reiteration theorem (see Bergh and Löfström [7, Theorem 3.5.3]) that

$$(\text{Höl}^0_{\eta_0,q_0}(\mathbb{R}), \text{Höl}^0_{\eta_1,q_1}(\mathbb{R}))_{\theta,q} = \text{Höl}^0_{\eta,q}(\mathbb{R}) \quad \text{for } \eta := (1-\theta)\eta_0 + \theta\eta_1,$$

where $\theta, \eta_0, \eta_1 \in (0,1)$ with $\eta_0 \neq \eta_1$, $q, q_0, q_1 \in [1,\infty]$, and the norms are equivalent up to a multiplicative constant. By the above definitions $(\text{Höl}^0_\eta(\mathbb{R}), |\cdot|_{\text{Höl}_\eta(\mathbb{R})})$ is a Banach space, and for $\eta \in (0,1)$ we have that $\text{Höl}^0_{\eta,\infty}(\mathbb{R}) = \text{Höl}^0_\eta(\mathbb{R})$ with equivalent norms up to a multiplicative constant.

2.5. Time-nets

Let $\mathcal{T}_{\text{det}}$ be the family of all *deterministic* time-nets $\tau = (t_i)_{i=0}^n$ on $[0,T]$ with $0 = t_0 < t_1 < \cdots < t_n = T$, $n \geqslant 1$.

The mesh size of $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\text{det}}$ is measured with respect to a $\theta \in (0,1]$ by

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

For $\theta \in (0,1]$ and for the *adapted time-nets* $\tau_n^\theta = (t_{i,n}^\theta)_{i=0}^n$ defined by

$$t_{i,n}^\theta := T(1 - \sqrt[\theta]{1 - i/n}), \tag{2.5.1}$$

we have

$$\|\tau_n^\theta\|_1 \leqslant T/(\theta n) \quad \text{and} \quad \|\tau_n^\theta\|_\theta \leqslant T^\theta/(\theta n). \tag{2.5.2}$$

One remarks that the smaller θ is, the more the time points of τ_n^θ are concentrated near T. The reason for using those adapted time-nets is to compensate the growth of gradient processes.

CHAPTER 3

Approximation, Riemann–Liouville type operator, and Interpolation

3.1. The L_2-setting revisited

The first result makes a link between the approximation, the Riemann–Liouville type operator, and interpolation in the L_2-setting. This will be extended later to the setting of weighted bounded mean oscillation.

Theorem 3.1.1. *Let $\theta \in (0,1)$. For a càdlàg martingale $\varphi = (\varphi_t)_{t\in[0,T)} \subseteq L_2(\mathbb{P}) =: H$ with the discrete-time version*

$$\varphi^d = (\varphi_{t_k})_{k=0}^\infty \quad \textit{with} \quad t_k := T(1-2^{-k}),$$

the following assertions are equivalent:

(1) *There exists a constant $c>0$ such that for all $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\rm det}$,*

$$\mathbb{E}\int_0^T \left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u \leqslant c^2 \|\tau\|_\theta.$$

(2) $\left(\mathcal{I}_t^{\frac{1-\theta}{2}}\varphi\right)_{t\in[0,T)}$ *is closable in $L_2(\mathbb{P})$.*

(3) $\varphi^d \in (\ell_2^{-1/2}(H), \ell_\infty(H))_{\theta,2}$.

Theorem 3.1.1(3) states that φ^d belongs to the space obtained by interpolating between two end-point spaces $\ell_2^{-1/2}(H)$ and $\ell_\infty(H)$ with the parameters $(\theta,2)$. Let us comment on these two end-points. By the definition of $\ell_2^{-1/2}(H)$ and the monotonicity of $t\mapsto \|\varphi_t\|_H$, we have

$$\varphi^d \in \ell_2^{-1/2}(H) \Longleftrightarrow \int_0^T \|\varphi_t\|_H^2 \mathrm{d}t < \infty,$$

and the condition $\int_0^T \|\varphi_t\|_H^2 \mathrm{d}t<\infty$ typically appears when φ is the integrand of certain stochastic integrals. On the other hand,

$$\varphi^d \in \ell_\infty(H) \Longleftrightarrow \sup\nolimits_{t\in[0,T)} \|\varphi_t\|_H < \infty,$$

and the finiteness implies that the martingale φ is closable in H.

Proof of Theorem 3.1.1. Because $(\|\varphi_{t_k}\|_H)_{k=0}^\infty$ is non-decreasing, we get for $s\in\mathbb{R}$ that

$$\frac{\|(\varphi_{t_k})_{k=0}^\infty\|^2_{\ell_2^s(H)}}{2T^{2s}} = \sum_{k=0}^\infty (T-t_k)^{-1-2s}(t_{k+1}-t_k)\|\varphi_{t_k}\|_H^2 \sim_{c_{T,s}} \int_0^T (T-t)^{-1-2s}\|\varphi_t\|_H^2 \mathrm{d}t \tag{3.1.1}$$

for some $c_{T,s}\geqslant 1$. For $s:=(1-\theta)\left(-\frac{1}{2}\right)+\theta 0$ (so that $-1-2s=-\theta$) and for $\alpha := \frac{1-\theta}{2}$, we use Proposition 2.3.2 (equation (2.3.3)) with $a=0$ to get

$$\int_0^T (T-t)^{-\theta}\|\varphi_t\|_H^2 \mathrm{d}t = \sup_{t\in[0,T)} \frac{T^{2\alpha}}{2\alpha}\mathbb{E}[|\mathcal{I}_t^\alpha\varphi - \varphi_0|^2 + |\varphi_0|^2] = \sup_{t\in[0,T)} \frac{T^{2\alpha}}{2\alpha}\mathbb{E}|\mathcal{I}_t^\alpha \varphi|^2.$$

Now the equivalence (2)$\Leftrightarrow$(3) follows from (2.4.2) and (3.1.1). The equivalence (1)$\Leftrightarrow$(2) follows from Theorem 3.3.2(3.3.1) below applied to $M := \varphi$, $\sigma \equiv 1$, $a := 0$, and $\mathcal{G} := \{\emptyset, \Omega\}$. $\square$

3.2. The weighted BMO-setting: Results with general random measures

We now turn in the weighted BMO-setting, which can be regarded as a *localization in time* of the L_2-setting above. Our next aim is to consider the equivalence Theorem 3.1.1((1)$\Leftrightarrow$(2)) in weighted BMO, and it turns out that the orthogonality structure behind this equivalence can be generalized by using two random measures Π and Υ as given in Assumption 3.2.1.

Assumption 3.2.1. We assume random measures

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

and a progressively measurable process $(\varphi_t)_{t\in[0,T)}$, and a constant $\kappa \geqslant 1$, such that

$$\Pi(\omega,(0,b]) + \Upsilon(\omega,(0,b]) + \sup\nolimits_{t\in[0,b]} |\varphi_t(\omega)| < \infty \tag{3.2.1}$$

for all $(\omega,b) \in \Omega \times (0,T)$ and such that, for $0 \leqslant s \leqslant a < b < T$, a.s.,

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,b]} |\varphi_u - \varphi_s|^2 \Pi(\cdot,\mathrm{d}u)\right] \sim_\kappa \mathbb{E}_{\mathcal{F}_a}\left[|\varphi_a - \varphi_s|^2 \Pi(\cdot,(a,b]) + \int_{(a,b]} (b-u)\Upsilon(\cdot,\mathrm{d}u)\right]. \tag{3.2.2}$$

If (3.2.2) holds with $\preceq_\kappa$ (resp. $\succeq_\kappa$), then we denote the inequality by (3.2.2)$^{\leqslant}$ (resp. (3.2.2)$^{\geqslant}$).

We will see later that the measure Π is related to the quadratic variation of the driving process of the stochastic integral and the measure Υ describes some kind of curvature of the stochastic integral.

Under condition (3.2.1), we define for $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ the non-negative, non-decreasing, and càdlàg process $[\varphi;\tau]^\pi = ([\varphi;\tau]_t^\pi)_{t\in[0,T)}$ by setting $[\varphi;\tau]_0^\pi \equiv 0$ and

$$[\varphi;\tau]_t^\pi := \int_{(0,t]} \left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \Pi(\cdot,\mathrm{d}u) \in [0,\infty), \quad t \in (0,T), \tag{3.2.3}$$

and let $[\varphi;\tau]_T^\pi := \lim_{t\uparrow T} [\varphi;\tau]_t^\pi \in [0,\infty]$.

The next two results, Theorems 3.2.2 and 3.2.3, are an important step to characterize the approximation in weighted BMO by means of the Riemann–Liouville type fractional integral. The original idea to come up with these results is due to S. Geiss and Hujo [26, Lemma 3.8], S. Geiss and Toivola [27, Lemma 5.6].

Theorem 3.2.2 (Upper estimate)**.** *Let Assumption 3.2.1 hold with* (3.2.2)$^{\leqslant}$*. For $\theta \in (0,1]$, $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ and $a \in [t_{k-1},t_k)$, one has, a.s.,*

$$\frac{\mathbb{E}_{\mathcal{F}_a}\big[[\varphi;\tau]_T^\pi - [\varphi;\tau]_a^\pi\big]}{\|\tau\|_\theta} \leqslant \kappa \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)} (T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u) + \frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}} |\varphi_a - \varphi_{t_{k-1}}|^2 \Pi(\cdot,(a,t_k])\right].$$

PROOF. See the proof of [29, Theorem 4.3]. $\square$

Theorem 3.2.3 (Lower estimate)**.** *Let Assumption 3.2.1 hold with* (3.2.2)$^{\geqslant}$*, and let $(\theta,a) \in (0,1] \times [0,T)$.*

(1) *If* $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$, $a \in [t_{k-1}, t_k)$, *and* $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$, *then, a.s.,*

$$\frac{\mathbb{E}_{\mathcal{F}_a}\big[[\varphi;\tau]^\pi_{t_k} - [\varphi;\tau]^\pi_a\big]}{\|\tau\|_\theta} \geqslant \frac{1}{\kappa}\mathbb{E}_{\mathcal{F}_a}\left[\frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}}|\varphi_a - \varphi_{t_{k-1}}|^2 \Pi(\cdot,(a,t_k])\right].$$

(2) *There exist* $\tau_n \in \mathcal{T}_{\det}$, $n \geqslant 1$, *with* $a \in \tau_n$ *and* $\lim_n \|\tau_n\|_\theta = 0$ *such that, a.s,*

$$\liminf_n \frac{\mathbb{E}_{\mathcal{F}_a}\big[[\varphi;\tau_n]^\pi_T - [\varphi;\tau_n]^\pi_a\big]}{\|\tau_n\|_\theta} \geqslant \frac{1}{\kappa 2^{\frac{1}{\theta}+2}}\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)} (T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\right].$$

PROOF. See the proof of [29, Theorem 4.4]. □

3.3. The weighted BMO-setting: A specification of random measures

We now specialize random measures Π and Υ to the settings that will be used in Section 3.5 (the Brownian case) and in Section 5.2 (the Lévy case). Another realization for those random measures in the exponential Lévy setting will be given in Chapters 6 and 7.

Assumption 3.3.1. We assume that there are

(1) a positive continuous and adapted process $(\sigma_t)_{t\in[0,T]}$ such that $\sup_{t\in[0,T]}\sigma_t \in L_2(\mathbb{P})$ and such that there is a constant $c_\sigma \geqslant 1$ with

$$\mathbb{E}_{\mathcal{F}_a}\left[\frac{1}{b-a}\int_a^b \sigma_u^2 \mathrm{d}u\right] \sim_{c_\sigma} \sigma_a^2 \quad \text{a.s., } \forall 0 \leqslant a < b \leqslant T.$$

(2) a square integrable martingale $M = (M_t)_{t\in[0,T)}$ with $M_0 \equiv 0$.
(3) a $\varphi \in \mathrm{CL}([0,T))$ with $\mathbb{E}\sup_{u\in[a,T)}|\varphi_a\sigma_u|^2 < \infty$ for all $a \in [0,T)$.

Assume that (3.2.2) is satisfied for

$$\Pi(\omega,\mathrm{d}u) := \sigma_u^2(\omega)\mathrm{d}u \quad \text{and} \quad \Upsilon(\omega,\mathrm{d}u) := \mathrm{d}\langle M,M\rangle_u(\omega), \quad u \in [0,T).$$

Since the measure Π is defined based on σ, we denote $[\varphi;\tau]^\sigma := [\varphi;\tau]^\pi$.

From Theorem 3.2.2 and Theorem 3.2.3 we immediately deduce:

Theorem 3.3.2. *Assume Assumption 3.3.1,* $(\theta,a) \in (0,1]\times[0,T)$, *and a* σ*-algebra* $\mathcal{G} \subseteq \mathcal{F}_a$. *Then there are constants* $c_{(3.3.1)}, c_{(3.3.2)} \geqslant 1$ *depending at most on* (θ,κ,c_σ) *such that, a.s.,*

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{\det},\tau\ni a} \frac{\mathbb{E}_{\mathcal{G}}\big[[\varphi;\tau]^\sigma_T - [\varphi;\tau]^\sigma_a\big]}{\|\tau\|_\theta} \sim_{c_{(3.3.1)}} \mathbb{E}_{\mathcal{G}}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M - \mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right], \tag{3.3.1}$$

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{\det}} \frac{\mathbb{E}_{\mathcal{F}_a}\big[[\varphi;\tau]^\sigma_T - [\varphi;\tau]^\sigma_a\big]}{\|\tau\|_\theta} \sim_{c_{(3.3.2)}} \mathbb{E}_{\mathcal{F}_a}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M - \mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right] + \sup_{s\in[0,a]}\frac{T-a}{(T-s)^\theta}|\varphi_a - \varphi_s|^2\sigma_a^2. \tag{3.3.2}$$

We remark that the inequality (3.3.1) is formulated for a more general σ-algebra $\mathcal{G}$ to prove Theorem 3.1.1. In (3.3.2) such a formulation is not necessary for us.

PROOF OF THEOREM 3.3.2. Relation (3.3.2): Let $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\rm det}$. For $t_{k-1} \leqslant a < t_k$, Assumption 3.3.1 implies that

$$\mathbb{E}_{\mathcal{F}_a}\left[\frac{(T-t_{k-1})^{1-\theta}}{t_k-t_{k-1}}|\varphi_a-\varphi_{t_{k-1}}|^2\Pi(\cdot,(a,t_k])\right] \sim_{c_\sigma} |\varphi_a-\varphi_{t_{k-1}}|^2\frac{(T-t_{k-1})^{1-\theta}}{t_k-t_{k-1}}\sigma_a^2(t_k-a) \text{ a.s.}$$

Maximizing the right-hand side over t_k gives $\frac{T-a}{(T-t_{k-1})^\theta}|\varphi_a-\varphi_{t_{k-1}}|^2\sigma_a^2$ a.s. Moreover, by Proposition 2.3.2(2.3.1) and conditional Itô's isometry we have, a.s.,

$$\mathbb{E}_{\mathcal{F}_a}\left[\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M-\mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right] = \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,t]}\left(\frac{T-u}{T}\right)^{1-\theta}\mathrm{d}\langle M,M\rangle_u\right]$$

for $0 \leqslant a < t < T$ so that

$$\mathbb{E}_{\mathcal{F}_a}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M-\mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right] \sim_4 \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}\left(\frac{T-u}{T}\right)^{1-\theta}\mathrm{d}\langle M,M\rangle_u\right]$$

by Doob's maximal inequality. Now we use Theorem 3.2.2 and Theorem 3.2.3.

Relation (3.3.1) for $\mathcal{G} = \mathcal{F}_a$ follows again from Theorem 3.2.2 and Theorem 3.2.3. In the case of $\mathcal{G} \subsetneq \mathcal{F}_a$ we argue as follows: let $c_{(3.3.1)} \geqslant 1$ be the constant in (3.3.1) for $\mathcal{F}_a$, then we get

$$\frac{\mathbb{E}_{\mathcal{G}}\big[[\varphi;\tau]_T^\sigma-[\varphi;\tau]_a^\sigma\big]}{\|\tau\|_\theta} \leqslant c_{(3.3.1)}\mathbb{E}_{\mathcal{G}}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M-\mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right]$$

as well for all τ with $a \in \tau$ which implies the general inequality $\leqslant$ in (3.3.1). Regarding the remaining inequality we choose the time-nets from Theorem 3.2.3(2) to get by Fatou's lemma that, a.s.,

$$\begin{aligned}\mathbb{E}_{\mathcal{G}}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M-\mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right] &\leqslant \kappa 2^{\frac{1}{\theta}+2}\mathbb{E}_{\mathcal{G}}\left[\liminf_n \mathbb{E}_{\mathcal{F}_a}\left[\frac{[\varphi;\tau_n]_T^\sigma-[\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right]\right]\\ &\leqslant \kappa 2^{\frac{1}{\theta}+2}\liminf_n \mathbb{E}_{\mathcal{G}}\left[\mathbb{E}_{\mathcal{F}_a}\left[\frac{[\varphi;\tau_n]_T^\sigma-[\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right]\right]\\ &= \kappa 2^{\frac{1}{\theta}+2}\liminf_n \mathbb{E}_{\mathcal{G}}\left[\frac{[\varphi;\tau_n]_T^\sigma-[\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right].\end{aligned}$$

□

We now are in a position to provide a weighted BMO-version for the equivalence Theorem 3.1.1((1)⇔(2)). One recalls $[\varphi;\tau]^\pi$ from (3.2.3).

Theorem 3.3.3. *Let Assumption 3.3.1 be satisfied. Then, for $\theta \in (0,1]$ and $\Phi \in \mathrm{CL}^+([0,T))$ the following assertions are equivalent:*

(1) *There is a constant $c > 0$ such that for all $\tau \in \mathcal{T}_{\rm det}$,*

$$\|[\varphi;\tau]^\sigma\|_{\mathrm{BMO}_1^{\Phi^2}([0,T))} \leqslant c^2\|\tau\|_\theta. \tag{3.3.3}$$

(2) *One has $\mathcal{I}^{\frac{1-\theta}{2}}M \in \mathrm{bmo}_2^\Phi([0,T))$ and there is a constant $c > 0$ such that*

$$|\varphi_a-\varphi_s| \leqslant c\frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}}\frac{\Phi_a}{\sigma_a} \quad \textit{a.s., } \forall 0 \leqslant s < a < T. \tag{3.3.4}$$

If $\Phi = (\sigma_t \Psi_t)_{t\in[0,T)}$, *where* $\Psi \in \mathrm{CL}^+([0,T))$ *is pathwise non-decreasing, then* (3.3.4) *is equivalent to the existence of constants* $c_\theta > 0$ *such that*

$$|\varphi_a - \varphi_0| \leqslant c_\theta (T-a)^{\frac{\theta-1}{2}} \Psi_a \quad \text{a.s., } \forall 0 \leqslant a < T \qquad \text{if } \theta \in (0,1), \tag{3.3.5}$$

$$|\varphi_a - \varphi_s| \leqslant c_1 \left(1 + \log \frac{T-s}{T-a}\right) \Psi_a \quad \text{a.s., } \forall 0 \leqslant s < a < T \qquad \text{if } \theta = 1. \tag{3.3.6}$$

PROOF. The equivalence between (1) and (2) follows directly from the second equivalence in Theorem 3.3.2 and [29, Proposition A.4]. The equivalence between (3.3.4) and (3.3.5)-(3.3.6) follows from [29, Lemma C.1]. □

3.4. Oscillation of stochastic processes and lower bounds

This section discusses some lower bounds for (3.3.3) in the non-weighted case (i.e. $\Phi \equiv 1$). It turns out that these lower bounds are closely related to some L_∞-oscillatory quantities of the integrands. Let us start to introduce the notion of maximal oscillation of a stochastic process.

Definition 3.4.1. If $\varphi = (\varphi_t)_{t\in[0,T)}$ is a stochastic process and $t \in (0,T)$, then we let

$$\underline{\mathrm{Osc}}_t(\varphi) := \inf_{s\in[0,t)} \|\varphi_t - \varphi_s\|_{L_\infty(\mathbb{P})} \in [0,\infty],$$

$$\overline{\mathrm{Osc}}_t(\varphi) := \inf_{s\in[0,t)} \sup_{u\in[s,t]} \|\varphi_t - \varphi_u\|_{L_\infty(\mathbb{P})} \in [0,\infty].$$

The process is called of *maximal oscillation* with constant $c \geqslant 1$ if for all $t \in (0,T)$ one has

$$\underline{\mathrm{Osc}}_t(\varphi) \geqslant \frac{1}{c} \|\varphi_t - \varphi_0\|_{L_\infty(\mathbb{P})}.$$

If both sides equal infinity, then we use $c = 1$ (however, this case is not of relevance for us).

Lemma 3.4.2. *For a stochastic process* $\varphi = (\varphi_t)_{t\in[0,T)}$ *the following holds:*

1. *One has* $\underline{\mathrm{Osc}}_t(\varphi) \leqslant \overline{\mathrm{Osc}}_t(\varphi)$ *for* $t \in (0,T)$.
2. *One has* $\overline{\mathrm{Osc}}_t(\varphi) \leqslant 2\underline{\mathrm{Osc}}_t(\varphi)$ *for* $t \in (0,T)$ *if* φ *is a martingale.*
3. *If* $\varphi_a \equiv \mathbb{1}_{\mathbb{Q}\cap[0,T)}(a)$ *for* $a \in [0,T)$, *then* $0 = \underline{\mathrm{Osc}}_t(\varphi) < \overline{\mathrm{Osc}}_t(\varphi) = 1$ *for all* $t \in (0,T)$.

PROOF. (1) follows from the definition. (2) If φ is a martingale and $0 \leqslant s < t < T$, then we have

$$\sup_{u\in[s,t]} \|\varphi_t - \varphi_u\|_{L_\infty(\mathbb{P})} \leqslant \|\varphi_t - \varphi_s\|_{L_\infty(\mathbb{P})} + \sup_{u\in[s,t]} \|\varphi_u - \varphi_s\|_{L_\infty(\mathbb{P})} \leqslant 2\|\varphi_t - \varphi_s\|_{L_\infty(\mathbb{P})}.$$

Taking the infimum on both sides over $s \in [0,t)$ yields the assertion. Item (3) is obvious. □

Remark 3.4.3. In the sequel we do not need the following two statements, so that we state them without proof:

1. It is possible to construct examples such that for a given $c \in [1,\infty)$ the constant c is optimal in the definition of maximal oscillation.
2. Again by examples one can see that the constant 2 in Lemma 3.4.2(2) is optimal.

To verify a maximal oscillation we make use of the following observation:

Lemma 3.4.4. *Assume two random variables* $A, B : \Omega \to \mathbb{R}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$. *Assume a probability measure* $\mathbb{Q} \ll \mathbb{P}$ *such that* $\mathbb{E}^{\mathbb{Q}}|B| < \infty$ *and* $\mathbb{E}^{\mathbb{Q}} B = 0$. *Then*

$$\|B - A\|_{L_\infty(\mathbb{P})} \geqslant \inf_{a\in\mathbb{R}} \|B - a\|_{L_\infty(\mathbb{P})} \quad \text{implies} \quad \|B - A\|_{L_\infty(\mathbb{P})} \geqslant \frac{1}{2}\|B\|_{L_\infty(\mathbb{P})}.$$

PROOF. We may assume that $\|B-A\|_{L_\infty(\mathbb{P})}<\infty$, otherwise there is nothing to prove. Because of our assumption, for all $\varepsilon>0$ there is an $a_\varepsilon\in\mathbb{R}$ such that we have

$$\|B-A\|_{L_\infty(\mathbb{P})}\geqslant\|B\|_{L_\infty(\mathbb{P})}-|a_\varepsilon|-\varepsilon$$

and

$$\|B-A\|_{L_\infty(\mathbb{P})}\geqslant\mathbb{E}^{\mathbb{Q}}|B-a_\varepsilon|-\varepsilon\geqslant|\mathbb{E}^{\mathbb{Q}}B-a_\varepsilon|-\varepsilon=|a_\varepsilon|-\varepsilon.$$

The combination of the inequalities implies

$$\|B-A\|_{L_\infty(\mathbb{P})}\geqslant\|B\|_{L_\infty(\mathbb{P})}-|a_\varepsilon|-\varepsilon\geqslant\|B\|_{L_\infty(\mathbb{P})}-\|B-A\|_{L_\infty(\mathbb{P})}-2\varepsilon$$

so that $2\|B-A\|_{L_\infty(\mathbb{P})}\geqslant\|B\|_{L_\infty(\mathbb{P})}-2\varepsilon$. By $\varepsilon\downarrow 0$ we get our statement. □

Now we consider two examples relevant for us:

Example 3.4.5 (Markov type processes)**.** Let $(Y_t)_{t\in[0,T]}$ be a process with values in $\mathcal{R}_Y$, where $\mathcal{R}_Y=\mathbb{R}$ or $\mathcal{R}_Y=(0,\infty)$, and $Y_0\equiv y_0\in\mathcal{R}_Y$. Assume continuous transition densities $\Gamma_Y:\{(s,t):0\leqslant s<t\leqslant T\}\times\mathcal{R}_Y\times\mathcal{R}_Y\to(0,\infty)$ such that

$$\mathbb{P}(Y_t\in B\,|\,Y_s)=\int_B\Gamma_Y(s,t;Y_s,y)\mathrm{d}y\qquad\text{a.s.}$$

for $B\in\mathcal{B}(\mathcal{R}_Y)$ and $0\leqslant s<t\leqslant T$. Then for $0<s<t\leqslant T$ and continuous $H,\tilde{H}:\mathcal{R}_Y\to\mathbb{R}$, one has

$$\|H(Y_t)-\tilde{H}(Y_s)\|_{L_\infty(\mathbb{P})}\geqslant\|H(Y_t)-\tilde{H}(y_0)\|_{L_\infty(\mathbb{P})}.$$

This follows from the fact that the density $D_{s,t}:\mathcal{R}_Y\times\mathcal{R}_Y\to[0,\infty)$ of $law(Y_s,Y_t)$ with respect to the Lebesgue measure $\lambda\otimes\lambda|_{\mathcal{R}_Y\times\mathcal{R}_Y}$ is the positive and continuous function

$$D_{s,t}(y_1,y_2):=\Gamma_Y(0,s;y_0,y_1)\Gamma_Y(s,t;y_1,y_2).$$

Consequently, if there is a probability measure $\mathbb{Q}\ll\mathbb{P}$ and if for all $t\in[0,T)$ one has that $H(t,\cdot):\mathcal{R}_Y\to\mathbb{R}$ is continuous, $\mathbb{E}^{\mathbb{Q}}|H(t,Y_t)|<\infty$, and $\mathbb{E}^{\mathbb{Q}}(H(t,Y_t)-H(0,y_0))=0$, then $(H(t,Y_t)-H(0,y_0))_{t\in[0,T)}$ is of maximal oscillation with constant 2 according to Lemma 3.4.4.

Example 3.4.6 (Lévy processes)**.** Let $(X_t)_{t\in[0,T]}$, $X_t:\Omega\to\mathbb{R}$, be a Lévy process. By [51, Theorem 61.2] there are $\ell\in\mathbb{R}$ and a closed non-empty $Q\subseteq\mathbb{R}$ such that $0\in Q$, $Q+Q=Q$, and $\operatorname{supp}(X_t)=Q+\ell t$ for $t\in(0,T]$. Define

$$Y_t:=(X_t-\ell t)\mathbb{1}_{\{X_t\in\operatorname{supp}(X_t)\}}$$

so that $Y_t(\Omega)\subseteq Q$ and $\operatorname{supp}(Y_t)=Q$ for all $t\in(0,T]$. Let $0<s<t\leqslant T$ and $H,\tilde{H}:Q\to\mathbb{R}$ be continuous on Q . Then

$$\|H(Y_t)-\tilde{H}(Y_s)\|_{L_\infty(\mathbb{P})}\geqslant\|H(Y_t)-\tilde{H}(0)\|_{L_\infty(\mathbb{P})}.$$

This can be seen from

$$\begin{aligned}\|H(Y_t)-\tilde{H}(Y_s)\|_{L_\infty(\mathbb{P})}&=\|H(Y_s+(Y_t-Y_s))-\tilde{H}(Y_s)\|_{L_\infty(\mathbb{P})}=\sup_{y,y'\in Q}|H(y'+y)-\tilde{H}(y')|\\&\geqslant\sup_{y\in Q}|H(y)-\tilde{H}(0)|=\|H(Y_t)-\tilde{H}(0)\|_{L_\infty(\mathbb{P})}.\end{aligned}$$

Consequently, if there is a probability measure $\mathbb{Q}\ll\mathbb{P}$ and if for all $t\in[0,T)$ one has that $H(t,\cdot):Q\to\mathbb{R}$ is continuous, $\mathbb{E}^{\mathbb{Q}}|H(t,Y_t)|<\infty$, and $\mathbb{E}^{\mathbb{Q}}(H(t,Y_t)-H(0,0))=0$, then $(H(t,Y_t)-H(0,0))_{t\in[0,T)}$ is of maximal oscillation with constant 2 according to Lemma 3.4.4.

Now we connect the notion of oscillation to the behavior of $[\varphi;\tau]$ (by the notation $[\varphi;\tau]$ we mean (3.2.3) with the measure $\Pi(\omega, \mathrm{d}u) = \mathrm{d}u$, $\forall \omega \in \Omega$), where we use extended conditional expectations for non-negative random variables.

Theorem 3.4.7. *Assume $\theta \in (0,1]$, $c_{(3.4.1)} > 0$, and $\varphi \in \mathrm{CL}([0,T))$ such that, a.s.,*

$$\frac{1}{c_{(3.4.1)}}|\varphi_a - Z|^2 \leqslant \mathbb{E}_{\mathcal{F}_a}\left[\frac{1}{b-a}\int_a^b |\varphi_u - Z|^2 \mathrm{d}u\right] \quad a.s. \tag{3.4.1}$$

for all $0 \leqslant a < b < T$ and all $\mathcal{F}_a$-measurable $Z : \Omega \to \mathbb{R}$. Consider the following assertions:

(1) $\inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\underline{\mathrm{Osc}}_t(\varphi) > 0$.

(2) *There is a $c_{(3.4.2)} > 0$ such that for all $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ with $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$ one has*

$$\inf_{\vartheta_{i-1}\in L_0(\mathcal{F}_{t_{i-1}})}\sup_{a\in[t_{k-1},t_k)}\left\|\mathbb{E}_{\mathcal{F}_a}\left[\int_a^T\left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2\mathrm{d}u\right]\right\|_{L_\infty(\mathbb{P})} \geqslant c_{(3.4.2)}^2\|\tau\|_\theta. \tag{3.4.2}$$

(3) *There is a constant $c_{(3.4.3)} > 0$ such that for all time-nets $\tau \in \mathcal{T}_{\det}$ one has*

$$\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} \geqslant c_{(3.4.3)}^2\|\tau\|_\theta. \tag{3.4.3}$$

(4) $\inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\overline{\mathrm{Osc}}_t(\varphi) > 0$.

Then we have (1) $\Rightarrow$ (2) $\Rightarrow$ (3). *If $\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} < \infty$ for all $\tau \in \mathcal{T}_{\det}$ and if*

$$\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} \to 0 \quad as \quad \|\tau\|_1 \to 0,$$

then (3) $\Rightarrow$ (4).

PROOF. See the proof of [29, Theorem 5.7]. $\square$

We remark that condition (3.4.1) above is satisfied if φ is an $L_2(\mathbb{P})$-martingale.

3.5. Approximation in the Brownian setting via gradient estimates

In this section, we extend the equivalence in the case **(c)** in Table 1.1, which is formulated for the geometric Brownian motion and for Lipschitz functionals, to more general frameworks. As a first result, Theorem 3.5.3 shows that the geometric Brownian motion can be replaced by a more general diffusion while keeping the equivalence. However, this extension is still in the Lipschitz framework and gives the impression that this approach is tight to Lipschitz functionals. Our next contribution is to move away from the Lipschitz framework, and this task is done in Theorem 3.5.4.

3.5.1. Setting. Let us recall the setting from C. Geiss and S. Geiss [21]: Let $X = (X_t)_{t\in[0,T]}$ be the solution of the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \hat{\sigma}(X_t)\mathrm{d}W_t + \hat{b}(X_t)\mathrm{d}t, \quad X_0 \equiv x_0 \in \mathbb{R},$$

where $\hat{\sigma} \in C_b^\infty(\mathbb{R})$, $\inf_{x\in\mathbb{R}}\hat{\sigma}(x) > 0$ and $\hat{b} \in C_b^\infty(\mathbb{R})$. Assume that Y solves the SDE

$$\mathrm{d}Y_t = \sigma(Y_t)\mathrm{d}W_t, \quad Y_0 \equiv y_0 \in \mathbb{R},$$

where two settings are used simultaneously:

- Case (C1): $Y = X$ with $\sigma \equiv \hat{\sigma}$, $\hat{b} \equiv 0$, and $\mathcal{R}_Y = \mathbb{R}$.

• Case (C2): $Y = \mathrm{e}^X$ with $\sigma(y) = y\hat{\sigma}(\ln y)$, $\hat{b}(x) = -\frac{1}{2}\hat{\sigma}(x)^2$, and $\mathcal{R}_Y = (0,\infty)$.
In both cases, denote by C_Y the set of all Borel functions $g\colon \mathcal{R}_Y \to \mathbb{R}$ such that

$$\sup_{x\in\mathbb{R}} \mathrm{e}^{-m|x|} \int_{\mathbb{R}} |g(\alpha(x+ty))|^2 \mathrm{e}^{-y^2} \mathrm{d}y < \infty, \quad \forall t > 0$$

for some $m > 0$, where $\alpha(x) = x$ in the case (C1) and $\alpha(x) = \mathrm{e}^x$ in the case (C2). It is clear that all polynomial growth functions belong to C_Y.

For $g \in C_Y$, we define

$$G(t,y) := \mathbb{E}(g(Y_T)|Y_t = y), \quad (t,y) \in [0,T]\times\mathcal{R}_Y.$$

Lemma 3.5.1. *If $g \in C_Y$, then Assumption 3.3.1 is satisfied for*

$$\sigma := (\sigma(Y_t))_{t\in[0,T]}, \ M := \left(\int_0^t \left(\sigma^2\partial_y^2 G\right)(u,Y_u)\mathrm{d}W_u\right)_{t\in[0,T)}, \ \varphi := (\partial_y G(t,Y_t))_{t\in[0,T)}.$$

PROOF. The assertion follows from [29, Lemma 6.8] and S. Geiss [24, Corollary 3.3]. □

For $g \in C_Y$ and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$, we define the error $E(g;\tau) = (E_t(g;\tau))_{t\in[0,T]}$ by

$$E_t(g;\tau) := \int_0^t \varphi_u \mathrm{d}Y_u - \sum_{i=1}^n \varphi_{t_{i-1}}(Y_{t_i\wedge t} - Y_{t_{i-1}\wedge t}), \quad t \in [0,T].$$

The following result verifies that gradient processes $(\partial_y G(t,Y_t))_{t\in[0,T)}$ have a large oscillation. Its proof can be found in [29, the proof of Theorem 6.3].

Theorem 3.5.2. *For $g \in C_Y$, the process $(\partial_y G(t,Y_t))_{t\in[0,T)}$ is of maximal oscillation with constant 2 in the sense of Definition 3.4.1.*

3.5.2. Approximation and gradient estimates. In this section, we use the processes σ, M, and φ as given in Lemma 3.5.1.

Now we discuss cases in which we get equivalences by choosing the weight Φ accordingly. For $\theta = 1$ we get a characterization in terms of Lipschitz functions that extends [25, Theorem 8].

Theorem 3.5.3 (Lipschitz case)**.** *For $g \in C_Y$ and $\Phi = \sigma$, the following assertions are equivalent:*

(1) *There is a constant $c > 0$ such that $\|E(g;\tau)\|_{\mathrm{BMO}_2^\Phi([0,T))} \leqslant c\sqrt{\|\tau\|_1}$ for all $\tau \in \mathcal{T}_{\det}$.*
(2) *There is a Lipschitz function $\tilde{g}\colon \mathcal{R}_Y \to \mathbb{R}$ such that $g = \tilde{g}$ a.e. on $\mathcal{R}_Y$ with respect to the Lebesgue measure.*

PROOF. See the proof of [29, Theorem 6.4]. □

In the case $\theta \in (0,1)$, which potentially includes the Hölder setting, we obtain an equivalence in terms of the Riemann–Liouville type integral of the gradient process.

Theorem 3.5.4 (non-Lipschitz case)**.** *Let $(\theta,q) \in (0,1)\times[2,\infty)$ and $\Phi = (\sigma_t\Psi_t)_{t\in[0,T)}$, where $\Psi \in \mathrm{CL}^+([0,T))$ is pathwise non-decreasing. If $g \in C_Y$ and if there is a constant $c > 0$ such that*

$$|\varphi_t| \leqslant c(T-t)^{-\frac{1}{2}}\Psi_t \quad a.s, \ \forall t \in [0,T), \tag{3.5.1}$$

then the following assertions are equivalent:

(1) *There is a constant $c > 0$ such that $\|E(g;\tau)\|_{\mathrm{BMO}_2^\Phi([0,T))} \leqslant c\sqrt{\|\tau\|_\theta}$ for any $\tau \in \mathcal{T}_{\det}$.*

(2) *One has for* $Z := \varphi\sigma$ *that* $\mathcal{I}^{\frac{1-\theta}{2}} Z - Z_0 \in \mathrm{BMO}_2^\Phi([0,T))$ *and there exists a constant* $c > 0$ *such that*

$$|\varphi_t| \leqslant c(T-t)^{\frac{\theta-1}{2}} \Psi_t \quad \textit{a.s.}, \ \forall t \in [0,T).$$

If the conditions (1) *and* (2) *are satisfied and* $\Phi \in \mathcal{SM}_q([0,T))$, *then* $\mathcal{I}_T^{\frac{1-\theta}{2}} Z := \lim_{t\uparrow T} \mathcal{I}_t^{\frac{1-\theta}{2}} Z$ *exists in* $L_q(\mathbb{P})$ *and a.s.*

PROOF. See the proof of [29, Theorem 6.5]. □

By an argument using conditional Itô's isometry, it holds that

$$\|E(g;\tau)\|^2_{\mathrm{BMO}_2^\Phi([0,T))} = \|[\varphi;\tau]^\sigma\|_{\mathrm{BMO}_1^{\Phi^2}([0,T))},$$

where $\Pi(\omega, \mathrm{d}u) = \sigma_u^2(\omega)\mathrm{d}u$ and $[\varphi;\tau]^\pi$, which is equal to $[\varphi;\tau]^\sigma$, is given in (3.2.3). In item (2) of Theorem 3.5.4, if the condition $\mathcal{I}^{\frac{1-\theta}{2}} Z - Z_0 \in \mathrm{BMO}_2^\Phi([0,T))$ is replaced by $\mathcal{I}^{\frac{1-\theta}{2}} M \in \mathrm{BMO}_2^\Phi([0,T))$, then the obtained statement is equivalent to item (1) without requiring (3.5.1). This observation can be verified by Lemma 3.5.1 and Theorem 3.3.3. Notice that M has continuous paths so that one can use $\mathrm{BMO}_2^\Phi([0,T))$ instead of $\mathrm{bmo}_2^\Phi([0,T))$. However, since we want to characterize the approximation statement in item (1) by means of the gradient process Z, the a priori estimate (3.5.1) enables to switch between Z and M in item (2).

Theorems 3.5.3 and 3.5.4 above are versions of the equivalence Theorem 3.1.1((1)⇔(2)) in the weighted BMO-setting where the approximation error estimates can be described via the Riemann–Liouville type operator. The next result provides a situation in which Theorem 3.5.4(2) is satisfied. In particular, Theorem 3.5.5(3) proves the implication (3) ⇒ (2) of Theorem 3.1.1 in the BMO-context.

Theorem 3.5.5. *Let* $(\theta, q) \in [0,1] \times [2,\infty)$ *and* $\Phi = (\sigma_t \Psi_t)_{t\in[0,T)}$ *with*

$$\Psi_t := \sup_{s\in[0,t]}(\sigma_s^{\theta-1}).$$

Then the following assertions hold:

(1) (*Weight regularity*) $\Phi \in \mathcal{SM}_q([0,T))$.

(2) (*Gradient estimates*) *If* $g \in \mathrm{Höl}_\theta(\mathbb{R})$, *then there is a constant* $c > 0$ *such that*

$$|\varphi_t| \leqslant c(T-t)^{\frac{\theta-1}{2}} \Psi_t \quad \textit{a.s.}, \ \forall t \in [0,T).$$

(3) *If* $g \in \mathrm{Höl}^0_{\theta,2}(\mathbb{R})$ *for some* $\theta \in (0,1)$, *then* $\mathcal{I}^{\frac{1-\theta}{2}} Z - Z_0 \in \mathrm{BMO}_q^\Phi([0,T))$.

PROOF. See the proof of [29, Theorem 6.6]. □

CHAPTER 4

Approximation in models with jumps: Jump adjusted method

Convention. From now until the end of this thesis, we only consider the time interval $[0,T]$, and the quantities $\|\cdot\|_{\mathrm{BMO}_p^{\Phi}([0,T])}$ and $\|\cdot\|_{\mathcal{SM}_p([0,T])}$ computed under $\mathbb{P}$ will be denoted respectively by $\|\cdot\|_{\mathrm{BMO}_p^{\Phi}(\mathbb{P})}$ and $\|\cdot\|_{\mathcal{SM}_p(\mathbb{P})}$ to indicate explicitly the reference measure.

4.1. Introduction

This chapter is concerned with discrete-time approximation problems for stochastic integrals and studies the error process $E = (E_t)_{t\in[0,T]}$ defined by

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

where $T \in (0,\infty)$ 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 is an approximation scheme for the stochastic integral.

The error represented by the difference between a stochastic integral and its discretisation has been extensively analysed 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 to C. Geiss and S. Geiss [21], Gobet and Temam [31], Zhang [62], where the error along with its convergence rates was examined. The weak convergence of the error was treated in S. Geiss and Toivola [28], Gobet and Temam [31]. When the driving process is a continuous semimartingale, the convergence in the L_2-sense was studied by Fukasawa [20], and in the almost sure sense it was considered by Gobet and Landon [30].

In this chapter, we allow the semimartingale to jump since many important processes used in financial modelling are not continuous (see Cont and Tankov [13], Schoutens [53]), 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 Schweizer [56] and the references therein), and it will be revisited in Chapters 6 and 7. The present chapter focuses on the first type of hedging error only. The discretisation error was studied within Lévy models in the weak convergence sense by Tankov and Voltchkova [59], in the L_2-sense by Brodén and Tankov [9], C. Geiss, S. Geiss and Laukkarinen [22], and for a more general jump model under the L_2-setting by Rosenbaum and Tankov [48].

In Section 3.5, the Riemann approximation errors with *deterministic* time-nets measured in weighted BMO are upper bounded by certain mesh sizes of the time-nets. In those results the continuity of the driving processes is crucial to obtain estimates. However, if the driving process has jumps, then such results might fail as asserted in the following example. Namely, it shows that the Riemann approximation error with respect to *deterministic* time-nets does in general *not* converge to zero if measured in weighted BMO.

Example 4.1.1 ([60], Example 3.7)**.** Let $\tilde{J} := (J_t - rt)_{t\in[0,T]}$ be a compensated Poisson process with intensity $r > 0$. Let $f : (0,T] \times \mathbb{N} \to \mathbb{R}$ be a Borel function with

$$\|f\|_\infty := \sup_{(t,k)\in(0,T]\times\mathbb{N}} |f(t,k)| < \infty \quad \text{and} \quad \varepsilon := \inf_{t\in(0,T]} |f(t,0)| > 0.$$

Assume that

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

Denote $\rho_1 := \inf\{t > 0 : \Delta J_t = 1\} \wedge T$ and $\rho_2 := \inf\{t > \rho_1 : \Delta J_t = 1\} \wedge T$. Let $\vartheta_0 \in \mathbb{R}$ and define

$$\vartheta_t = \vartheta_0 + \int_{(0,t\wedge\rho_2]} f(s, J_{s-})\mathrm{d}\tilde{J}_s, \quad t \in (0,T].$$

It is not difficult to check that $\vartheta = (\vartheta_t)_{t\in[0,T]}$ is a càdlàg martingale with $\|\vartheta_T\|_{L_\infty(\mathbb{P})} < \infty$.

Let $E^{\mathrm{Rm}}(\vartheta,\tau) = (E^{\mathrm{Rm}}_t(\vartheta,\tau))_{t\in[0,T]}$ be the error resulting from the Riemann approximation of $\int_0^T \vartheta_{t-}\mathrm{d}\tilde{J}_t$ with the deterministic time-net $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$. Namely,

$$E^{\mathrm{Rm}}_t(\vartheta,\tau) := \int_0^t \vartheta_{u-}\mathrm{d}\tilde{J}_u - \sum_{i=1}^n \vartheta_{t_{i-1}-}(\tilde{J}_{t_i\wedge t} - \tilde{J}_{t_{i-1}\wedge t}), \quad t \in [0,T].$$

On the set $\{0 < \rho_1 < \rho_2 < t_1\}$ we have

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

Since $\mathbb{P}(0 < \rho_1 < \rho_2 < t_1) > 0$, it implies that $\inf_{\tau\in\mathcal{T}_{\mathrm{det}}} \|\Delta E^{\mathrm{Rm}}_{\rho_2}(\vartheta,\tau)\|_{L_\infty(\mathbb{P})} \geqslant \delta$. Then it follows from Proposition 2.2.2 that

$$\inf_{\tau\in\mathcal{T}_{\mathrm{det}}} \|E^{\mathrm{Rm}}(\vartheta,\tau)\|_{\mathrm{BMO}_p(\mathbb{P})} > 0, \quad \forall p \in (0,\infty).$$

Therefore, we need to look for another approximation scheme in jump models to exploit benefits of weighted BMO. Before we do that, let us give the family of stochastic integrals used for approximation.

4.1.1. Stochastic integrals to be approximated. The stochastic integrals we are going to approximate are of the form

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

where the assumptions for S and ϑ are as follows.

(1) $S \in \mathrm{CL}([0,T])$ satisfies the SDE[1]

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

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

[1] See, for example, Protter [47, Ch.V, Sec.3], for the existence and uniqueness of S.

(2) $Z \in \mathrm{CL}([0,T])$ is a square integrable semimartingale with the representation

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

where $Z_0 \in \mathbb{R}$, V is a progressively measurable process, Z^{c} is a pathwise continuous square integrable 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 for Z are the following:

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

$$\pi_Z(\omega, \mathrm{d}t, \mathrm{d}z) = \nu_t(\omega, \mathrm{d}z)\mathrm{d}t, \tag{4.1.1}$$

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

(b) There is a progressively measurable process C such that $\langle Z^{\mathrm{c}}, Z^{\mathrm{c}} \rangle = \int_0^{\cdot} C_u^2 \mathrm{d}u$.

(c) The processes V and K, where $K_t := (C_t^2 + \int_{\mathbb{R}} z^2 \nu_t(\mathrm{d}z))^{1/2}$, satisfy that

$$\| \|V\|_{L_2([0,T],\lambda)} \|_{L_\infty(\mathbb{P})} < \infty, \quad \|K\|_{L_\infty(\Omega \times [0,T], \mathbb{P} \otimes \lambda)} < \infty.$$

(3) ϑ belongs to the family Σ_S^{adm} of *admissible integrands*, where

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

For $t \in [0,T]$, it follows from (4.1.1) 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. Thus, it is natural to assume $\Delta\vartheta_t = 0$ a.s. for admissible integrands.

In particular, when Z is a square integrable Lévy process (we will consider this case in Sections 6 and 7), then the assumptions for Z in item (2) are satisfied in the view of the Lévy–Itô decomposition of Z (see (5.1.1)).

4.2. Approximation scheme with jump adjustment

As we have already seen in Example 4.1.1 that, in models with jumps, deterministic time-nets are not suitable for the Riemann approximation measured in weighted BMO because of the possibly large jumps of the driving process. To overcome this problem, we exploit an idea of Dereich and Heidenreich [18] and propose an approximation scheme based on a correction of the 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. Because of the assumptions imposed on S in Subsection 4.1.1, one has $\sigma(S_-) > 0$ and

$$\Delta S = \sigma(S_-)\Delta Z$$

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

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

[3] See Jacod and Shiryaev [37, Ch.II, Sec.1] for more details.

of S rather than of Z. Therefore, we define the random times $\rho(\varepsilon,\kappa) = (\rho_i(\varepsilon,\kappa))_{i\geqslant 0}$ based on tracking the jumps of S as follows (recall that $\inf\emptyset := \infty$).

Definition 4.2.1. For $\varepsilon > 0$ and $\kappa \geqslant 0$, let $\rho_0(\varepsilon,\kappa) := 0$ and

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

The quantity $\varepsilon(T-t)^\kappa$ above is the level at time t where we decide which jumps of S are (relatively) large, and moreover, this level shrinks when t approaches the terminal time T in the case $\kappa > 0$. 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 $\kappa = 0$, the control parameter ε can be interpreted as the *jump size threshold*.

Definition 4.2.2 (*Jump adjusted approximation*)**.** Let $\varepsilon > 0$, $\kappa \in [0, \frac{1}{2})$, and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$.

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

(2) The discretised strategy ϑ^τ, the Riemann appoximation A^{Rm}, the approximation with corrections A^{adj} and the corresponding error E^{adj} are defined as follows: For $t \in [0,T]$,

$$\vartheta^\tau_t := \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(t), \quad A^{\mathrm{Rm}}_t(\vartheta,\tau) := \sum_{i=1}^n \vartheta_{t_{i-1}-}(S_{t_i\wedge t} - S_{t_{i-1}\wedge t}),$$

$$A^{\mathrm{adj}}_t(\vartheta,\tau|\varepsilon,\kappa) := A^{\mathrm{Rm}}_t(\vartheta,\tau) + \sum_{\rho_i(\varepsilon,\kappa)\in[0,t]\cap[0,T)} \left(\vartheta_{\rho_i(\varepsilon,\kappa)-} - \vartheta^\tau_{\rho_i(\varepsilon,\kappa)}\right)\Delta S_{\rho_i(\varepsilon,\kappa)}, \tag{4.2.1}$$

$$E^{\mathrm{adj}}_t(\vartheta,\tau|\varepsilon,\kappa) := \int_0^t \vartheta_{u-}\mathrm{d}S_u - A^{\mathrm{adj}}_t(\vartheta,\tau|\varepsilon,\kappa).$$

As verified in [60, Subsection 5.2], each $\rho_i(\varepsilon,\kappa)$ is a stopping time. Moreover, in our setting the sum on the right-hand side of (4.2.1) is a finite sum a.s. as a consequence of [60, Proposition 5.3]. 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, the strategy ϑ does not necessarily have the left-limit at T, and secondly, since $\Delta S_T = 0$ a.s. as mentioned in Subsection 4.1.1, any value of the form $a\Delta S_T$ ($a \in \mathbb{R}$) added to the correction term does not affect the approximation in our context.

4.3. Approximation with corrections in weighted BMO

We now use the jump adjusted approximation and apply the results in Section 3.2 to obtain the main results in this part. First, for reader's convenience, let us adapt Assumption 3.2.1 to this section. Since we are only interested in *upper estimates* for the approximation error and the random measure Π we are going to choose is

$$\Pi(\omega,\mathrm{d}t) := \sigma(S_t(\omega))^2\mathrm{d}t,$$

Assumption 3.2.1 becomes the following form:

Assumption 4.3.1. For $\vartheta \in \Sigma^{\mathrm{adm}}_S$, we assume that there exists a random measure

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

such that

$$\Upsilon(\omega,(0,t]) < \infty, \quad \forall(\omega,t) \in \Omega\times(0,T),$$

and such that there exists a constant $c > 0$ such that for any $0 \leqslant a < b < T$,

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,b]} |\vartheta_t - \vartheta_a|^2 \sigma(S_t)^2 \mathrm{d}t\right] \leqslant c^2 \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,b]} (b-t)\Upsilon(\cdot, \mathrm{d}t)\right] \quad \text{a.s.} \tag{4.3.1}$$

We now provide the formula for Υ which is used in the (exponential) Lévy setting later.

Example 4.3.2 ([60], Example 3.2)**.** Assume that $M := \vartheta\sigma(S) \in \mathrm{CL}([0,T))$ is an $L_2(\mathbb{P})$-martingale. Then (4.3.1) is satisfied for the random measure Υ defined by

$$\Upsilon(\omega, \mathrm{d}t) := \mathrm{d}\langle M, M\rangle_t(\omega) + |\sigma|^2_{\mathrm{Lip}} |M_t(\omega)|^2 \mathrm{d}t,$$

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

In view of Theorem 3.2.2, the following assumption enables approximation results.

Assumption 4.3.3. Let $\theta \in (0,1]$. Assume that Assumption 4.3.1 is satisfied and there is an a.s. non-decreasing process $\Theta \in \mathrm{CL}^+([0,T])$ such that the following two conditions hold:

(1) (*Growth condition*) There is a constant $c > 0$ such that

$$|\vartheta_a| \leqslant c(T-a)^{\frac{\theta-1}{2}} \Theta_a \quad \text{a.s., } \forall a \in [0,T). \tag{4.3.2}$$

(2) (*Curvature condition*) For $\Phi := \Theta\sigma(S)$, there is a constant $c > 0$ such that

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta} \Upsilon(\cdot, \mathrm{d}t)\right] \leqslant c^2 \Phi_a^2 \quad \text{a.s., } \forall a \in [0,T). \tag{4.3.3}$$

The parameter θ in Assumption 4.3.3 describes the growth (pathwise and relative to Θ) of ϑ when the time variable a approaches the terminal time T. For the Black–Scholes model with the delta-hedging strategy ϑ, the parameter θ can be interpreted as the fractional smoothness of the payoff in the sense of [21, 27].

To formulate main results, we need to modify the weight processes. For $\Phi \in \mathrm{CL}^+([0,T])$ and $t \in [0,T]$, we define

$$\overline{\Phi}_t := \Phi_t + \sup_{s \in [0,t]} |\Delta \Phi_s|.$$

The reason to consider $\overline{\Phi}$ is that in the proof of main results we will end up with Φ_- which is not càdlàg, and therefore is not a candidate for a weight process. For $\overline{\Phi}$, it is straightforward to check that $\overline{\Phi} \in \mathrm{CL}^+([0,T])$ with $\Phi \vee \Phi_- \leqslant \overline{\Phi}$, and $\Phi \equiv \overline{\Phi}$ if and only if Φ is continuous. Moreover, $\Phi \in \mathcal{SM}_p(\mathbb{P})$ implies $\overline{\Phi} \in \mathcal{SM}_p(\mathbb{P})$ (see [60, Propostion 7.1]).

Theorem 4.3.4. *Assume that Assumption 4.3.3 holds for some $\theta \in (0,1]$ and $\Phi \in \mathcal{SM}_2(\mathbb{P})$.*

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

$$\left\| (\omega,t) \mapsto \textstyle\int_{|z|\leqslant 1} |z|^\alpha \nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T], \mathbb{P}\otimes\lambda)} < \infty, \tag{4.3.4}$$

then a constant $c_{(4.3.5)} > 0$ exists such that for all $\tau \in \mathcal{T}_{\det}$, $\varepsilon > 0$,

$$\left\| E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(4.3.5)} \max\left\{ \varepsilon^{1-\alpha}\sqrt{\|\tau\|_\theta},\ \sqrt{\|\tau\|_\theta},\ \varepsilon \right\}. \tag{4.3.5}$$

(2) *If there is a constant $c_{(4.3.6)} > 0$ such that for $\mathbb{P}\otimes\lambda$-a.e. $(\omega,t) \in \Omega\times[0,T]$,*

$$\sup_{r>0} \left| \textstyle\int_{|z|>r} z\nu_t(\omega, \mathrm{d}z) \right| \leqslant c_{(4.3.6)}, \tag{4.3.6}$$

then a constant $c_{(4.3.7)} > 0$ exists such that for all $\tau \in \mathcal{T}_{\mathrm{det}}$, $\varepsilon > 0$,

$$\left\| E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(4.3.7)} \max\left\{\sqrt{\|\tau\|_\theta},\, \varepsilon\right\}. \tag{4.3.7}$$

PROOF. See the proof of [60, Theorem 3.10]. □

Minimizing the right-hand side of (4.3.5) (resp. (4.3.7)) over $\varepsilon > 0$ leads us to the selection $\varepsilon = \sqrt[2\alpha]{\|\tau\|_\theta}$ (resp. $\varepsilon = \sqrt{\|\tau\|_\theta}$). Then we have the following:

Corollary 4.3.5. *Assume that Assumption 4.3.3 holds for some $\theta \in (0,1]$ and $\Phi \in \mathcal{SM}_2(\mathbb{P})$.*

(1) *If* (4.3.4) *is satisfied for some $\alpha \in [1,2]$, then*

$$\left\| E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \sqrt[2\alpha]{\|\tau\|_\theta}, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant T^{\frac{\theta}{2}(1-\frac{1}{\alpha})} c_{(4.3.5)} \sqrt[2\alpha]{\|\tau\|_\theta}.$$

(2) *If* (4.3.6) *is satisfied, then*

$$\left\| E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \sqrt{\|\tau\|_\theta}, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(4.3.7)} \sqrt{\|\tau\|_\theta}.$$

The time-net used in Theorem 4.3.4 is $\tau \sqcup \rho(\varepsilon, \frac{1-\theta}{2})$. Due to the randomness, a simple way to quantify the cardinality of this combined time-net is to compute its expected cardinality, i.e. $\mathbb{E}\left[\#\tau \sqcup \rho(\varepsilon, \frac{1-\theta}{2})\right]$ (see, e.g., Fukasawa [19]). We provide in the next result an estimate for certain moments of the cardinality. Since we aim to apply Proposition 2.2.4(3) later, changes of the underlying measure are also taken into account.

Proposition 4.3.6. *Let $q \in [1,2]$, $r \in [2,\infty]$ with $\frac{q}{2} + \frac{1}{r} = 1$. Assume that $\mathbb{Q}$ is a probability measure absolutely continuous with respect to $\mathbb{P}$ and $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L_r(\mathbb{P})$. For $\theta \in (0,1]$ and $(\varepsilon_n)_{n\geqslant 1} \subset (0,\infty)$ with $\inf_{n\geqslant 1} \sqrt{n}\varepsilon_n > 0$, there is a constant $c_{(4.3.8)} > 0$ such that for any $n \geqslant 1$, $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$,*

$$\left\| \#\tau_n \sqcup \rho\left(\varepsilon_n, \tfrac{1-\theta}{2}\right) \right\|_{L_q(\mathbb{Q})} \sim_{c_{(4.3.8)}} n. \tag{4.3.8}$$

PROOF. See the proof of [60, Proposition 3.13]. □

Using the adapted time-nets τ_n^θ given in (2.5.1), we have the following:

Theorem 4.3.7. *Assume that Assumption 4.3.3 holds for some $\theta \in (0,1]$ and $\Phi \in \mathcal{SM}_2(\mathbb{P})$.*

(1) *If* (4.3.4) *is satisfied for some $\alpha \in [1,2]$, then*

$$\sup_{n\geqslant 1} n^{\frac{1}{2\alpha}} \left\| E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \middle| n^{-\frac{1}{2\alpha}}, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty.$$

(2) *If* (4.3.6) *is satisfied, then*

$$\sup_{n\geqslant 1} n^{\frac{1}{2}} \left\| E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \middle| n^{-\frac{1}{2}}, \tfrac{1-\theta}{2}\right) \right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty.$$

(3) *If in addition $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in (2,\infty)$, then the conclusions of items* (1)–(2) *hold for the $L_p(\mathbb{P})$-norm in place of the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$-norm.*

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

PROOF. Items (1), (2) follow directly from combining Theorem 4.3.4 with (2.5.2). Items (3), (4) are due to Proposition 2.2.4 and the fact that $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$ (see [60, Proposition 7.1]). □

In Theorem 4.3.7, applying Proposition 4.3.6 with $q = 2, r = \infty$ and $\mathbb{Q} = \mathbb{P}$ implies that the $L_2(\mathbb{P})$-norm of the cardinality of the combined time-net used in the corresponding approximation schemes is comparable to n up to a multiplicative constant. In item (4), if $s \in [2, \infty)$, then applying Proposition 4.3.6 with $q = 1, r = 2$ yields that the $L_1(\mathbb{Q})$-norm of the combined time-net is comparable to n.

We derive from Theorem 4.3.7(2) the convergence rate of order $n^{-1/2}$ which is *asymptotically optimal* in general (e.g., see C. Geiss, S. Geiss and Laukkarinen [22, Theorem 5] in the Lévy case), while this rate is achieved in item (1) for $\alpha = 1$. Obviously, the convergence rate in item (1) depends on the small jumps intensity of the underlying process Z, which is characterised by α. If we define

$$\beta^Z := \inf\left\{\alpha \in [0,2] : \left\|(\omega, t) \mapsto \int_{|z|\leqslant 1} |z|^\alpha \nu_t(\omega, \mathrm{d}z)\right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty\right\},$$

then it follows from Theorem 4.3.7(1) that

$$\inf\left\{\alpha \in [1,2] : \sup_{n\geqslant 1} n^{\frac{1}{2\alpha}} \left\| E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \,\middle|\, n^{-\frac{1}{2\alpha}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty\right\} \leqslant 1 \vee \beta^Z.$$

Note that when Z is a Lévy process, then β^Z is the *Blumenthal–Getoor index* of Z (see Blumenthal and Getoor [8]).

CHAPTER 5

Gradient type estimates in the Lévy–Itô space

Before proceeding to apply the results in Chapter 4 for the (exponential) Lévy setting, in view of (4.3.2), we need some gradient type estimates in the Lévy–Itô space.

5.1. Lévy process and Itô's chaos expansion

5.1.1. Lévy process. Let $T > 0$ be a fixed finite time horizon. Let $X = (X_t)_{t\in[0,T]}$ be a real-valued Lévy process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $X_0 = 0$, X has independent and stationary increments and X has càdlàg paths. Assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of X. According to the *Lévy–Khintchine formula* (see, e.g., Sato [51, Theorem 8.1]), the *characteristic exponent* ψ of X, which is defined by

$$\mathbb{E}\mathrm{e}^{\mathrm{i}uX_t} = \mathrm{e}^{-t\psi(u)}, \quad u \in \mathbb{R}, t \geqslant 0,$$

is of the form

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

Here, $\gamma \in \mathbb{R}$, while $\sigma \geqslant 0$ is the coefficient of the Brownian component, and $\nu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ is a Lévy measure (i.e. $\nu(\{0\}) := 0$ and $\int_{\mathbb{R}}(x^2 \wedge 1)\nu(\mathrm{d}x) < \infty$). The triplet (γ, σ, ν) is also called the characteristics of X. To indicate explicitly the characteristics of X under $\mathbb{P}$, we write

$$(X|\mathbb{P}) \sim (\gamma, \sigma, \nu) \quad \text{or} \quad (X|\mathbb{P}) \sim \psi.$$

It is known that paths of X can be described by the following *Lévy–Itô decomposition*

$$X_t = \gamma t + \sigma W_t + \int_0^t \int_{|x|\leqslant 1} x\widetilde{N}(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_{|x|>1} xN(\mathrm{d}s, \mathrm{d}x), \quad t \geqslant 0, \tag{5.1.1}$$

where W is a standard Brownian motion, N is the Poisson random measure of X, i.e. $N((s,t] \times B) := \#\{u \in (s,t] : \Delta X_u \in B\}$ for $0 \leqslant s < t$, $B \in \mathcal{B}(\mathbb{R}_0)$, and $\widetilde{N}(\mathrm{d}s, \mathrm{d}x) := N(\mathrm{d}s, \mathrm{d}x) - \mathrm{d}s\nu(\mathrm{d}x)$.

5.1.2. Itô's chaos expansion. We present briefly the Malliavin calculus for Lévy processes by means of Itô's chaos expansion which is the main tool to establish an explicit form for the gradient process (Proposition 5.2.2 and Theorem 6.2.3) and to prove the martingale representation formula (Proposition 7.2.2) later. For further details, we refer to [57, 45, 46, 1].

We assume the Lévy process X as in Subsection 5.1.1 and assume that $\mathcal{F} = \mathcal{F}_T$. Define the σ-finite measures μ on $\mathcal{B}(\mathbb{R})$ and $\mathbb{m}$ on $\mathcal{B}([0,T] \times \mathbb{R})$ by setting

$$\mu(\mathrm{d}x) := \sigma^2\delta_0(\mathrm{d}x) + x^2\nu(\mathrm{d}x) \quad \text{and} \quad \mathbb{m} := \lambda \otimes \mu,$$

where δ_0 is the Dirac measure at zero. For $B \in \mathcal{B}([0,T] \times \mathbb{R})$ with $\mathbb{m}(B) < \infty$, the random measure M is defined by

$$M(B) := \sigma \int_{\{t\in[0,T]:(t,0)\in B\}} \mathrm{d}W_t + L_2(\mathbb{P})\text{-}\lim_{n\to\infty} \int_{B\cap([0,T]\times\{\frac{1}{n}<|x|<n\})} x\widetilde{N}(\mathrm{d}t, \mathrm{d}x).$$

Set $L_2(\mu^0) = L_2(\mathbbm{m}^0) := \mathbb{R}$. For $n \geqslant 1$, we denote

$$L_2(\mu^{\otimes n}) := L_2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu^{\otimes n}),$$
$$L_2(\mathbbm{m}^{\otimes n}) := L_2(([0,T]\times\mathbb{R})^n, \mathcal{B}(([0,T]\times\mathbb{R})^n), \mathbbm{m}^{\otimes n}).$$

The multiple integral $I_n : L_2(\mathbbm{m}^{\otimes n}) \to L_2(\mathbb{P})$ is defined in the sense of Itô [35] by using an approximation argument, where it is given for simple functions as follows: For

$$\xi_n^m := \sum_{k=1}^m a_k \mathbb{1}_{B_1^k\times\cdots\times B_n^k},$$

where $a_k \in \mathbb{R}$, $B_i^k \in \mathcal{B}([0,T]\times\mathbb{R})$ with $\mathbbm{m}(B_i^k) < \infty$ and $B_i^k \cap B_j^k = \emptyset$ for $k = 1,\dots,m$, $i, j = 1,\dots,n$, $i \neq j$ and $m \geqslant 1$, we define

$$I_n(\xi_n^m) := \sum_{k=1}^m a_k M(B_1^k)\cdots M(B_n^k).$$

Then [35, Theorem 2] asserts the following Itô chaos expansion

$$L_2(\mathbb{P}) = \bigoplus_{n=0}^{\infty}\{I_n(\xi_n) : \xi_n \in L_2(\mathbbm{m}^{\otimes n})\},$$

where $I_0(\xi_0) := \xi_0 \in \mathbb{R}$. For $n \geqslant 1$, the symmetrization $\tilde{\xi}_n$ of a $\xi_n \in L_2(\mathbbm{m}^{\otimes n})$ is

$$\tilde{\xi}_n((t_1,x_1),\dots,(t_n,x_n)) := \frac{1}{n!}\sum_{\pi}\xi_n((t_{\pi(1)},x_{\pi(1)}),\dots,(t_{\pi(n)},x_{\pi(n)})),$$

where the sum is taken over all permutations π of $\{1,\dots,n\}$, so that $I_n(\xi_n) = I_n(\tilde{\xi}_n)$ a.s. The Itô chaos decomposition verifies that $\xi \in L_2(\mathbb{P})$ if and only if there are $\xi_n \in L_2(\mathbbm{m}^{\otimes n})$ such that $\xi = \sum_{n=0}^{\infty} I_n(\xi_n)$ a.s., and this expansion is unique if every ξ_n is symmetric, i.e. $\xi_n = \tilde{\xi}_n$. Furthermore, $\|\xi\|^2_{L_2(\mathbb{P})} = \sum_{n=0}^{\infty} n!\|\tilde{\xi}_n\|^2_{L_2(\mathbbm{m}^{\otimes n})}$.

Definition 5.1.1. Let $\mathbb{D}_{1,2}$ be the Malliavin–Sobolev space of all $\xi = \sum_{n=0}^{\infty} I_n(\xi_n) \in L_2(\mathbb{P})$ with

$$\|\xi\|^2_{\mathbb{D}_{1,2}} := \sum_{n=0}^{\infty}(n+1)!\|\tilde{\xi}_n\|^2_{L_2(\mathbbm{m}^{\otimes n})} < \infty.$$

The *Malliavin derivative operator* $D : \mathbb{D}_{1,2} \to L_2(\mathbb{P}\otimes\mathbbm{m})$, where $L_2(\mathbb{P}\otimes\mathbbm{m}) := L_2(\Omega\times[0,T]\times\mathbb{R}, \mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}), \mathbb{P}\otimes\mathbbm{m})$, is defined for $\xi = \sum_{n=0}^{\infty} I_n(\xi_n) \in \mathbb{D}_{1,2}$ by

$$D_{t,x}\xi := \sum_{n=1}^{\infty} n I_{n-1}(\tilde{\xi}_n((t,x),\cdot)), \quad (\omega,t,x) \in \Omega\times[0,T]\times\mathbb{R}.$$

Proposition 5.1.2 below was obtained by Laukkarinen [39, Corollary 3.1 in the second article of this thesis] and it provides an equivalent condition such that a functional of X_t belongs to $\mathbb{D}_{1,2}$. We refer to Malliavin, Airault, Kay and Letac [43, Proposition V.2.3.1] when X is a Brownian motion, and refer to C. Geiss and Steinicke [23, Lemma 3.2] when X has no Brownian component.

Proposition 5.1.2 ([39]). *Let $t \in (0,T]$ and a Borel function $f : \mathbb{R} \to \mathbb{R}$ with $f(X_t) \in L_2(\mathbb{P})$. Then $f(X_t) \in \mathbb{D}_{1,2}$ if and only if the following two assertions hold:*

(a) *when* $\sigma > 0$, f *has a weak derivative*[1] f'_w *on* $\mathbb{R}$ *with* $f'_w(X_t) \in L_2(\mathbb{P})$,

(b) *the map* $(s,x) \mapsto \frac{f(X_t+x)-f(X_t)}{x} \mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x)$ *belongs to* $L_2(\mathbb{P}\otimes\mathbb{m})$.

Furthermore, if $f(X_t) \in \mathbb{D}_{1,2}$, *then for* $\mathbb{P}\otimes\mathbb{m}$*-a.e.* $(\omega,s,x) \in \Omega\times[0,T]\times\mathbb{R}$ *one has*

$$D_{s,x} f(X_t) = f'_w(X_t)\mathbb{1}_{[0,t]\times\{0\}}(s,x) + \frac{f(X_t+x)-f(X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x),$$

where we set, by convention, $f'_w := 0$ *when* $\sigma = 0$.

The first item in the following result verifies that the kernels in the chaos expansion of $f(X_T)$ do not depend on time variables, and this is a key observation for us to establish some gradient processes in Proposition 5.2.2 and Theorem 6.2.3.

Lemma 5.1.3. *If a Borel function* $f:\mathbb{R}\to\mathbb{R}$ *satisfies* $f(X_T) \in L_2(\mathbb{P})$, *then there exist symmetric* $\tilde{f}_n \in L_2(\mu^{\otimes n})$ *such that the following holds:*

(1) *One has* $f(X_T) = \mathbb{E} f(X_T) + \sum_{n=1}^{\infty} I_n(\tilde{f}_n \mathbb{1}_{(0,T]}^{\otimes n})$ *a.s.*

(2) *For any* $t \in [0,T)$ *one has* $\mathbb{E}_{\mathcal{F}_t}[f(X_T)] = \mathbb{E} f(X_T) + \sum_{n=1}^{\infty} I_n(\tilde{f}_n \mathbb{1}_{(0,t]}^{\otimes n})$ *a.s. Consequently,* $\mathbb{E}_{\mathcal{F}_t}[f(X_T)] \in \mathbb{D}_{1,2}$ *for any* $t \in [0,T)$.

PROOF. (1) follows from [5, Theorem 4]. (2) The first claim is known. For the latter consequence we use the isometry to obtain

$$\begin{aligned}\sum_{n=1}^{\infty}(n+1)\|I_n(\tilde{f}_n\mathbb{1}_{(0,t]}^{\otimes n})\|_{L_2(\mathbb{P})}^2 &= \sum_{n=1}^{\infty}(n+1)!t^n\|\tilde{f}_n\|_{L_2(\mu^{\otimes n})}^2 \\ &= \sum_{n=1}^{\infty}(n+1)\frac{t^n}{T^n}\|I_n(\tilde{f}_n\mathbb{1}_{(0,T]}^{\otimes n})\|_{L_2(\mathbb{P})}^2 < \infty,\end{aligned}$$

which verifies $\mathbb{E}_{\mathcal{F}_t}[f(X_T)] \in \mathbb{D}_{1,2}$ for $t \in [0,T)$. □

5.1.3. Exponential Lévy process. We present here the relation between two exponential processes induced by a Lévy process, the ordinary exponential process and the stochastic exponential process, which will be exploited later.

Let X be a Lévy process with $(X|\mathbb{P}) \sim (\gamma,\sigma,\nu)$. The *stochastic exponential* of X, denoted by $\mathcal{E}(X)$, is the càdlàg process that satisfies the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{E}(X) = \mathcal{E}(X)_-\mathrm{d}X, \quad \mathcal{E}(X)_0 = 1.$$

We apply [1, Theorem 5.1.6] with the truncation function $x\mathbb{1}_{\{|x|\leqslant 1\}}$ instead of $x\mathbb{1}_{\{|x|<1\}}$ to obtain that, if $\mathcal{E}(X) > 0$, then there exists a Lévy process Y with $(Y|\mathbb{P}) \sim (\gamma_Y,\sigma_Y,\nu_Y)$ such that $\mathcal{E}(X) = \mathrm{e}^Y$, where $\sigma_Y = \sigma$ and

$$\begin{aligned}\nu_Y(B) &= \int_{\mathbb{R}} \mathbb{1}_{\{\ln(1+x)\in B\}}\nu(\mathrm{d}x), \quad B \in \mathcal{B}(\mathbb{R}),\\ \gamma_Y &= \gamma - \frac{\sigma^2}{2} + \int_{\mathbb{R}}\left(\mathbb{1}_{\{|\ln(1+x)|\leqslant 1\}}\ln(1+x) - x\mathbb{1}_{\{|x|\leqslant 1\}}\right)\nu(\mathrm{d}x).\end{aligned}$$

[1] A locally integrable function h is called a *weak derivative* of a locally integrable funtion f on $\mathbb{R}$ if for all smooth functions ϕ with compact support in $\mathbb{R}$ one has $\int_{\mathbb{R}} f(x)\phi'(x)\mathrm{d}x = -\int_{\mathbb{R}} h(x)\phi(x)\mathrm{d}x$. When such an h exists (unique up to a λ-null set), then we denote $f'_w := h$.

Conversely, there is a Lévy process Z with $(Z|\mathbb{P}) \sim (\gamma_Z, \sigma_Z, \nu_Z)$ such that $e^X = \mathcal{E}(Z)$. Moreover, one has $\sigma_Z = \sigma$ and

$$\nu_Z(B) = \int_{\mathbb{R}} \mathbb{1}_{\{e^x - 1 \in B\}} \nu(\mathrm{d}x), \quad B \in \mathcal{B}(\mathbb{R}),$$

$$\gamma_Z = \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left((e^x - 1)\mathbb{1}_{\{|e^x-1|\leqslant 1\}} - x\mathbb{1}_{\{|x|\leqslant 1\}}\right) \nu(\mathrm{d}x).$$

5.2. Lévy setting: Directional gradient estimates

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process as in Subsection 5.1.1 with $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$. We recall the Borel measure

$$\mu(\mathrm{d}x) = \sigma^2 \delta_0(\mathrm{d}x) + x^2 \nu(\mathrm{d}x).$$

To avoid the degenerate case we assume that $\mu(\mathbb{R}) \in (0, \infty]$.

Definition 5.2.1. A Borel function $f : \mathbb{R} \to \mathbb{R}$ belongs to $\mathcal{D}_X$ if $\mathbb{E}|f(x + X_s)| < \infty$ for all $(s, x) \in [0, T] \times \mathbb{R}$. For $f \in \mathcal{D}_X$ we define $F : [0, T] \times \mathbb{R} \to \mathbb{R}$ by

$$F(t, x) := \mathbb{E} f(x + X_{T-t}). \tag{5.2.1}$$

5.2.1. Galtchouk–Kunita–Watanabe (GKW) projection. We additionally assume that $X = (X_t)_{t\in[0,T]}$ is an $L_2(\mathbb{P})$-martingale so that $\mu(\mathbb{R}) \in (0, \infty)$ and assume that $f \in L_2(\mathbb{R}, \mathbb{P}_{X_T})$.

Let $D \in L_2(\mathbb{R}, \mu)$ such that $D \geqslant 0$ and $\int_{\mathbb{R}} D^2(z)\mu(\mathrm{d}z) > 0$, and define

$$\mathrm{d}\rho := \frac{D\mathrm{d}\mu}{\int_{\mathbb{R}} D\mathrm{d}\mu}.$$

According to Lemma 5.1.3(1), the chaos expansion of $f(X_T) \in L_2(\mathbb{P})$ is of the form

$$f(X_T) = \mathbb{E} f(X_T) + \sum_{n=1}^{\infty} I_n(\tilde{f}_n \mathbb{1}_{(0,T]}^{\otimes n}) \quad \text{a.s.},$$

where $\tilde{f}_n \in L_2(\mu^{\otimes n})$ are symmetric. Then we define[2]

$$h_0 := \int_{\mathbb{R}} \tilde{f}_1(z)\rho(\mathrm{d}z) \quad \text{and} \quad h_n(x_1, \ldots, x_n) := \int_{\mathbb{R}} \tilde{f}_{n+1}(x_1, \ldots, x_n, z)\rho(\mathrm{d}z), \quad n \geqslant 1,$$

and define the càdlàg $L_2(\mathbb{P})$-martingale $\varphi(f, \rho) = (\varphi_t(f, \rho))_{t\in[0,T)}$ by the chaos expansion

$$\varphi_t(f, \rho) := \sum_{n=0}^{\infty} (n+1) I_n(h_n \mathbb{1}_{(0,t]}^{\otimes n}),$$

and the càdlàg martingale $X^D = (X_t^D)_{t\in[0,T]}$ by $X_0^D \equiv 0$ and $X_t^D := I_1(\mathbb{1}_{(0,t]} \otimes D)$ a.s., where the integral I_1 is introduced in Subsection 5.1.2.

Denote by $\mathcal{P}_{X^D} : L_2(\mathbb{P}) \to I(X^D) \subseteq L_2(\mathbb{P})$ the orthogonal projection onto the closed subspace

$$I(X^D) := \left\{ \int_{(0,T)} \vartheta_t \mathrm{d}X_t^D : \vartheta \text{ is predictable with } \mathbb{E}\int_0^T \vartheta_t^2 \mathrm{d}t < \infty \right\}.$$

[2]There might be a symmetric $\mu^{\otimes n}$-null-set in $(x_1, \ldots, x_n)$ on which the integral does not exist. On this set we set h_n to be 0.

Then

$$\mathcal{P}_{X^D}(f(X_T)) = \frac{\int_{\mathbb{R}} D \mathrm{d}\mu}{\int_{\mathbb{R}} D^2 \mathrm{d}\mu} \int_{(0,T)} \varphi_{t-}(f,\rho) \mathrm{d}X_t^D \text{ a.s.}$$

For $D \equiv 1$ this was shown by S. Geiss, C. Geiss and Laukkarinen [22, (8), (10), Example (c1) on. p. 209, Lemma 4]. We omit the proof of this extension. The following statement is one motivation of Section 5.2 and will be used in Subsection 5.2.4.

Proposition 5.2.2 (Gradient of GKW-projection)**.** *Assume that the Lévy process X is an $L_2(\mathbb{P})$-martingale, that $f \in \mathcal{D}_X \cap L_2(\mathbb{R}, \mathbb{P}_{X_T})$ and F is given by* (5.2.1)*, that* $\mathrm{d}\rho = D\mathrm{d}\mu / \int_{\mathbb{R}} D\mathrm{d}\mu$ *as above, and that $t \in (0,T)$. Then there is a null-set $N_t \in \mathcal{F}$ such that for $\omega \notin N_t$ one has*[3]

$$\varphi_t(f,\rho)(\omega) = \rho(\{0\}) \partial_x F(t, X_t(\omega)) + \int_{\mathbb{R}_0} \frac{F(t, X_t(\omega)+z) - F(t, X_t(\omega))}{z} \rho(\mathrm{d}z). \tag{5.2.2}$$

Proposition 5.2.2 is proved in [29, Appendix D] by using Malliavin calculus. Results related to Proposition 5.2.2 are provided in Jacob, Méléard and Protter [36, Theorem 2.4], Benth et al. [6, Theorems 2.1, 3.11, 4.1], Cont, Tankov and Voltchkova [14, Proposition 2], and Theorem 6.2.3. Other techniques use the Fourier transform (see, e.g., Brodén and Tankov [9]).

5.2.2. Upper bounds for the gradient process. Gradient estimates in the Lévy setting are studied in different ways in the literature. In [10, Theorem 1.1 and Remark 2.4], Hölder regularities are studied, where one looks for an improvement of the Hölder regularity caused by the transition group. In a way, this is opposite to our question. The result from the literature we contribute to is [52, Theorem 1.3] (see Remark 5.2.21 below). Finally, Laukkarinen [40] investigates when $f(X_T)$ belongs to $\mathbb{D}_{1,2}$ or $(L_2(\mathbb{P}), \mathbb{D}_{1,2})_{\theta,\infty}$ in dependence on $f \in \mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})$ and properties of the underlying Lévy process X. In our context, we look for L_∞- and BMO-bounds for vector-valued gradient processes generated by an $f(X_T)$ when $f \in \mathrm{Höl}^0_{\eta,2}(\mathbb{R})$, where we do not need and consider any Malliavin smoothness of $f(X_T)$ itself. Moreover, for a given $f(X_T)$ the fractional smoothness of the gradient process depends on the direction in which the gradient process is tested. So far, we do not see a way to exploit the results from [40] for our purpose, but it would be worthy to understand connections.

For this section we assume the following setting:

(1) $X = (X_t)_{t\in[0,T]}$ is a Lévy process with $\mu(\mathbb{R}) \in (0,\infty]$.
(2) ρ is a probability measure on $\mathcal{B}(\mathbb{R})$.

Let us start by formalizing the right-hand side of (5.2.2):

Definition 5.2.3. For an $F : [0,T) \times \mathbb{R} \to \mathbb{R}$, such that $x \mapsto F(t,x)$ is measurable for all $t \in [0,T)$, and for $(t,x) \in [0,T) \times \mathbb{R}$ we define

$$D_\rho F(t,x) := \int_{\mathbb{R}_0} \frac{F(t,x+z) - F(t,x)}{z} \rho(\mathrm{d}z) \quad \text{if} \quad \int_{\mathbb{R}_0} \frac{|F(t,x+z) - F(t,x)|}{|z|} \rho(\mathrm{d}z) < \infty.$$

If additionally we have that $F(t,\cdot) \in C^1(\mathbb{R})$, then we let

$$\overline{D}_\rho F(t,x) := \rho(\{0\}) \partial_x F(t,x) + \int_{\mathbb{R}_0} \frac{F(t,x+z) - F(t,x)}{z} \rho(\mathrm{d}z).$$

[3] The integral with respect to $\rho(\mathrm{d}z)$ exists for $\omega \notin N_t$ and we omit $\rho(\{0\})\partial_x F(t, X_t(\omega))$ if $\rho(\{0\}) = 0$.

One point of this definition is that the measure ρ is general. This allows us to capture different aspects: If ρ is as in Proposition 5.2.2, then we can study GKW projections, if ρ is a Dirac measure at $z \in \mathbb{R}_0$, then we study the point-wise behaviour of $(F(t,x+z)-F(t,x))/z$. A general background is provided in [29, Appendix D.3] in terms of a vector-valued gradient process associated to a functional $f(X_T)$.

We recall a class of functions that are of local bounded variation:

Definition 5.2.4. A Borel function $f:\mathbb{R}\to\mathbb{R}$ belongs to $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ provided that f is right continuous and there are Borel measures μ^+ and μ^- on $\mathcal{B}(\mathbb{R})$, finite on each compact interval, and disjoint $S^+, S^- \in \mathcal{B}(\mathbb{R})$ with $S^+ \cup S^- = \mathbb{R}$ and $\mu^+(S^-) = \mu^-(S^+) = 0$, such that

$$f(b) - f(a) = \mu^+((a,b]) - \mu^-((a,b]) \quad \text{for all} \quad -\infty < a < b < \infty.$$

We let $|f'| := \mu^+ + \mu^-$ and, for a Borel function $g:\mathbb{R}\to\mathbb{R}$ with $\int_{\mathbb{R}} |g(x)||f'|(\mathrm{d}x) < \infty$,

$$\int_{\mathbb{R}} g(x) f'(\mathrm{d}x) := \int_{\mathbb{R}} g(x)\mu^+(\mathrm{d}x) - \int_{\mathbb{R}} g(x)\mu^-(\mathrm{d}x).$$

The pair of measures (μ^+,μ^-) is unique and we will identify f' with (μ^+,μ^-). The space $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ consists of functions that are of bounded variation of on each compact interval (see Rudin [50, Chapter 8]). The next definition is the key for what follows and defines two functionals to obtain $D_\rho F(t,x)$, the second term on the right-hand side of (5.2.2), from a given terminal condition f. The first functional simply rephrases $D_\rho F$, the second one uses some kind of partial integration.

Definition 5.2.5. (1) For $t \in [0,T)$ we define $\Gamma^0_{t,\rho} : \mathrm{Dom}(\Gamma^0_\rho) \to \mathbb{R}$ by

$$\mathrm{Dom}(\Gamma^0_\rho) := \Bigg\{ f \in \mathcal{D}_X \text{ and } \forall s \in [0,T)\ \forall\, 0 \leqslant \delta \leqslant s < T\ \forall x \in \mathbb{R} : \\ \mathbb{E}\int_{\mathbb{R}_0} \left| \frac{F(s,x+X_\delta+z) - F(s,x+X_\delta)}{z} \right| \rho(\mathrm{d}z) < \infty \Bigg\},$$

$$\langle f, \Gamma^0_{t,\rho}\rangle := D_\rho F(t,0).$$

(2) For $t \in [0,T)$ we define the Borel function $\gamma_{t,\rho} : \mathbb{R} \to [0,\infty]$ and $\Gamma^1_{t,\rho} : \mathrm{Dom}(\Gamma^1_\rho) \to \mathbb{R}$ by

$$\gamma_{t,\rho}(v) := \int_{\mathbb{R}_0} \frac{\mathbb{P}(X_{T-t} \in J(v;z))}{|z|} \rho(\mathrm{d}z) \quad \text{with} \quad J(v;z) := v + [-z^+, z^-),$$

$$\mathrm{Dom}(\Gamma^1_\rho) := \Bigg\{ f \in \mathcal{D}_X \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}) \text{ and } \forall\, 0 \leqslant \delta \leqslant s < T\ \forall x \in \mathbb{R} : \\ \mathbb{E}\int_{\mathbb{R}} \gamma_{s,\rho}(v - x - X_\delta)|f'|(\mathrm{d}v) < \infty \Bigg\},$$

$$\langle f, \Gamma^1_{t,\rho}\rangle = \langle f', \gamma_{t,\rho}\rangle := \int_{\mathbb{R}} \gamma_{t,\rho}(v)\, f'(\mathrm{d}v).$$

In Definition 5.2.5 we use L_1-conditions instead of L_2-conditions which is sufficient at this point. The L_1-conditions are chosen to guarantee a point-wise definition of $D_\rho F$ and the properties stated in Remark 5.2.6 below.

In Theorem 5.2.10 we prove $\int_{\mathbb{R}} \gamma_{t,\rho}(v)\mathrm{d}v = \rho(\mathbb{R}_0)$, $\mathrm{Dom}(\Gamma^1_\rho) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$, and that

$$\langle f', \gamma_{t,\rho}\rangle = \langle f, \Gamma^0_{t,\rho}\rangle \quad \text{for} \quad f \in \mathrm{Dom}(\Gamma^1_\rho).$$

If $D(\mathbb{R})$ is the test function space that consists of all $f \in C^\infty(\mathbb{R})$ with compact support, then $D(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^1_\rho)$ (for $f \in D(\mathbb{R})$ we have $f'(\mathrm{d}v) = f'(v)\mathrm{d}v$ and $|f'|(\mathrm{d}v) = |f'(v)|\mathrm{d}v$, where f' on the right-hand sides is the classical derivative). If we consider $\gamma_{t,\rho} \in L_1(\mathbb{R})$ as distribution $\gamma_{t,\rho} \in D'(\mathbb{R})$ (see Rudin [49, Section 6.11]), then we have the interpretation

$$D\gamma_{t,\rho} = -\Gamma^0_{t,\rho}, \tag{5.2.3}$$

see [49, Section 6.12] and $\Gamma^0_{t,\rho}$ can be seen as distributional derivative of a distribution of L_1-type.

Before we continue, let us list some facts we exploit later:

Remark 5.2.6. For $f \in \mathrm{Dom}(\Gamma^0_\rho)$ the following holds:

(1) $D_\rho F(t,x) = \langle f(x+\cdot), \Gamma^0_{t,\rho}\rangle$.
(2) One has that $t \mapsto d(t) := \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \in [0,\infty]$ is non-decreasing.
(3) The process $(D_\rho F(t,X_t))_{t\in[0,T)}$ is a martingale.
(4) There exists a càdlàg modification $\varphi = (\varphi_t)_{t\in[0,T)}$ of $(D_\rho F(t,X_t))_{t\in[0,T)}$ such that

$$|\varphi_t| \leqslant d(t+) \quad \text{on} \quad [0,T)\times\Omega.$$

It will be useful to consider $\Gamma^0_{t,\rho}$ as linear functional on semi-normed spaces:

Definition 5.2.7. For $t \in [0,T)$ and a linear space $E \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ equipped with a semi-norm $|\cdot|_E$, we let $\|\Gamma^0_{t,\rho}\|_{E^*} := \inf c$, where the infimum is taken over all $c > 0$ such that

$$|\langle f, \Gamma^0_{t,\rho}\rangle| \leqslant c|f|_E \quad \text{for all} \quad f \in E.$$

In this section we aim for estimates of type

$$\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \leqslant c_{(5.2.4)}(t)|f|_E \quad \text{for all} \quad f \in E. \tag{5.2.4}$$

If E contains only functions f such that $f(0) = 0$ (to have a norm $\|\cdot\|_E$ rather than a semi-norm $|\cdot|_E$ later) and are "translation invariant" in the sense that $\|f\|_E = \|x \mapsto f(x_0+x) - f(x_0)\|_E$ for any $x_0 \in \mathbb{R}$, then the estimate (5.2.4) is equivalent to

$$|D_\rho F(t,0)| = |\langle f, \Gamma^0_{t,\rho}\rangle| \leqslant c_{(5.2.4)}(t)\|f\|_E \quad \text{for all} \quad f \in E.$$

This is the reasoning for the definition of $\langle f, \Gamma^0_{t,\rho}\rangle$, i.e. for the estimates (5.2.4) one does not need to work with the Banach space $B_b(\mathbb{R})$. One application of the results of this section are the upper gradient estimates provided by Corollary 5.2.13 that can be seen as a counterpart to Theorem 3.5.5 proved on the Wiener space. To prove Corollary 5.2.13 we use the interpolation result [29, Theorem 7.1] with end-point estimates derived by Theorem 5.2.9 and Theorem 5.2.12. As an application, inequality (5.2.11) of Corollary 5.2.13 allows for BMO-estimates of $(D_\rho F(t,X_t))_{t\in[0,T)}$ after applying our Riemann–Liouville type operators to its càdlàg version by exploiting Theorem 5.2.11.

To start with, we introduce a variational quantity that is one key for us to obtain upper bounds for gradient processes:

Definition 5.2.8. For $\eta \in [0,1]$ and $s \in [0,T]$, we let

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)} := \inf_P \left\{ \int_{\mathbb{R}_0} P(z)^{1-\eta} \rho(\mathrm{d}z) \right\} \in [0,\infty],$$

where the infimum is taken over all measurable $P : \mathbb{R}_0 \to [0,\infty)$ such that

$$\frac{\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} \leqslant P(z) \quad \text{for} \quad z \in \mathbb{R}_0.$$

We use the potentials P to avoid a discussion about the measurability of the map $z \mapsto \|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}$ (which would not be necessary for us). We have the following special cases:

(1) $\|X_s\|_{\mathrm{TV}(\rho,1)} = \rho(\mathbb{R}_0) < \infty$ for $s \in [0,T]$.

(2) $\|X_0\|_{\mathrm{TV}(\rho,\eta)} = 2^{1-\eta} \int_{\mathbb{R}_0} |z|^{\eta-1} \rho(\mathrm{d}z)$ for $\eta \in [0,1]$.

(3) $\|X_s\|_{\mathrm{TV}(\delta_z,\eta)} = \left(\frac{\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} \right)^{1-\eta} < \infty$, $\eta \in [0,1]$, if δ_z is the Dirac measure at $z \in \mathbb{R}_0$.

We will not use $\|X_0\|_{\mathrm{TV}(\rho,\eta)}$, whereas our idea is to use $\|X_s\|_{\mathrm{TV}(\rho,\eta)}$ for $s \in (0,T]$, where we exploit the behaviour of $\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}$. This enables us to obtain the correct blow-up of gradient processes when considering β-stable-like processes. Upper bounds for $\|X_s\|_{\mathrm{TV}(\delta_z,\eta)}$ can be found in the literature, see Schilling, Sztonyk and Wang [52, Theorem 3.1], Theorem 5.2.9(2) is a variant for our setting.

In Theorem 5.2.9 and Theorem 5.2.10 below we provide basic properties of $\Gamma^0_{t,\rho}$ and $\Gamma^1_{t,\rho}$. We will use Theorem 5.2.9 to deduce upper and Theorem 5.2.10 to deduce lower bounds for our gradient processes. Moreover, Theorem 5.2.10 gives the interpretation (5.2.3) of $\Gamma^0_{t,\rho}$ and $\Gamma^1_{t,\rho}$ as distributions.

Theorem 5.2.9 (Properties of the functional $\Gamma^0_{t,\rho}$)**.** *Assume $\eta \in [0,1]$ and $(X_t)_{t\in[0,T]} \subseteq L_\eta(\mathbb{P})$.*

(1) *If $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for $s \in (0,T]$, then* $\mathrm{H\ddot{o}l}_\eta(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ *and*

$$\|\Gamma^0_{t,\rho}\|_{(\mathrm{H\ddot{o}l}_\eta(\mathbb{R}))^*} \leqslant \|X_{T-t}\|_{\mathrm{TV}(\rho,\eta)}, \tag{5.2.5}$$

where $\mathrm{H\ddot{o}l}_\eta(\mathbb{R})$ *is equipped with the semi-norm*

$$|f|_{\eta,\infty} := \|f - f(0)\|_{\mathrm{H\ddot{o}l}^0_{\eta,\infty}(\mathbb{R})} \quad \text{if} \quad \eta \in (0,1).$$

(2) *If $t \in [0,T)$ and X_{T-t} has a $C^1(\mathbb{R})$-density $y \mapsto p_{T-t}(y)$, then*

$$\|X_{T-t}\|_{\mathrm{TV}(\rho,\eta)} \leqslant \int_{\mathbb{R}_0} \left(\min \left\{ \frac{2}{|z|}, \|\partial_y p_{T-t}\|_{L_1(\mathbb{R})} \right\} \right)^{1-\eta} \rho(\mathrm{d}z).$$

In particular, if $\sigma > 0$, then $p_{T-t} \in C^1(\mathbb{R})$ with $\|\partial_y p_{T-t}\|_{L_1(\mathbb{R})} \leqslant \sqrt{\frac{2}{\pi\sigma^2}}(T-t)^{-\frac{1}{2}}$.

PROOF. (1) First we remark that $(X_t)_{t\in[0,T]} \subseteq L_\eta(\mathbb{P})$ implies that $\mathrm{H\ddot{o}l}_\eta(\mathbb{R}) \subseteq \mathcal{D}_X$. Moreover, for fixed $z \in \mathbb{R}_0$, $t \in [0,T)$, and $f \in \mathrm{H\ddot{o}l}_1(\mathbb{R})$ we obtain the estimate

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant |f|_{\mathrm{H\ddot{o}l}_1(\mathbb{R})} \tag{5.2.6}$$

and, for $f \in B_b(\mathbb{R})$ and $x' \in \mathbb{R}$,

$$|F(t,x+z) - F(t,x)| = \left| \int_{\mathbb{R}} (f(x+y) - f(x'))\mathbb{P}_{z+X_{T-t}}(\mathrm{d}y) - \int_{\mathbb{R}} (f(x+y) - f(x'))\mathbb{P}_{X_{T-t}}(\mathrm{d}y) \right|$$

$$\leqslant \int_{\mathbb{R}} |f(x+y) - f(x')|\,|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}|(\mathrm{d}y)$$
$$\leqslant \|f - f(x')\|_{B_b(\mathbb{R})} \|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}.$$

Therefore,

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant c_{(5.2.7)} \frac{\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}}{|z|} \tag{5.2.7}$$

for $c_{(5.2.7)} := \|f\|_{C_b^0(\mathbb{R})}$ if $f \in C_b^0(\mathbb{R})$ (take $x' = 0$) and $c_{(5.2.7)} := |f|_0$ if $f \in \mathrm{Höl}_0(\mathbb{R})$ (take the supremum over $x' \in \mathbb{R}$ on the right-hand side). Moreover, real interpolation between (5.2.7) for $C_b^0(\mathbb{R})$ and (5.2.6) for $\mathrm{Höl}_1^0(\mathbb{R})$ (for fixed x and z) implies that

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant \|f\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})} \left[\frac{\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}}{|z|} \right]^{1-\eta} \tag{5.2.8}$$

for $\eta \in (0,1)$ by (2.4.1). From (5.2.7) and (5.2.6) we deduce $\mathrm{Höl}_\eta(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma_\rho^0)$ and (5.2.5) for $\eta \in \{0,1\}$. If $\eta \in (0,1)$, then (5.2.8) implies $\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma_\rho^0)$ and (5.2.5) with $\mathrm{Höl}_\eta(\mathbb{R})$ replaced by $\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})$. But if $f \in \mathrm{Höl}_\eta(\mathbb{R})$, then we replace f by $f_0 := f - f(0) \in \mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})$ and get (5.2.8) with constant $\|f - f(0)\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})}$. This concludes the proof of (1).

(2) We observe that

$$\begin{aligned}
\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}} = \|p_{T-t}(\cdot - z) - p_{T-t}\|_{L_1(\mathbb{R})} &= \int_{\mathbb{R}} \left| \int_{x-z}^{x} \partial_y p_{T-t}(y)\mathrm{d}y \right| \mathrm{d}x \\
&\leqslant \mathrm{sign}(z) \int_{\mathbb{R}} \int_{x-z}^{x} \left|\partial_y p_{T-t}(y)\right| \mathrm{d}y\mathrm{d}x \\
&= |z| \int_{\mathbb{R}} \left|\partial_y p_{T-t}(y)\right| \mathrm{d}y.
\end{aligned}$$

As we have $\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}} \leqslant 2$ as well, we obtain the first part of item (2). If $\sigma > 0$ and $s \in (0,T]$, then the density of X_s is given by $p_s(y) := \mathbb{E} p_{\sigma W_s}(y - J_s)$ where $p_{\sigma W_s}$ is the C^∞-density of σW_s and satisfies

$$\|\partial_y p_s\|_{L_1(\mathbb{R})} = \|\mathbb{E}\partial_y p_{\sigma W_s}(\cdot - J_s)\|_{L_1(\mathbb{R})} \leqslant \|\partial_y p_{\sigma W_s}\|_{L_1(\mathbb{R})} = \sqrt{\frac{2}{\pi\sigma^2}} s^{-\frac{1}{2}}. \qquad \square$$

Theorem 5.2.10 (Properties of the functional $\Gamma^1_{t,\rho}$). *Let $t \in [0,T)$.*

(1) *One has* $\int_{\mathbb{R}} \gamma_{t,\rho}(v)\mathrm{d}v = \rho(\mathbb{R}_0)$.

(2) *One has* $\mathrm{Dom}(\Gamma^1_\rho) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ *and for* $f \in \mathrm{Dom}(\Gamma^1_{t,\rho})$ *and* $x \in \mathbb{R}$ *that*

$$D_\rho F(t,x) = \langle f^x, \Gamma^1_{t,\rho} \rangle = \langle f^x, \Gamma^0_{t,\rho} \rangle \quad \textit{if} \quad f^x(\cdot) := f(\cdot + x).$$

(3) *If $q,r \in [1,\infty]$, X_{T-t} has a density $p_{T-t} \in L_r(\mathbb{R})$, and $s := r \wedge q$, then*

$$\|\gamma_{t,\rho}\|_{L_q(\mathbb{R})} \leqslant \|p_{T-t}\|_{L_s(\mathbb{R})} \int_{\mathbb{R}_0} |z|^{\frac{1}{q} - \frac{1}{s}} \rho(\mathrm{d}z).$$

PROOF. See the proof of [29, Theorem 8.10]. $\square$

We return to the Riemann–Liouville type operators and aim for correct upper bounds for (say) $\|\mathcal{I}^\alpha \varphi - \varphi_0\|_{\mathrm{BMO}_2([0,T))}$. *Point-wise* bounds for $\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}$, in the sense that $t \in$

$[0,T)$ is *fixed*, will not yield to optimal results. Instead, we exploit integral bounds expressed by $|||f|||_{\rho,\alpha}$ below.

Theorem 5.2.11. *Assume that* $\alpha > 0$, $f \in \mathrm{Dom}(\Gamma_\rho^0)$, *and*

$$|||f|||_{\rho,\alpha}^2 := \frac{2\alpha}{T^{2\alpha}} \int_0^T (T-t)^{2\alpha-1} \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}^2 \mathrm{d}t < \infty,$$

and define

$$\varepsilon(a)^2 := \frac{2\alpha}{T^{2\alpha}} \int_a^T (T-t)^{2\alpha-1} \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}^2 \mathrm{d}t \leqslant |||f|||_{\rho,\alpha}^2$$

so that $\varepsilon(a) \downarrow 0$ *if* $a \uparrow T$. *For a càdlàg modification* $\varphi = (\varphi_t)_{t\in[0,T)}$ *of* $(D_\rho F(t,X_t))_{t\in[0,T)}$ *one has*

(1) $\left(\frac{T-a}{T}\right)^\alpha \|D_\rho F(a,\cdot)\|_{B_b(\mathbb{R})} \leqslant \varepsilon(a)$ *for* $a \in [0,T)$,

(2) $\mathbb{E}_{\mathcal{F}_a}\left[|\mathcal{I}_t^\alpha \varphi - \mathcal{I}_a^\alpha \varphi|^2\right] \leqslant \varepsilon(a)^2$ *a.s. for* $0 \leqslant a < t < T$,

(3) $\left\|(\mathcal{I}_t^\alpha \varphi - \mathcal{I}_a^\alpha \varphi)_{t\in[a,T)}\right\|_{\mathrm{BMO}_2([a,T))} \leqslant 3\varepsilon(a)$ *for* $a \in [0,T)$.

PROOF. (1) follows from

$$\begin{aligned}\frac{(T-a)^{2\alpha}}{2\alpha}\|D_\rho F(a,\cdot)\|_{B_b(\mathbb{R})}^2 &= \int_a^T (T-t)^{2\alpha-1}\|D_\rho F(a,\cdot)\|_{B_b(\mathbb{R})}^2 \mathrm{d}t \\ &\leqslant \int_a^T (T-t)^{2\alpha-1}\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}^2 \mathrm{d}t = \frac{T^{2\alpha}}{2\alpha}\varepsilon(a)^2.\end{aligned}$$

(2) We assume $B \in \mathcal{F}_a$ of positive measure and apply Proposition 2.3.2, formula (2.3.2), to get

$$\begin{aligned}\int_B |\mathcal{I}_t^\alpha \varphi - \mathcal{I}_a^\alpha \varphi|^2 \mathrm{d}\mathbb{P}_B &= 2\gamma T^{-2\alpha} \int_B \int_a^T (T-u)^{2\alpha-1} |\varphi_{u\wedge t} - \varphi_a|^2 \mathrm{d}u \mathrm{d}\mathbb{P}_B \\ &\leqslant 2\alpha T^{-2\alpha} \int_B \int_a^T (T-u)^{2\alpha-1} |\varphi_{u\wedge t}|^2 \mathrm{d}u \mathrm{d}\mathbb{P}_B \\ &\leqslant 2\alpha T^{-2\alpha} \int_a^T (T-u)^{2\alpha-1} \|D_\rho F(u,\cdot)\|_{B_b(\mathbb{R})}^2 \mathrm{d}u \\ &= \varepsilon(a)^2.\end{aligned}$$

(3) Because the $\mathrm{BMO}_2([a,T))$-norm is invariant when passing to càdlàg modifications, we may assume the bound from Remark 5.2.6(4) for φ in order to get

$$|\Delta \mathcal{I}_t^\alpha \varphi| = \left(\frac{T-t}{T}\right)^\alpha |\Delta \varphi_t| \leqslant 2\varepsilon(t) \quad \text{on} \quad [0,T)\times\Omega.$$

The statement follows from item (2), (2.2.1), and Proposition 2.2.2(1) (applied to the time interval $[a,T)$). □

Theorem 5.2.12 (End-point estimate)**.** *Let* $X = (X_t)_{t\in[0,T]}$ *be a Lévy process. If there are* $\varepsilon \in (0,1)$ *and* $\beta \in (0,\infty]$ *such that*

$$c_{(5.2.9)} := \sup_{n\in\mathbb{N}} 2^{\varepsilon n} \rho(\{2^{-n} \leqslant |z| < 2^{-n+1}\}) < \infty, \tag{5.2.9}$$

$$c_{(5.2.10)} := \sup_{s\in(0,T]} \sup_{z\in\mathrm{supp}(\rho)\setminus\{0\}} s^{\frac{1}{\beta}} \frac{\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} < \infty, \tag{5.2.10}$$

then, for $\eta \in [0,1-\varepsilon)$ *there is a constant* $c = c(\varepsilon,\beta,\eta,c_{(5.2.9)},c_{(5.2.10)}) > 0$ *such that*

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)} \leqslant c\, s^{\frac{\epsilon+\eta-1}{\beta}} \quad \textit{for} \quad s \in (0,T].$$

PROOF. See the proof of [29, Theorem 8.12]. □

By using the interpolation technique and end-point estimates, we deduce the following result (see [29, Corollary 8.3] for the detailed proof):

Corollary 5.2.13. *Assume that* $(X_t)_{t\in[0,T]} \subseteq L_1(\mathbb{P})$ *and either that*

(1) $\sigma > 0$, $(\varepsilon,\beta) = (0,2)$, *or*
(2) $\sigma = 0$, $(\varepsilon,\beta) \in (0,1)\times(1,2)$, *and that* (5.2.9) *and* (5.2.10) *hold.*

Then one has for $\eta \in (0,1-\varepsilon)$, $\alpha := \frac{1-(\varepsilon+\eta)}{\beta} \in \left(0,\frac{1}{\beta}\right)$, *and* $q \in [1,\infty]$ *that* $\mathrm{Höl}^0_{\eta,q}(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ *and*

$$\left\| t \mapsto (T-t)^\alpha \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \right\|_{L_q([0,T),\frac{dt}{T-t})} \leqslant c^{(q)}_{(5.2.11)} \|f\|_{\mathrm{Höl}^0_{\eta,q}(\mathbb{R})} \tag{5.2.11}$$

for $f \in \mathrm{Höl}^0_{\eta,q}(\mathbb{R})$, *where* $c^{(q)}_{(5.2.11)} > 0$ *is a constant independent from* f. *In particular, for* $q = 2$ *we obtain*

$$|||f|||_{\rho,\alpha} \leqslant \frac{\sqrt{2\alpha}}{T^\alpha} c^{(2)}_{(5.2.11)} \|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})},$$

for $q = \infty$ *we obtain*

$$\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \leqslant c^{(\infty)}_{(5.2.11)} \|f\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})} (T-t)^{\frac{\varepsilon+\eta-1}{\beta}}, \quad t \in [0,T).$$

5.2.3. Lower bounds for the oscillation of gradient processes. Theorem 5.2.18 and Theorem 5.2.19 are the main results of this subsection. Their background is from Proposition 5.2.2 where we compute the gradient process of the GKW projection. Theorem 5.2.18 proves the maximal oscillation of these gradients and Theorem 5.2.19 determines the quantitative behaviour of the maximal oscillation as a counterpart to Corollary 5.2.13.

To handle the oscillation we exploit the supports of the laws $\mathbb{P}_{X_t}$ and transform the Lévy process $(X_t)_{t\in[0,T]}$ into the process $(Y_t)_{t\in[0,T]}$ below which has independent and stationary increments as well. The statements Theorem 5.2.15, Example 5.2.16, and Example 5.2.17, are formulated for the Y-process, before we return to the X-process. Let us start with the basic setting of this subsection:

Assumption 5.2.14. (1) In the notation of Example 3.4.6 we use

$$\mathrm{supp}(X_t) = Q + \ell t,\ t \in (0,T], \quad \text{and} \quad Y_t = (X_t - \ell t)\mathbb{1}_{\{X_t \in \mathrm{supp}(X_t)\}} \quad \text{for } t \in [0,T].$$

(2) The function $H\colon [0,T)\times Q \to \mathbb{R}$ is Y*-consistent*, which means
 (a) $H(t,\cdot)$ is continuous on Q for all $t \in [0,T)$,
 (b) $\mathbb{E}|H(t,y+Y_{t-s})| < \infty$ for all $0 \leqslant s \leqslant t < T$ and $y \in Q$,
 (c) $\mathbb{E}H(t,y+Y_{t-s}) = H(s,y)$ for all $0 \leqslant s \leqslant t < T$ and $y \in Q$.
(3) ρ is a probability measure on $\mathcal{B}(\mathbb{R})$.

The reason for this definition is the following statement, where we recall Definition 3.4.1.

Theorem 5.2.15. *Let* H *be* Y*-consistent and* $\varphi_t := H(t,Y_t)$, $t \in [0,T)$. *Then* $\varphi = (\varphi_t)_{t\in[0,T)}$ *is a martingale of maximal oscillation with constant* 2. *Moreover, if for all* $t \in [0,T)$ *there is an* $\bar{t} \in (t,T)$ *such that* $H(t,Y_{\bar{t}}) \in L_2(\mathbb{P})$, *then the following assertions are equivalent:*

(1) $\inf_{t\in(0,T)} \underline{\mathrm{Osc}}_t(\varphi) = 0$.
(2) $\varphi_t = \varphi_0$ *a.s. for all* $t \in (0,T)$.

Item (2) in Theorem 5.2.15 implies a *forward uniqueness*: If there is an $s \in (0,T)$ such that $\varphi_0 = \varphi_s$ a.s., then the martingale is constant a.s.

PROOF OF THEOREM 5.2.15. The martingale property follows by the definition and the maximal oscillation with constant 2 follows from Example 3.4.6. Regarding the equivalence we only need to show (1)⇒(2). For $0 < s < t < T$, $y_1', y_2' \in Q$ and $\omega \in (Y_t - Y_s)^{-1}(Q)$ we obtain that

$$\begin{aligned}\|\varphi_t - \varphi_s\|_{L_\infty(\mathbb{P})} &= \sup_{y,y'\in Q} |H(t,y+y') - H(s,y')| \\ &\geqslant |H(t,y_1' + (Y_t - Y_s)(\omega) + y_2' + Y_s(\omega)) - H(s,y_2' + Y_s(\omega))| \\ &= |H(t,y_1' + y_2' + Y_t(\omega)) - H(s,y_2' + Y_s(\omega))|,\end{aligned}$$

where the first inequality comes from $\varphi_t - \varphi_s = H(t,Y_t - Y_s + Y_s) - H(s,Y_s)$, $\mathrm{supp}(Y_t - Y_s, Y_s) = Q \times Q$, and from the continuity of $Q \times Q \ni (y,y') \mapsto H(t,y+y') - H(s,y')$. This implies

$$\begin{aligned}\|\varphi_t - \varphi_s\|_{L_\infty(\mathbb{P})} &\geqslant \sup_{y,y'\in Q} |\mathbb{E}H(t,y+y'+Y_t) - \mathbb{E}H(s,y'+Y_s)| \\ &= \sup_{y,y'\in Q} |H(0,y+y') - H(0,y')| \\ &\geqslant \sup_{y\in Q} |H(0,y) - H(0,0)|.\end{aligned}$$

For $s = 0$ we use the same idea with $y' = y_2' = 0$ to get $\|\varphi_t - \varphi_0\|_{L_\infty(\mathbb{P})} \geqslant \sup_{y\in Q} |H(0,y) - H(0,0)|$. So (1) yields to $C := H(0,0) = H(0,y)$ for all $y \in Q$. Fix $0 \leqslant t < \bar{t} < T$ as in our assumption. According to Lemma 5.1.3, we have a chaos expansion

$$H(t,Y_{\bar{t}}) = \mathbb{E}H(t,Y_{\bar{t}}) + \sum_{n=1}^{\infty} I_n\left(\tilde{h}_n \mathbb{1}_{[0,\bar{t}]}^{\otimes n}\right)$$

with $\tilde{h}_n \in L_2(\mu^{\otimes n})$. Let $\widetilde{Y}$ be an independent copy of Y with the corresponding expectation $\widetilde{\mathbb{E}}$. For $\Delta t := \bar{t} - t > 0$ this implies $\mathbb{E}_{\mathcal{F}_{\Delta t}}[H(t,Y_{\bar{t}})] = \widetilde{\mathbb{E}}H(t,Y_{\Delta t} + \widetilde{Y}_t) = H(0,Y_{\Delta t}) = C$ a.s. and

$$C = \mathbb{E}_{\mathcal{F}_{\Delta t}}[H(t,Y_{\bar{t}})] = \mathbb{E}H(t,Y_{\bar{t}}) + \sum_{n=1}^{\infty} I_n\left(\tilde{h}_n \mathbb{1}_{[0,\Delta t]}^{\otimes n}\right) \quad \text{a.s.}$$

Therefore, $\tilde{h}_n = 0$ in $L_2(\mu^{\otimes n})$ for all $n \geqslant 1$, which yields $H(t,Y_{\bar{t}}) = C$ a.s. Since $\mathrm{supp}(Y_{\bar{t}}) = Q = \mathrm{supp}(Y_t)$, together with the continuity of $H(t,\cdot)$ on Q, we derive that $H(t,y) = C$ for all $y \in Q$. Therefore $\varphi_t = H(t,Y_t) = C$ a.s. □

The next two results provide the fundamental examples of Y-consistent functions:

Example 5.2.16 ([29], Example 8.17)**.** We assume

(1) that $k\colon Q \to \mathbb{R}$ is a Borel function with $\mathbb{E}|k(y+Y_s)| < \infty$ for $(s,y) \in [0,T] \times Q$ and that $K\colon [0,T) \times Q \to \mathbb{R}$ with $K(t,y) := \mathbb{E}k(y + Y_{T-t})$ satisfies

$$\mathbb{E}\int_{Q\setminus\{0\}} \left|\frac{K(t,y+Y_\delta+z) - K(t,y+Y_\delta)}{z}\right| \rho(\mathrm{d}z) < \infty \quad \text{for} \quad 0 \leqslant \delta \leqslant t < T,$$

(2) that $y \mapsto K(t,y)$ is continuous on Q for $t \in [0,T)$,

(3) that for all $(t,y) \in [0,T) \times Q$ there is an $\varepsilon > 0$ such that the family of functions
$$z \mapsto \frac{K(t,y'+z)-K(t,y')}{z},$$
indexed by $y' \in Q$ with $|y-y'| < \varepsilon$, is uniformly integrable on $(Q\backslash\{0\},\rho)$.

Then we obtain a Y-consistent function by
$$H(t,y) := \int_{Q\backslash\{0\}} \frac{K(t,y+z)-K(t,y)}{z}\rho(\mathrm{d}z) \quad \text{for} \quad (t,y) \in [0,T) \times Q.$$

Example 5.2.17 ([29], Example 8.18)**.** Let $\sigma > 0$. Then $Q = \mathbb{R}$ and the following holds:

(1) If $k:\mathbb{R} \to \mathbb{R}$ is a Borel function with $\mathbb{E}|k(Y_T)|^q < \infty$ for some $q \in (1,\infty)$, then
$$\mathbb{E}|k(y+Y_{T-t})| < \infty \quad \text{for} \quad (t,y) \in [0,T] \times \mathbb{R}.$$
If $K(t,y) := \mathbb{E}k(y+Y_{T-t})$ on $[0,T] \times \mathbb{R}$, then $K(t,\cdot) \in C^\infty(\mathbb{R})$ for $t \in [0,T)$ and we obtain a Y-consistent function $H : [0,T) \times \mathbb{R} \to \mathbb{R}$ by
$$H(t,y) := \partial_y K(t,y) \quad \text{with} \quad H(t,y) = \frac{1}{\sigma}\mathbb{E}\left[k(y+Y_{T-t})\frac{W_{T-t}}{T-t}\right].$$

(2) If $k \in \mathrm{Höl}_\eta(\mathbb{R})$ for some $\eta \in [0,1]$ (and $\mathbb{E}|k(Y_T)|^q < \infty$ as above if $\eta \in (0,1]$), then
$$\|H(t,\cdot)\|_{B_b(\mathbb{R})} \leqslant |k|_{\mathrm{Höl}_\eta(\mathbb{R})}\sigma^{\eta-1}(T-t)^{\frac{\eta-1}{2}} \int_{\mathbb{R}} |x|^{\eta+1}\mathrm{e}^{-\frac{x^2}{2}} \frac{\mathrm{d}x}{\sqrt{2\pi}}.$$

Now we are in a position to return to the setting of Proposition 5.2.2:

Theorem 5.2.18 (Maximal oscillation)**.** *Suppose that*

(a) *the Lévy process $(X_t)_{t\in[0,T]}$ is an $L_2(\mathbb{P})$-martingale and $\rho := \mu/\mu(\mathbb{R})$,*

(b) *$\eta \in [0,1]$ and $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for all $s \in (0,T]$ if $\eta \in [0,1)$,*

(c) *$f \in \mathrm{Höl}_\eta(\mathbb{R})$, where we additionally assume that $y \mapsto f(y+\ell T)$ is continuous on Q if $\eta = \sigma = 0$.*

Then $f \in \mathrm{Dom}(\Gamma_\rho^0)$ and, additionally, $F(t,\cdot) \in C^\infty(\mathbb{R})$ for $t \in [0,T)$ if $\sigma > 0$. Letting $\varphi_t := \overline{D}_\rho F(t,X_t)$ for $t \in [0,T)$, the following holds:

(1) *$\|\varphi_t\|_{L_\infty(\mathbb{P})} = \sup_{x\in\mathrm{supp}(X_t)} |\overline{D}_\rho F(t,x)|$ for $t \in [0,T)$.*

(2) *$(\varphi_t)_{t\in[0,T)}$ is an $L_2(\mathbb{P})$-martingale of maximal oscillation with constant 2.*

(3) *Unless $\varphi_t = \varphi_0$ a.s. for all $t \in [0,T)$, one has $\inf_{t\in(0,T)} \underline{\mathrm{Osc}}_t(\varphi) > 0$.*

PROOF. See the proof of [29, Theorem 8.19]. □

Now we provide in Theorem 5.2.19 the corresponding lower bounds for Corollary 5.2.13. The conditions (5.2.12) and (5.2.13) are a counterpart to (5.2.9) and (5.2.10) assumed in Corollary 5.2.13. The proof of this result is given in [29, Theorem 8.20].

Theorem 5.2.19 (Size of maximal oscillation)**.** *Suppose that*

(a) *the Lévy process $(X_t)_{t\in[0,T]}$ is an $L_2(\mathbb{P})$-martingale and $\rho := \mu/\mu(\mathbb{R})$,*

(b) *$\eta \in [0,1)$ and $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for all $s \in (0,T]$,*

(c) *$f_\eta : \mathbb{R} \to \mathbb{R} \in \mathrm{Höl}_\eta(\mathbb{R})$ is given by $f_\eta(x) := (x \vee 0)^\eta$ if $\eta \in (0,1)$ and $f_0(x) := \mathbb{1}_{[0,\infty)}(x)$.*

If $F_\eta(t,x) := \mathbb{E} f_\eta(x+X_{T-t})$ for $(t,x) \in [0,T) \times \mathbb{R}$, then one has

(1) $\inf_{t\in[0,T)}(T-t)^{1-\frac{1+\eta}{2}}\partial_x F_\eta(t,0)>0$ *if* $\sigma>0$,

(2) $\inf_{t\in[0,T)}(T-t)^{1-\frac{1+\eta}{\beta}}D_\rho F_\eta(t,0)>0$ *if* $\sigma=0$ *and* $\beta\in[1+\eta,2)$, *and if*

$$\rho([-d,d])\geqslant c_{(5.2.12)}d^{2-\beta}\quad \textit{for}\quad d\in(0,d_{(5.2.12)}], \tag{5.2.12}$$

$$\inf_{|v|\vee|z|\leqslant\tilde c_{(5.2.13)}s^{\frac{1}{\beta}},z\neq 0}\frac{\mathbb{P}(X_s\in J(v;z))}{|z|}\geqslant c_{(5.2.13)}s^{-\frac{1}{\beta}}\quad \textit{for}\quad s\in(0,T], \tag{5.2.13}$$

for some constants $c_{(5.2.12)}$, $d_{(5.2.12)}$, $c_{(5.2.13)}$, $\tilde c_{(5.2.13)}>0$ *and where* $J(v;z)=v+[-z^+,z^-)$.

5.2.4. Sharpness of the results - β-stable-like processes. In this section we assume a Lévy process $X=(X_t)_{t\in[0,T]}$ with $\sigma=0$, which is an $L_2(\mathbb{P})$-martingale, and $\beta\in(1,2)$ such that the Lévy measure satisfies $\nu(\mathrm{d}z)=p_\nu(z)\mathrm{d}z$, where p_ν is symmetric and

$$0<\liminf_{|z|\to 0}|z|^{1+\beta}p_\nu(z)\leqslant\limsup_{|z|\to 0}|z|^{1+\beta}p_\nu(z)<\infty.$$

We consider a $D\in L_2(\mathbb{R},\mu)$ with $D\geqslant 0$ and $\int_{\mathbb{R}}D^2\mathrm{d}\mu>0$, and set

$$\mathrm{d}\rho:=\frac{D\mathrm{d}\mu}{\int_{\mathbb{R}}D\mathrm{d}\mu}.$$

Given $\varepsilon\in(0,1)$, the small-ball assumption (5.2.9) reads as

$$\left(\int_{\mathbb{R}}D\mathrm{d}\mu\right)\sup_{n\in\mathbb{N}}2^{\varepsilon n}\rho(\{2^{-n}\leqslant|z|<2^{-n+1}\})=\sup_{n\in\mathbb{N}}2^{\varepsilon n}\int_{[2^{-n},2^{-n+1})}D\mathrm{d}\mu<\infty. \tag{5.2.14}$$

Given $f\in\mathcal{D}_X\cap L_2(\mathbb{R},\mathbb{P}_{X_T})$, we also discuss the Riemann approximation of the stochastic integral

$$\int_{(0,T)}\varphi_{t-}(f,\rho)\mathrm{d}X_t^D$$

that represents by Proposition 5.2.2 the GKW projection of $f(X_T)$ on $I(X^D)$ up to a factor. The corresponding error process with respect to the time-net $\tau=(t_i)_{i=0}^n\in\mathcal{T}_{\mathrm{det}}$ is

$$E_t(f;\tau,D):=\int_{(0,t]}\varphi_{s-}(f,\rho)\mathrm{d}X_s^D-\sum_{i=1}^n\varphi_{t_{i-1}-}(f,\rho)(X^D_{t_i\wedge t}-X^D_{t_{i-1}\wedge t}),\quad t\in[0,T).$$

Theorem 5.2.20. *Let* $\eta\in(0,1-\varepsilon)$, $\alpha:=\frac{1-(\varepsilon+\eta)}{\beta}$, $\theta:=1-2\alpha$, *and assume that the functional* D *satisfies the* ε*-small ball property* (5.2.14). *Then* $\mathrm{Höl}^0_{\eta,2}(\mathbb{R})\subseteq\mathrm{Dom}(\Gamma^0_\rho)$ *and the following holds:*

<u>Upper bounds:</u> *For* $f\in\mathrm{Höl}^0_{\eta,2}(\mathbb{R})$ *and the parameters* $\Theta:=(\beta,\nu,\varepsilon,D,\eta,T)$ *the following holds:*

(1) *There is a* $c_{(5.2.15)}=c(\Theta)>0$ *such that one has*

$$\sup_{t\in[0,T)}(T-t)^\alpha\|\varphi_t(f,\rho)\|_{L_\infty(\mathbb{P})}+\|\mathcal{I}^\alpha\varphi(f,\rho)-\varphi_0(f,\rho)\|_{\mathrm{BMO}_2([0,T))}\leqslant c_{(5.2.15)}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}, \tag{5.2.15}$$

$$\lim_{a\uparrow T}\left\|\left(\mathcal{I}^\alpha_t\varphi(f,\rho)-\mathcal{I}^\alpha_a\varphi(f,\rho)\right)_{t\in[a,T)}\right\|_{\mathrm{BMO}_2([a,T))}=0. \tag{5.2.16}$$

(2) *There is a* $c_{(5.2.17)} = c(\Theta) > 0$ *such that one has*

$$\|E(f;\tau,D)\|_{\mathrm{bmo}_2([0,T))} \leqslant c_{(5.2.17)} \sqrt{\|\tau\|_\theta}\, \|f\|_{\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})}. \tag{5.2.17}$$

(3) $\varphi(f,\rho)$ *has maximal oscillation with constant 2.*

(4) *Unless* $\varphi(f,\rho)$ *is almost surely constant, one has* $\inf_{t\in[0,T)} \underline{\mathrm{Osc}}_t(\varphi(f,\rho)) > 0$.

(5) *If* $p \in [2,\infty)$, *then there is a* $c_{(5.2.18)} > 0$ *such that for* $0 \leqslant a < t < T$, $\Phi \in \mathrm{CL}^+([0,t])$ *with* $1 \vee |\Delta X_s| \leqslant \Phi_s$ *on* $[0,t]$, $\sup_{u\in[0,t]} \Phi_u \in L_p(\mathbb{P})$, *and* $\lambda > 0$ *one has*

$$\mathbb{P}_{\mathcal{F}_a}\left(|E_t(f;\tau_n^\theta,D) - E_a(f;\tau_n^\theta,D)| > \lambda\right) \leqslant c_{(5.2.18)} \min\left\{\frac{1}{n\lambda^2}, \frac{\mathbb{E}_{\mathcal{F}_a}\left[\sup_{u\in[a,t]}\Phi_u^p\right]}{\lambda^p(T-t)^{p\alpha}}\right\} \text{ a.s.} \tag{5.2.18}$$

<u>LOWER BOUNDS:</u> *For* $D \equiv 1$ *we can take* $\varepsilon = 2-\beta$ *and there is an* $f_\eta \in \mathrm{H\ddot{o}l}_\eta(\mathbb{R})$ *such that for* $\varphi_t := \varphi_t(f_\eta,\rho)$ *one has:*

(6) $\inf_{t\in(0,T)} (T-t)^\alpha \underline{\mathrm{Osc}}_t(\varphi) > 0$.

(7) *There is a* $c_{(5.2.19)} > 0$ *such that for all* $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ *with* $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$,

$$\inf_{\vartheta_{i-1}\in L_0(\mathcal{F}_{t_{i-1}})} \sup_{a\in[t_{k-1},t_k)} \left\| \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)} \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1} \mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|_{L_\infty(\mathbb{P})} \geqslant c_{(5.2.19)}^2 \|\tau\|_\theta. \tag{5.2.19}$$

(8) $\|E(f;\tau,1)\|_{\mathrm{bmo}_2([0,T))} \geqslant \sqrt{\mu(\mathbb{R}) c_{(5.2.19)}} \sqrt{\|\tau\|_\theta}$ *for all* $\tau \in \mathcal{T}_{\mathrm{det}}$.

Remark 5.2.21. From the above theorem we get that

$$\|\varphi_t(f,\rho)\|_{L_\infty(\mathbb{P})} \leqslant c_{(5.2.15)} (T-t)^{-\frac{1-(\varepsilon+\eta)}{\beta}} \|f\|_{\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})}.$$

Let us take a sequence of real numbers $|z_l| = 2^{-l}$, $l \in \mathbb{N}$, and consider the corresponding Dirac-measures $\rho_l = \delta_{z_l}$. Suppose that the small ball condition

$$\rho_l(\{2^{-n} \leqslant |z| < 2^{-n+1}\}) \leqslant c_{(5.2.9)} 2^{-\varepsilon n}$$

holds uniformly in l and n. Because $\rho_l(\{2^{-l} \leqslant |z| < 2^{-l+1}\}) = 1$ this implies that $1 \leqslant c_{(5.2.9)} 2^{-\varepsilon n}$ for all $n \in \mathbb{N}$ and finally $\varepsilon = 0$. If we interpret $f \in B_b(\mathbb{R})$ as $\eta = 0$, then we would get an exponent

$$(T-t)^{-\frac{1-(\varepsilon+\eta)}{\beta}} = (T-t)^{-\frac{1}{\beta}}$$

which is the upper bound of [52, Theorem 1.3].

5.3. Gradient type estimates in the exponential Lévy setting

Because of the weighted setting which is caused by the usage of exponential Lévy processes, it seems that we cannot use the interpolation method as in Section 5.2 to derive gradient estimates, at least in a straightforward way.

We first introduce some sub-classes of Hölder continuous functions and bounded Borel functions, where the payoff functions are contained in.

Definition 5.3.1 ([60])**.** For $\eta \in [0,1]$ and $q \in [1,\infty]$, and for a non-empty open interval $U \subseteq \mathbb{R}$, we define

$$\mathrm{Höl}_\eta(U) := \left\{ f : U \to \mathbb{R} \text{ Borel} : |f|_{\mathrm{Höl}_\eta(U)} := \sup_{x,y\in U, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\eta} < \infty \right\},$$

$$\mathring{W}^{1,q}(U) := \left\{ f : U \to \mathbb{R} : \exists k \in L_q(U),\ f(y)-f(x) = \int_x^y k(u)\mathrm{d}u, \forall x,y \in U, x<y \right\},$$

and let $|f|_{\mathring{W}^{1,q}(U)} := \|k\|_{L_q(U)}$.

For $\eta \in [0,1]$, Hölder's inequality implies that

$$\mathring{W}^{1,\frac{1}{1-\eta}}(U) \subseteq \mathrm{Höl}_\eta(U) \quad \text{with} \quad |f|_{\mathrm{Höl}_\eta(U)} \leqslant |f|_{\mathring{W}^{1,\frac{1}{1-\eta}}(U)}, \quad \forall f \in \mathring{W}^{1,\frac{1}{1-\eta}}(U).$$

In particular, $\mathring{W}^{1,\infty}(U) = \mathrm{Höl}_1(U)$, which is the space of Lipschitz functions on U.

Definition 5.3.2 ([60], *α-stable-like Lévy measures*)**.** Let ν be a Lévy measure and $\alpha \in (0,2)$.

(1) We let $\nu \in \mathcal{S}_1(\alpha)$ if one can decompose $\nu = \nu_1 + \nu_2$, where ν_1, ν_2 are Lévy measures that satisfy

$$\limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu_2(\mathrm{d}x) < \infty, \tag{5.3.1}$$

$$\nu_1(\mathrm{d}x) = \frac{k(x)}{|x|^{\alpha+1}} \mathbb{1}_{\{x\neq 0\}}\mathrm{d}x, \tag{5.3.2}$$

where $0 < \liminf_{x\to 0} k(x) \leqslant \limsup_{x\to 0} k(x) < \infty$, and the function $x \mapsto \frac{k(x)}{|x|^\alpha}$ is non-decreasing on $(-\infty,0)$ and non-increasing on $(0,\infty)$.

(2) We let $\nu \in \mathcal{S}_2(\alpha)$ if

$$0 < \liminf_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu(\mathrm{d}x) \leqslant \limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu(\mathrm{d}x) < \infty.$$

Remark 5.3.3. Let $\alpha \in (0,2)$.

(1) One has $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$. Indeed, for $\nu \in \mathcal{S}_1(\alpha)$ with the decomposition $\nu = \nu_1 + \nu_2$, a computation shows that $\nu_1 \in \mathcal{S}_2(\alpha)$. Hence, $\nu \in \mathcal{S}_2(\alpha)$. Moreover, since $\nu(\mathrm{d}x) := x^{-1-\alpha}\mathbb{1}_{(0,1)}(x)\mathrm{d}x$ belongs to $\mathcal{S}_2(\alpha)\backslash\mathcal{S}_1(\alpha)$, the inclusion $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$ is *strict*.

(2) According to [8, Theorem 3.2], if $\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in (0,2)$, then α is equal to the *Blumenthal–Getoor index* of ν, i.e. $\alpha = \inf\{r \in [0,2] : \int_{|x|\leqslant 1} |x|^r \nu(\mathrm{d}x) < \infty\}$.

We provide a sufficient condition for a Lévy measure in $\mathcal{S}_1(\alpha)$.

Lemma 5.3.4. *If a Lévy measure ν has a density $p(x) := \frac{\nu(\mathrm{d}x)}{\mathrm{d}x}$ which satisfies*

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

for some $\alpha \in (0,2)$, then $\nu \in \mathcal{S}_1(\alpha)$.

PROOF. By assumption, there exist constants $0 < c \leqslant C < \infty$ and $\varepsilon > 0$ such that

$$c|x|^{-1-\alpha} \leqslant p(x) \leqslant C|x|^{-1-\alpha}, \quad \forall |x| \leqslant \varepsilon.$$

We let

$$\nu_1(\mathrm{d}x) := c\mathbb{1}_{\{0<|x|\leqslant\varepsilon\}}|x|^{-1-\alpha}\mathrm{d}x \quad \text{and} \quad \nu_2(\mathrm{d}x) := \nu(\mathrm{d}x) - \nu_1(\mathrm{d}x).$$

Then ν_1 satisfies (5.3.2). For ν_2, we have

$$\int_{\mathbb{R}}(1-\cos(ux))\nu_2(\mathrm{d}x) \leqslant (C-c)\int_{|x|\leqslant\varepsilon}\frac{1-\cos(ux)}{|x|^{1+\alpha}}\mathrm{d}x + 2\int_{|x|>\varepsilon}\nu(\mathrm{d}x),$$

which implies that (5.3.1) holds for ν_2. Hence, $\nu \in \mathcal{S}_1(\alpha)$. □

Example 5.3.5. Let us provide some examples for those classes of Hölder functions and of α-stable-like processes used in financial modelling.

(1) The *European call* and *put* are Lipschitz, hence they belong to $\mathring{W}^{1,\infty}(\mathbb{R}_+)$.
The *power call* $g(y) := ((y-K)\vee 0)^\eta$ with $K>0$ and $\eta \in (0,1)$ belongs to $C^{0,\eta}(\mathbb{R}_+)$, but $g \notin \mathring{W}^{1,q}(\mathbb{R}_+)$ for any $q \in (1,\infty)$. However, we can decompose $g = g_1 + g_2$, where $g_1(y) := ((y-K)\vee 0)^\eta \wedge 1$ and $g_2 := g - g_1$, so that $g_1 \in \cap_{1\leqslant q<\frac{1}{1-\eta}}\mathring{W}^{1,q}(\mathbb{R}_+)$ and g_2 is Lipschitz. This decomposition of g fits well with the quadratic hedging approach we choose later, which asserts that the hedging strategy of g is the sum of that of g_1 and g_2.
The *binary option* $g(y) := \mathbb{1}_{[K,\infty)}(y)$ belongs to $C^{0,0}(\mathbb{R}_+)$ obviously.

(2) The *CGMY process* with parameters $C, G, M > 0$ and $Y \in (0,2)$ (see Schoutens [53, Section 5.3.9]) 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 $\mathcal{S}_1(Y)$ due to Lemma 5.3.4.
The *Normal Inverse Gaussian* (NIG) process (see Schoutens [53, Section 5.3.8]) has the Lévy density $p_{\mathrm{NIG}}(x) := \nu_{\mathrm{NIG}}(\mathrm{d}x)/\mathrm{d}x$ that satisfies

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

Hence, Lemma 5.3.4 verifies that $\nu_{\mathrm{NIG}} \in \mathcal{S}_1(1)$.

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process with $(X|\mathbb{P}) \sim (\gamma,\sigma,\nu)$. Proposition 5.3.6 below is a variant of Theorem 5.2.9 in the exponential Lévy setting.

Proposition 5.3.6 (Hölder estimates)**.** *Let $g \in C^{0,\eta}(\mathbb{R}_+)$ with $\eta \in [0,1]$. Define*

$$P_t g(y) := \mathbb{E}g(y\mathrm{e}^{X_t}), \quad y>0,\ t \in (0,T].$$

Then there exists a constant $c_{(5.3.3)} > 0$ such that for any $z>0$, $y>0$ and any $t \in (0,T]$ one has

$$|P_t g(z) - P_t g(y)| \leqslant c_{(5.3.3)} U_t(y,z), \tag{5.3.3}$$

where the cases for $U_t(y,z)$ are provided as follows:

(1) *If $\sigma>0$ and $\int_{|x|>1}\mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$, then $U_t(y,z) = \big(t^{\frac{\eta-1}{2}}\frac{|z^\eta - y^\eta|}{\eta}\big) \wedge |z-y|^\eta$.*

(2) *When $\sigma = 0$ and $\int_{|x|>1}\mathrm{e}^{x}\nu(\mathrm{d}x) < \infty$:*

(a) *If $\nu \in \mathcal{S}_1(\alpha)$ for some $\alpha \in (0,2)$, then $U_t(y,z) = \big(t^{\frac{\eta-1}{\alpha}}\frac{|z^\eta - y^\eta|}{\eta}\big) \wedge |z-y|^\eta$.*

(b) *If $\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in (0,2)$ and $g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$, then*

$$U_t(y,z) = \big(t^{\frac{\eta-1}{\alpha}}|\ln z - \ln y|^{1-\eta}|z-y|^\eta\big) \wedge |z-y|^\eta.$$

Here, we set $0^0 := 1$ *and* $\frac{|z^0-y^0|}{0} := \lim_{\eta\downarrow 0}\frac{|z^\eta-y^\eta|}{\eta} = |\ln z - \ln y|$ *by convention.*

PROOF. See the proof of [60, Proposition 8.5]. □

Motivated by the hedging strategies established later in (6.2.2) and (7.2.1), we write symbolically for a Lévy measure ℓ and a Borel function g the formula

$$\Gamma_\ell(t,y) := \sigma^2\partial_y P_t g(y) + \int_{\mathbb{R}} \frac{P_t g(\mathrm{e}^x y) - P_t g(y)}{y}(\mathrm{e}^x-1)\ell(\mathrm{d}x) \tag{5.3.4}$$

for $(t,y)\in(0,T]\times\mathbb{R}_+$, where $P_t g(y) = \mathbb{E}g(y\mathrm{e}^{X_t})$, and we set $\partial_y P_t g(y) := 0$ if $\sigma = 0$.

Theorem 5.3.7 (Gradient type estimates)**.** *Let ℓ be a Lévy measure and $g\in \mathrm{H\ddot{o}l}_\eta(\mathbb{R}_+)$ with $\eta\in[0,1]$. Assume $\int_{|x|>1}\mathrm{e}^{(\eta+1)x}\ell(\mathrm{d}x)<\infty$. Then $\Gamma_\ell(t,y)$ is well-defined for all $(t,y)\in(0,T]\times\mathbb{R}_+$, and there is a constant $c_{(5.3.5)}>0$ such that*

$$|\Gamma_\ell(t,y)| \leqslant c_{(5.3.5)} V_t y^{\eta-1}, \quad \forall (t,y)\in(0,T]\times\mathbb{R}_+, \tag{5.3.5}$$

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 $\int_{|x|\leqslant 1}|x|^{\eta+1}\ell(\mathrm{d}x)<\infty$, then $V_t=1$.*
(3) *If $\sigma=0$ and if the following two conditions hold:*
 (a) *$\nu\in\mathcal{S}_1(\alpha)$ for some $\alpha\in(0,2)$ and $\int_{|x|>1}\mathrm{e}^x\nu(\mathrm{d}x)<\infty$,*
 (b) *there is a $\beta\in(1+\eta,2]$ such that*
$$0<\sup_{r\in(0,1]} r^\beta\int_{|x|\leqslant 1}\left(\left|\frac{x}{r}\right|^2\wedge\left|\frac{x}{r}\right|^{\eta+1}\right)\ell(\mathrm{d}x)<\infty, \tag{5.3.6}$$
 then one has $V_t = t^{\frac{\eta+1-\beta}{\alpha}}$.
(4) *If $\sigma=0$ and $g\in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$, and if the following two conditions hold:*
 (a) *$\nu\in\mathcal{S}_2(\alpha)$ for some $\alpha\in(0,2)$ and $\int_{|x|>1}\mathrm{e}^x\nu(\mathrm{d}x)<\infty$,*
 (b) *there is a $\beta\in(1+\eta,2]$ such that* (5.3.6) *is satisfied,*
 then one has $V_t = t^{\frac{\eta+1-\beta}{\alpha}}$.

Here, the constant $c_{(5.3.5)}$ may depend on β in items (3) *and* (4).

Remark 5.3.8. Since $|\frac{x}{r}|^2\wedge|\frac{x}{r}|^{\eta+1}\leqslant|\frac{x}{r}|^\beta$ for $\beta\in(1+\eta,2]$, a sufficient condition for (5.3.6) is that $0<\int_{|x|\leqslant 1}|x|^\beta\ell(\mathrm{d}x)<\infty$.

Remark 5.3.9. Let us discuss the connection between item (2) in Corollary 5.2.13 and item (3) in Theorem 5.3.7. Let $\eta\in(0,1)$ and let $\nu(\mathrm{d}x) = p_\nu(x)\mathrm{d}x$ be a Lévy measure satisfying

$$0<\liminf_{|x|\to 0}|x|^{1+\kappa}p_\nu(x)\leqslant\limsup_{|x|\to 0}|x|^{1+\kappa}p_\nu(x)<\infty \quad \text{for some } \kappa\in(1+\eta,2).$$

In Corollary 5.2.13(2), we let $\rho(\mathrm{d}x) := x^2\nu(\mathrm{d}x)$, and choose $\varepsilon = 2-\kappa$, $\beta=\kappa$ to obtain that the growth rate in time of the gradient process is $\frac{\eta+1}{\kappa}-1$. On the other hand, in Theorem 5.3.7(3), we choose $\ell=\nu$ and $\alpha=\beta=\kappa$ to get $V_t = t^{\frac{\eta+1}{\kappa}-1}$. Therefore, the growth in time of both gradient processes coincides.

PROOF OF THEOREM 5.3.7. In the sequel, we use the following inequality without mentioning it again:

$$\frac{|\mathrm{e}^{\eta x}-1|}{\eta} \leqslant \mathrm{e}^{\eta}|x|, \quad \forall |x|\leqslant 1, \eta\in[0,1],$$

where $\frac{|\mathrm{e}^{0x}-1|}{0} := \lim_{\eta\downarrow 0}\frac{|\mathrm{e}^{\eta x}-1|}{\eta} = |x|$.

(1) Since $\sigma>0$ and $\int_{|x|>1}\mathrm{e}^{2x}\nu(\mathrm{d}x)<\infty$, Proposition 5.3.6(1) implies that

$$|P_t g(z)-P_t g(y)| \leqslant c_{(5.3.3)}\left(\left(t^{\frac{\eta-1}{2}}\frac{|z^\eta-y^\eta|}{\eta}\right)\wedge|z-y|^\eta\right) \tag{5.3.7}$$

for all $z>0$, $y>0$, $t\in(0,T]$. Moreover, since $P_t g\in C^\infty(\mathbb{R}_+)$ due to $\sigma>0$, we divide both side of (5.3.7) by $|z-y|$ and then let $z\to y$ to obtain that

$$|\partial_y P_t g(y)| \leqslant c_{(5.3.3)} t^{\frac{\eta-1}{2}} y^{\eta-1}, \quad \forall (t,y)\in(0,T]\times\mathbb{R}_+.$$

Hence, we separate $\int_{\mathbb{R}} = \int_{|x|\leqslant 1}+\int_{|x|>1}$ and apply (5.3.7) with $z=y\mathrm{e}^x$ to obtain

$$\begin{aligned}|\Gamma_\ell(t,y)| \leqslant{}& c_{(5.3.3)}\left(\sigma^2+\int_{|x|\leqslant 1}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}|\mathrm{e}^x-1|\ell(\mathrm{d}x)\right)t^{\frac{\eta-1}{2}}y^{\eta-1}\\ &+c_{(5.3.3)}y^{\eta-1}\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x).\end{aligned} \tag{5.3.8}$$

Since $0<\sigma^2+\int_{|x|\leqslant 1}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}|\mathrm{e}^x-1|\ell(\mathrm{d}x)\leqslant\sigma^2+\mathrm{e}^{\eta+1}\int_{|x|\leqslant 1}|x|^2\ell(\mathrm{d}x)<\infty$ and $\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)<\infty$, together with $\inf_{t\in(0,T]}t^{\frac{\eta-1}{2}}>0$, the second term on the right-hand side of (5.3.8) can be upper bounded by the first term up to a positive constant. Hence, the desired conclusion follows.

(2) One has $\mathrm{e}^{-t\psi(-\eta\mathrm{i})}=\mathbb{E}\mathrm{e}^{\eta X_t}<\infty$ for $t>0$. The Hölder continuity of g implies that $|P_t g(\mathrm{e}^x y)-P_t g(y)|\leqslant|g|_{C^{0,\eta}(\mathbb{R}_+)}\mathbb{E}\mathrm{e}^{\eta X_t}|\mathrm{e}^x-1|^\eta y^\eta$, and hence

$$\begin{aligned}|\Gamma_\ell(t,y)| &\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\mathbb{E}\mathrm{e}^{\eta X_t}y^{\eta-1}\int_{\mathbb{R}}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\\ &\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\mathrm{e}^{T|\psi(-\eta\mathrm{i})|}\left(\mathrm{e}^{\eta+1}\int_{|x|\leqslant 1}|x|^{\eta+1}\ell(\mathrm{d}x)+\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\right)y^{\eta-1},\end{aligned}$$

which implies the assertion.

(3) Let $t\in(0,T]$ and $y>0$. We separate $\int_{\mathbb{R}}=\int_{|x|\leqslant 1}+\int_{|x|>1}$, and then apply Proposition 5.3.6(2a) with $z=y\mathrm{e}^x$ to obtain

$$\begin{aligned}|\Gamma_\ell(t,y)| &\leqslant c_{(5.3.3)}y^{\eta-1}\Bigg(\int_{|x|\leqslant 1}\left(\left(t^{\frac{\eta-1}{\alpha}}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}\right)\wedge|\mathrm{e}^x-1|^\eta\right)|\mathrm{e}^x-1|\ell(\mathrm{d}x)\\ &\qquad\qquad+\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\Bigg)\\ &\leqslant c_{(5.3.3)}y^{\eta-1}\Bigg(\mathrm{e}^{\eta+1}t^{\frac{\eta+1}{\alpha}}\int_{|x|\leqslant 1}\left(\left|\frac{x}{t^{1/\alpha}}\right|^2\wedge\left|\frac{x}{t^{1/\alpha}}\right|^{\eta+1}\right)\ell(\mathrm{d}x)\\ &\qquad\qquad+\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\Bigg)\end{aligned}$$

$$\leqslant c_{(5.3.3)} y^{\eta-1} \left(c_{(5.3.9)} t^{\frac{\eta+1-\beta}{\alpha}} + \int_{|x|>1} |e^x - 1|^{\eta+1} \ell(dx) \right), \tag{5.3.9}$$

where $c_{(5.3.9)} := e^{\eta+1}(T^{\frac{\beta-2}{\alpha}} \vee T^{\frac{\beta-\eta-1}{\alpha}}) \sup_{r\in(0,1]} r^\beta \int_{|x|\leqslant 1} (|\frac{x}{r}|^2 \wedge |\frac{x}{r}|^{\eta+1}) \ell(dx) \in (0,\infty)$ by (5.3.6). Since $\inf_{(t,\beta)\in(0,T]\times(1+\eta,2]} t^{\frac{\eta+1-\beta}{\alpha}} > 0$, the desired conclusion follows from (5.3.9).

(4) Let $t \in (0,T]$ and $y > 0$. We apply Proposition 5.3.6(2b) with $z = ye^x$ and use the same argument as in the proof of item (3) to obtain

$$\begin{aligned} |\Gamma_\ell(t,y)| &\leqslant c_{(5.3.3)} y^{\eta-1} \Bigg(\int_{|x|\leqslant 1} \left(\left(t^{\frac{\eta-1}{\alpha}} |x|^{1-\eta} |e^x - 1|^\eta \right) \wedge |e^x - 1|^\eta \right) |e^x - 1| \ell(dx) \\ &\qquad\qquad + \int_{|x|>1} |e^x - 1|^{\eta+1} \ell(dx) \Bigg) \\ &\leqslant c_{(5.3.3)} y^{\eta-1} \left(c_{(5.3.9)} t^{\frac{\eta+1-\beta}{\alpha}} + \int_{|x|>1} |e^x - 1|^{\eta+1} \ell(dx) \right). \end{aligned}$$

Again, a similar argument as in the one after inequality (5.3.9) yields the assertion. $\square$

CHAPTER 6

Hedging in exponential Lévy models: The martingale setting

6.1. Introduction

We consider two types of risk for hedging an option in exponential Lévy models. The first type comes from the incompleteness of the market and the second one is due to the impossibility of continuously rebalancing a hedging portfolio. We now briefly discuss the first type of hedging error.

It is known that exponential Lévy models correspond to incomplete markets in general, there is no hedging strategy which is self-financing and replicates an option at maturity. Therefore, one has to look for certain strategies that minimize some types of risk. In this thesis (Chapters 6 and 7), we choose the quadratic hedging approach which is a popular method to deal with the problem in models with jumps. We refer the reader to the survey article of Schweizer [56] for this approach.

Two typical types of quadratic hedging strategies are the *mean-variance hedging* (MVH) strategies and the *local risk-minimizing* (LRM) strategies. Roughly speaking, the MVH strategy is self-financing and minimizes the *global* hedging error in the mean square sense, while the LRM strategy is mean-self-financing, replicates an option at maturity and minimizes the riskiness of the cost process *locally in time*.

Two cases are considered:

• The martingale setting: The driving process is assumed to be a (local) martingale. In this case, the MVH strategy and the LRM strategy can be determined via the *Galtchouk–Kunita–Watanabe (GKW) decomposition*. We deal with the martingale case in Chapter 6 and establish an explicit form for the GKW decomposition.

• The semimartingale setting: The driving process is assumed to be a semimartingale. Both types of those strategies are intimately related to the so-called *Föllmer–Schweizer (FS) decomposition*, which is an extension of the GKW decomposition in the semimartingale framework. Namely, in our (exponential Lévy) setting, the FS decomposition gives directly the LRM strategy, and the MVH strategy can be determined in a *feedback form* based on this decomposition (see Schweizer [54, Theorem 3]). We discuss in Chapter 7 the semimartingale case and provide an explicit form for the FS decomposition.

6.2. Galtchouk–Kunita–Watanabe (GKW) decomposition and explicit MVH strategies

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of X. We let $\mathcal{F} = \mathcal{F}_T$ and $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$. The following assumption is imposed in this section:

Assumption 6.2.1. Assume that the exponential $S = \mathrm{e}^X$ is a *square integrable* $\mathbb{P}$*-martingale*. To prevent the triviality, let us assume in addition that X is not a.s. deterministic.

Let us define the family of *admissible strategies* as

$$\mathcal{AS}(\mathbb{P}) := \left\{\vartheta \text{ predictable } : \mathbb{E}\int_0^T \vartheta_t^2 S_{t-}^2 \mathrm{d}t < \infty\right\}.$$

Definition 6.2.2 (Schweizer [56], p.546)**.**

(1) The *GKW decomposition* of an $H \in L_2(\mathbb{P})$ is of the form

$$H = \mathbb{E}H + \int_0^T \vartheta_t^{\mathrm{GKW}}(H)\mathrm{d}S_t + L_T^{\mathrm{GKW}}, \tag{6.2.1}$$

where $\vartheta^{\mathrm{GKW}}(H) \in \mathcal{AS}(\mathbb{P})$, L^{GKW} is an $L_2(\mathbb{P})$-martingale with zero mean and is strongly orthogonal to S, i.e. the product SL^{GKW} is a local martingale.

(2) The integrand $\vartheta^{\mathrm{GKW}}(H)$ is called the *MVH strategy* of H.

According to the GKW decomposition, it turns out that the MVH strategy $\vartheta^{\mathrm{GKW}}(H)$ is the minimizer, which is unique up to a $\mathbb{P} \otimes \lambda$-null set, for the problem

$$\min_{\vartheta \in \mathcal{AS}(\mathbb{P})} \mathbb{E}\left|H - \mathbb{E}H - \int_0^T \vartheta_t \mathrm{d}S_t\right|^2.$$

Our aim is to apply the approximation results obtained in Chapter 4 for the stochastic integral term in (6.2.1), which can be interpreted in mathematical finance as the hedgeable part of H. To do that, one of the main tasks for us is to find a representation of $\vartheta^{\mathrm{GKW}}(H)$ which is convenient for verifying the conditions in Assumption 4.3.3. This issue is handled in the next subsection in which we focus on the European type options $H = g(S_T)$.

In the literature, there are several methods to determine an explicit form for the MVH strategy of a European type option $H = g(S_T)$. Let us briefly discuss 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., Jacob, Méléard and Protter [36], Cont, Tankov and Voltchkova [14]) 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., Brodén and Tankov [9], Tankov [58]). 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., Benth et al. [6], Løkka [41]), 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, the result 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.

Theorem 6.2.3 (Explicit MVH strategies)**.** *For a Borel function $g\colon\mathbb{R}_+ \to \mathbb{R}$ with $g(S_T) \in L_2(\mathbb{P})$, there is a $\tilde{\vartheta}^{\mathrm{GKW}}(g) \in \mathrm{CL}([0,T))$ such that the following assertions hold:*

(1) *$\tilde{\vartheta}_-^{\mathrm{GKW}}(g)$ is a MVH strategy of $g(S_T)$.*

(2) *$\tilde{\vartheta}^{\mathrm{GKW}}(g)S$ is an $L_2(\mathbb{P})$-martingale and $\Delta\tilde{\vartheta}_t^{\mathrm{GKW}}(g) = 0$ a.s. for each $t \in [0,T)$.*

(3) *For any $t \in (0,T)$, a.s.,*

$$\tilde{\vartheta}_t^{\mathrm{GKW}}(g) = \frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2\partial_y G(t,S_t) + \int_{\mathbb{R}} \frac{G(t,\mathrm{e}^x S_t) - G(t,S_t)}{S_t}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right), \tag{6.2.2}$$

where $\|(\sigma,\nu)\| := \sigma^2 + \int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x) \in (0,\infty)$, 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.*

Formula (6.2.2) was also given in [14, Section 4] and in [58, Proposition 7] under some extra conditions for g and S. A similar formula as (6.2.2) in a general framework can be found in [36, Theorem 2.4].

PROOF OF THEOREM 6.2.3. We only sketch here the main idea of the proof, and the reader is referred to [60, Section 6] for more details.

Step 1. According to Subsection 5.1.3, the exponential $S = \mathrm{e}^X$ satisfies the SDE

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

where Z is another Lévy process. Since S is an $L_2(\mathbb{P})$-martingale due to assumption, it implies that Z is also an $L_2(\mathbb{P})$-martingale with zero mean.

Step 2. For each $t \in (0,T)$ one determines the chaos expansion of $\vartheta_t^{\mathrm{GKW}}(g)S_t$ with respect to the Lévy process Z in the way introduced by C. Geiss, S. Geiss and Laukkarinen [22].

Step 3. We translate this chaos expansion (with respect to Z) to an expansion with respect to X from which one can use Proposition 5.1.2. □

6.3. Weight regularity

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process with $(X|\mathbb{P}) \sim (\gamma,\sigma,\nu)$ and ψ the characteristic exponent. One also remarks that Assumption 6.2.1 is not necessarily satisfied in this section.

For $\eta \in [0,1]$, we define the processes $\Theta(\eta)$ and $\Phi(\eta)$ by setting

$$\Theta(\eta)_t = \sup_{u\in[0,t]}(S_u^{\eta-1}), \quad \Phi(\eta)_t = \Theta(\eta)_t S_t, \quad t \in [0,T]. \tag{6.3.1}$$

We will see later that $\Theta(\eta)$ and $\Phi(\eta)$ appear as the weight processes in the discrete-time approximation using the mean-variance hedging strategies (Theorem 6.4.1) or the local risk-minimizing strategies (Theorem 7.3.1).

Proposition 6.3.1. *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,*

$$\|\Phi(\eta)\|^q_{\mathcal{SM}_q(\mathbb{P})} \leqslant \mathrm{e}^{T|\psi(-\mathrm{i})|(2q+1)}2^{1-\eta}\left(\frac{q}{q-1}\right)^{2q}\|S_T\|^q_{L_q(\mathbb{P})}.$$

PROOF. The first step considers the particular case when S is a martingale, and the general case is handled in the second step.

Step 1. Assume that S is a $\mathbb{P}$-martingale. Due to Sato [51, Theorem 25.3], the assumption $\int_{|x|>1} \mathrm{e}^{qx}\nu(\mathrm{d}x) < \infty$ implies that $\mathrm{e}^{X_t} \in L_q(\mathbb{P})$ for all $t > 0$. Denote $c_q := (\frac{q}{q-1})^q$ and define $M = (M_t)_{t\in[0,T]}$ by

$$M_t := \sup_{u\in[0,t]} \mathrm{e}^{X_t - X_u}.$$

We show that M is a positive $L_q(\mathbb{P})$-submartingale. The adaptedness and positivity are clear. Pick a $t \in (0,T]$. Since $(X_t - X_{t-u})_{u\in[0,t]}$ is *càglàd* (left-continuous with right limits) and $(X_u)_{u\in[0,t]}$ is càdlàg, and both processes have the same finite-dimensional distribution, applying Doob's maximal inequality yields

$$\begin{aligned}\mathbb{E}M_t^q &= \mathbb{E}\Big[\sup_{u\in[0,t]} \mathrm{e}^{q(X_t - X_u)}\Big] = \mathbb{E}\Big[\sup_{u\in[0,t]} \mathrm{e}^{q(X_t - X_{t-u})}\Big] \\ &= \mathbb{E}\big[\sup_{u\in[0,t]} \mathrm{e}^{qX_u}\big] \leqslant c_q \mathbb{E}\mathrm{e}^{qX_t} < \infty.\end{aligned} \tag{6.3.2}$$

For $0 \leqslant s \leqslant t \leqslant T$ one has, a.s.,

$$\mathbb{E}_{\mathcal{F}_s}[M_t] \geqslant \mathbb{E}_{\mathcal{F}_s}\big[\sup_{u\in[0,s]} \mathrm{e}^{X_t - X_u}\big] = \sup_{u\in[0,s]} \mathrm{e}^{X_s - X_u}\mathbb{E}\mathrm{e}^{X_t - X_s} = M_s,$$

where we use $\mathbb{E}e^{X_t - X_s} = \mathbb{E}S_{t-s} = 1$.

We observe that the process $\Phi(\eta)$ can be re-written as

$$\Phi(\eta)_t = e^{\eta X_t} \sup_{s\in[0,t]} e^{(1-\eta)(X_t - X_s)} = e^{\eta X_t} M_t^{1-\eta}.$$

Let us fix $\eta \in (0,1)$ and $a \in [0,T]$. For $e^{\eta X} = (e^{\eta X_t})_{t\in[0,T]}$, applying Doob's maximal inequality and Jensen's inequality we obtain that, a.s.,

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_a}\Big[\sup_{t\in[a,T]}(e^{\eta X_t})^{\frac{q}{\eta}}\Big] &= e^{qX_a}\mathbb{E}\Big[\sup_{t\in[a,T]} e^{q(X_t - X_a)}\Big] \leqslant c_q e^{qX_a}\mathbb{E}e^{q(X_T - X_a)}\\ &= c_q e^{qX_a}\mathbb{E}e^{qX_{T-a}} \leqslant c_q e^{qX_a}\mathbb{E}e^{qX_T},\end{aligned}$$

which implies

$$\|e^{\eta X}\|_{\mathcal{SM}_{q/\eta}(\mathbb{P})} \leqslant (c_q \mathbb{E}e^{qX_T})^{\frac{\eta}{q}}.$$

For $M^{1-\eta} = (M_t^{1-\eta})_{t\in[0,T]}$, one has that, a.s.,

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_a}\Big[\sup_{t\in[a,T]}(M_t^{1-\eta})^{\frac{q}{1-\eta}}\Big] &= \mathbb{E}_{\mathcal{F}_a}\big[\sup_{t\in[a,T]} M_t^q\big] \leqslant c_q \mathbb{E}_{\mathcal{F}_a}\big[M_T^q\big]\\ &\leqslant c_q\mathbb{E}_{\mathcal{F}_a}\Big[\sup_{s\in[0,a]} e^{q(X_T - X_s)}\Big] + c_q\mathbb{E}_{\mathcal{F}_a}\Big[\sup_{s\in[a,T]} e^{q(X_T - X_s)}\Big]\\ &= c_q \sup_{s\in[0,a]} e^{q(X_a - X_s)}\mathbb{E}e^{q(X_T - X_a)} + c_q\mathbb{E}\Big[\sup_{s\in[a,T]} e^{q(X_T - X_s)}\Big]\\ &\leqslant 2c_q \sup_{s\in[0,a]} e^{q(X_a - X_s)}\mathbb{E}\Big[\sup_{s\in[a,T]} e^{q(X_T - X_s)}\Big]\\ &\leqslant \Big(2c_q\mathbb{E}\Big[\sup_{s\in[0,T]} e^{q(X_T - X_s)}\Big]\Big) M_a^q\\ &\leqslant \Big(2c_q^2\mathbb{E}e^{qX_T}\Big) M_a^q,\end{aligned}$$

where the conditional Doob maximal inequality is applied for the positive sub-martingale M to obtain the first inequality, and the last one comes from (6.3.2). Hence,

$$\|M^{1-\eta}\|_{\mathcal{SM}_{q/(1-\eta)}(\mathbb{P})} \leqslant (2c_q^2\mathbb{E}e^{qX_T})^{\frac{1-\eta}{q}}.$$

Applying a version of Hölder's inequality for $\|\cdot\|_{\mathcal{SM}_q(\mathbb{P})}$ given in [29, Proposition A.2] with $\frac{1}{q} = \frac{1}{q/\eta} + \frac{1}{q/(1-\eta)}$, we obtain

$$\|\Phi(\eta)\|_{\mathcal{SM}_q(\mathbb{P})} \leqslant \|e^{\eta X}\|_{\mathcal{SM}_{q/\eta}(\mathbb{P})}\|M^{1-\eta}\|_{\mathcal{SM}_{q/(1-\eta)}(\mathbb{P})} \leqslant 2^{\frac{1-\eta}{q}}\left(\frac{q}{q-1}\right)^2 \|S_T\|_{L_q(\mathbb{P})} < \infty,$$

which asserts $\Phi(\eta) \in \mathcal{SM}_q(\mathbb{P})$. When $\eta = 0$ or $\eta = 1$, the desired conclusion is straightforward as $\Phi(0) = M$, $\Phi(1) = e^X$.

Step 2. In the general case, we define

$$\widetilde{S}_t := e^{t\psi(-\mathrm{i})} S_t.$$

Then it is known that $\widetilde{S}$ is a martingale under $\mathbb{P}$. Some standard calculations yield

$$e^{-T|\psi(-\mathrm{i})|}\widetilde{\Phi}(\eta)_t \leqslant \Phi(\eta)_t \leqslant e^{T|\psi(-\mathrm{i})|}\widetilde{\Phi}(\eta)_t,$$

where $\widetilde{\Phi}(\eta)_t := \widetilde{S}_t \sup_{u\in[0,t]}(\widetilde{S}_u^{\eta-1})$. Applying *Step 1* for $\mathbb{P}$-martingale $\widetilde{S}$ we derive that $\widetilde{\Phi}(\eta) \in \mathcal{SM}_q(\mathbb{P})$. Hence, for $a \in [0,T]$, one has, a.s.,

$$\mathbb{E}_{\mathcal{F}_a}\big[\sup_{t\in[a,T]} \Phi(\eta)_t^q\big] \leqslant e^{qT|\psi(-\mathrm{i})|}\mathbb{E}_{\mathcal{F}_a}\big[\sup_{t\in[a,T]} \widetilde{\Phi}(\eta)_t^q\big]$$

$$\leqslant \mathrm{e}^{qT|\psi(-\mathrm{i})|}\|\widetilde{\Phi}(\eta)\|^q_{\mathcal{SM}_q(\mathbb{P})}\widetilde{\Phi}(\eta)^q_a$$
$$\leqslant \mathrm{e}^{2qT|\psi(-\mathrm{i})|}2^{1-\eta}\left(\frac{q}{q-1}\right)^{2q}\|\widetilde{S}_T\|^q_{L_q(\mathbb{P})}\Phi(\eta)^q_a$$
$$\leqslant \mathrm{e}^{T|\psi(-\mathrm{i})|(2q+1)}2^{1-\eta}\left(\frac{q}{q-1}\right)^{2q}\|S_T\|^q_{L_q(\mathbb{P})}\Phi(\eta)^q_a,$$

which proves the desired conclusion. □

6.4. Discretisation of MVH strategies in the martingale setting

As mentioned earlier, Assumption 4.3.3 is crucial to obtain approximation results for the jump adjusted method in weighted BMO spaces. We now provide examples for Assumption 4.3.3 in the exponential Lévy model using the MVH strategies given in Theorem 6.2.3. Once Assumption 4.3.3 is satisfied, one may apply Theorems 4.3.4 and 4.3.7 to derive the results accordingly. We recall $\mathring{W}^{1,q}(\mathbb{R}_+)$ from Definition 5.3.1, $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ from Definition 5.3.2, and $\Theta(\eta)$, $\Phi(\eta)$ from (6.3.1).

Theorem 6.4.1. *Assume Assumption 6.2.1. Let $\eta \in [0,1]$. Then the following assertions hold:*

(1) (*MVH strategy growth*) *If $g \in C^{0,\eta}(\mathbb{R}_+)$, then there exist a $\hat{\theta} \in [0,1]$ and a constant $c_{(6.4.1)} > 0$, which might depend on $\hat{\theta}$, such that for $\tilde{\vartheta}^{\mathrm{GKW}}(g)$ given in* (6.2.2) *one has*

$$|\tilde{\vartheta}^{\mathrm{GKW}}_t(g)| \leqslant c_{(6.4.1)}(T-t)^{\frac{\hat{\theta}-1}{2}}\Theta(\eta)_t \quad a.s., \ \forall t \in [0,T), \tag{6.4.1}$$

where $\hat{\theta}$ is provided in Table 6.1:

Table 6.1: Values of $\hat{\theta}$

	σ and η	*Small jump condition for X*	*Regularity of g*	*Conclusion for $\hat{\theta}$*
C1	$\sigma > 0$ $\eta \in [0,1]$		$g \in C^{0,\eta}(\mathbb{R}_+)$	$\hat{\theta} = \eta$
C2	$\sigma = 0$ $\eta \in [0,1]$	$\int_{\lvert x\rvert\leqslant 1}\lvert x\rvert^{1+\eta}\nu(\mathrm{d}x) < \infty$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\hat{\theta} = 1$
C3	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_1(\alpha)$ *for some* $\alpha \in [1+\eta, 2)$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\forall \hat{\theta} \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$
C4	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_2(\alpha)$ *for some* $\alpha \in [1+\eta, 2)$	$g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$	$\forall \hat{\theta} \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$

(2) *Denote $M := \tilde{\vartheta}^{\mathrm{GKW}}(g)S$. Then Assumption 4.3.3 is satisfied for*

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

and for $\theta = 1$ if $\hat{\theta} = 1$, and for any $\theta \in (0,\hat{\theta})$ if $\hat{\theta} \in (0,1)$.

Remark 6.4.2. In Table 6.1, the larger $\hat{\theta}$ is, the better estimate one can get for $\tilde{\vartheta}^{\mathrm{GKW}}(g)$. Moreover, the parameter $\hat{\theta}$ comes from the interplay between the small jump intensity of the underlying Lévy process and the regularity of the payoff function which affects the convergence rate of the approximation error.

PROOF OF THEOREM 6.4.1. We recall from Assumption 6.2.1 that $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$.

(1) We let $\ell := \nu$ in (5.3.4) and obtain from (6.2.2) that

$$\tilde{\vartheta}_t^{\mathrm{GKW}}(g) = \|(\sigma,\nu)\|^{-1}\varGamma_\nu(T-t,S_t) \quad \text{a.s., } \forall t\in[0,T).$$

We consider each case in Table 6.1 as follows. We apply Theorem 5.3.7(1) to get C1. The case C2 follows from Theorem 5.3.7(2). For C3, since $\nu\in\mathcal{S}_1(\alpha)$, Remark 5.3.3(2) implies that $0<\int_{|x|\leqslant 1}|x|^{\alpha+\varepsilon}\nu(\mathrm{d}x)<\infty$ for any $\varepsilon\in(0,2-\alpha]$. Moreover, applying Theorem 5.3.7(3) and Remark 5.3.8 with $\beta=\alpha+\varepsilon$ yields

$$|\tilde{\vartheta}_t^{\mathrm{GKW}}(g)| \leqslant c(\varepsilon)(T-t)^{\frac{\eta+1}{\alpha}-1-\frac{\varepsilon}{\alpha}}S_t^{\eta-1} \leqslant c(\varepsilon)(T-t)^{\frac{1}{2}\left(\left(\frac{2(\eta+1)}{\alpha}-1-\frac{2\varepsilon}{\alpha}\right)-1\right)}\Theta(\eta)_t \text{ a.s., } \forall t\in[0,T),$$

where $c(\varepsilon)>0$ is a constant depending on ε. Since $\varepsilon>0$ can be arbitrarily small, C3 follows. The case C4 is similar to C3 where we use Theorem 5.3.7(4) and Remark 5.3.8.

(2) According to Theorem 6.2.3(2), M is an $L_2(\mathbb{P})$-martingale. Then Assumption 4.3.1 holds because of Example 4.3.2. We now only need to check (4.3.3). If $\hat{\theta}=1$, then the martingale M is closed by $M_T := L_2(\mathbb{P})\text{-}\lim_{t\uparrow T} M_t$ due to (6.4.1) and $\Phi(\eta)\in\mathcal{SM}_2(\mathbb{P})$. Then for $\theta=1$ and for any $a\in[0,T)$ one has, a.s.,

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}\Upsilon(\cdot,\mathrm{d}t)\right] &= \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}\mathrm{d}\langle M,M\rangle_t + \int_{(a,T)} M_t^2\mathrm{d}t\right]\\ &\leqslant \mathbb{E}_{\mathcal{F}_a}\left[|M_T-M_a|^2 + c_{(6.4.1)}^2(T-a)\sup_{t\in(a,T)}\Phi(\eta)_t^2\right]\\ &\leqslant c_{(6.4.1)}^2(T+1)\|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2\Phi(\eta)_a^2.\end{aligned}$$

If $\hat{\theta}\in(0,1)$, then for any $\theta\in(0,\hat{\theta})$ and any $a\in[0,T)$ one has, a.s.,

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}(T-t)^{1-\theta}M_t^2\mathrm{d}t\right] \leqslant c_{(6.4.1)}^2 T^{\hat{\theta}-\theta+1}\|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2\Phi(\eta)_a^2. \tag{6.4.2}$$

We apply conditional Itô's isometry and Proposition 2.3.2(2.3.1) to obtain that, a.s.,

$$\begin{aligned}&\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}(T-t)^{1-\theta}\mathrm{d}\langle M,M\rangle_t\right] = \lim_{b\uparrow T}\mathbb{E}_{\mathcal{F}_a}\left[\left|\int_{(a,b]}(T-t)^{\frac{1-\theta}{2}}\mathrm{d}M_t\right|^2\right]\\ &\leqslant (1-\theta)\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)}(T-t)^{-\theta}M_t^2\mathrm{d}t\right] \leqslant (1-\theta)c_{(6.4.1)}^2\|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2\Phi(\eta)_a^2\int_{(a,T)}(T-t)^{\hat{\theta}-\theta-1}\mathrm{d}t\\ &\leqslant \frac{T^{\hat{\theta}-\theta}}{\hat{\theta}-\theta}(1-\theta)c_{(6.4.1)}^2\|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2\Phi(\eta)_a^2. \end{aligned}\tag{6.4.3}$$

Combining (6.4.2) with (6.4.3) yields the desired conclusion. □

CHAPTER 7

Hedging in exponential Lévy models: The semimartingale setting

7.1. Föllmer–Schweizer (FS) decomposition

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of X. We assume that $\mathcal{F} = \mathcal{F}_T$ and let $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$.

Assume that X is not a.s. deterministic and that $S = \mathrm{e}^X$ is square integrable under $\mathbb{P}$. Then S is a semimartingale with the representation (due to Itô's formula)

$$\begin{aligned} S &= 1 + \left(\int_0^\cdot \sigma S_{t-}\mathrm{d}W_t + \int_0^\cdot\int_{\mathbb{R}_0} S_{t-}(\mathrm{e}^x - 1)\widetilde{N}(\mathrm{d}t, \mathrm{d}x)\right) + \int_0^\cdot \gamma_S S_{t-}\mathrm{d}t \\ &=: 1 + S^{\mathrm{m}} + S^{\mathrm{fv}}, \end{aligned}$$

where S^{m} and S^{fv} respectively denote the martingale part and the predictable finite variation part in the representation of S. We denote

$$\gamma_S := \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}}(\mathrm{e}^x - 1 - x\mathbb{1}_{\{|x|\leqslant 1\}})\nu(\mathrm{d}x), \tag{7.1.1}$$

and use again the notation

$$\|(\sigma,\nu)\| := \sigma^2 + \int_{\mathbb{R}}(\mathrm{e}^x - 1)^2\nu(\mathrm{d}x) \in (0,\infty). \tag{7.1.2}$$

7.1.1. FS decomposition. We briefly present the FS decomposition of a random variable and the notion of minimal local martingale measure which is the key tool to determine the FS decomposition. The reader is referred to Schweizer [56] for more information about these objects.

In the exponential Lévy setting, we follow Hubalek, Kallsen and Krawczyk [34, p.863] and use the family of *admissible strategies* as

$$\mathcal{AS}(\mathbb{P}) := \left\{\vartheta \text{ predictable} : \mathbb{E}\int_0^T \vartheta_t^2 S_{t-}^2\mathrm{d}t < \infty\right\}.$$

It turns out that if $\vartheta \in \mathcal{AS}(\mathbb{P})$, then

$$\begin{aligned} \mathbb{E}\int_0^T \vartheta_t^2\mathrm{d}[S,S]_t &= \mathbb{E}\int_0^T \vartheta_t^2\mathrm{d}[S^{\mathrm{m}}, S^{\mathrm{m}}]_t \\ &= \mathbb{E}\int_0^T \vartheta_t^2\mathrm{d}\langle S^{\mathrm{m}}, S^{\mathrm{m}}\rangle_t = \|(\sigma,\nu)\|\mathbb{E}\int_0^T \vartheta_t^2 S_{t-}^2\mathrm{d}t < \infty. \end{aligned} \tag{7.1.3}$$

Definition 7.1.1 (Schweizer [56], *FS decomposition and LRM strategy*)**.**

(1) An $H \in L_2(\mathbb{P})$ admits a *FS decomposition* if H can be written as

$$H = H_0 + \int_0^T \vartheta_t^{\mathrm{FS}}(H)\mathrm{d}S_t + L_T^{\mathrm{FS}},$$

where $H_0 \in \mathbb{R}$, $\vartheta^{\mathrm{FS}}(H) \in \mathcal{AS}(\mathbb{P})$, and where L^{FS} is a square integrable $\mathbb{P}$-martingale starting at zero and L^{FS} is $\mathbb{P}$-orthogonal to the martingale part S^{m} of S.

(2) The integrand $\vartheta^{\mathrm{FS}}(H)$ is called the *local risk-minimizing* strategy of H.

We remark that in the exponential Lévy setting, S satisfies the *structure condition*, and the *mean-variance trade-off process* $\widehat{K}$ of S in the sense of [56, p.553] is

$$\widehat{K}_t = \frac{\gamma_S^2}{\|(\sigma,\nu)\|}t,$$

which is uniformly bounded in $(\omega,t) \in \Omega \times [0,T]$. Hence, any $H \in L_2(\mathbb{P})$ admits a unique FS decomposition (see Monat and Stricker [44, Theorem 3.4]).

The original definition of LRM strategies is quite involved and it was shown in [56, Theorem 3.3 and Proposition 3.4] that the LRM strategy of an $H \in L_2(\mathbb{P})$ can be determined via the FS decomposition of H. In fact, the LRM strategy of H is the pair $(\vartheta^{\mathrm{FS}}(H), \eta^H)$ (see [56, p.553]), where $\vartheta^{\mathrm{FS}}(H)$ is the integrand of the integral term in the FS decomposition of H, and η^H is determined by $\eta^H = H_0 + \int_0^\cdot \vartheta_u^{\mathrm{FS}}(H)\mathrm{d}S_u + L^{\mathrm{FS}} - \vartheta^{\mathrm{FS}}(H)S$. Since η^H can be computed by knowing $\vartheta^{\mathrm{FS}}(H)$, we identify $\vartheta^{\mathrm{FS}}(H)$ with the LRM strategy of H.

We continue with the notion of minimal local martingale measure.

Definition 7.1.2 (Schweizer [55], Section 2)**.** Let $\mathcal{E}(U) \in \mathrm{CL}([0,T])$ be the stochastic exponential of U, i.e. $\mathrm{d}\mathcal{E}(U) = \mathcal{E}(U)_-\mathrm{d}U$ with $\mathcal{E}(U)_0 = 1$, where

$$U = -\frac{\gamma_S}{\|(\sigma,\nu)\|}\left(\sigma W + \int_0^\cdot\int_{\mathbb{R}_0}(\mathrm{e}^x - 1)\widetilde{N}(\mathrm{d}s,\mathrm{d}x)\right). \tag{7.1.4}$$

If $\mathcal{E}(U) > 0$, then the probability measure $\mathbb{P}^*$ defined by

$$\mathrm{d}\mathbb{P}^* = \mathcal{E}(U)_T\mathrm{d}\mathbb{P}$$

is called the *minimal local martingale measure* for S.

The following assumption, which is imposed on the characteristics of X, ensures that $\mathcal{E}(U) > 0$, and hence $\mathbb{P}^*$ exists:

Assumption 7.1.3. $\gamma_S(\mathrm{e}^x - 1) < \|(\sigma,\nu)\|$ for all $x \in \mathrm{supp}(\nu)$.

Remark that a sufficient condition for Assumption 7.1.3 is

$$0 \geqslant \gamma_S \geqslant -\|(\sigma,\nu)\|.$$

Assume that Assumption 7.1.3 holds true. Then by an application of Girsanov's theorem (see, e.g., Esche and Schweizer [17, Propositions 2 and 3]), X is also a Lévy process under $\mathbb{P}^*$ with $(X|\mathbb{P}^*) \sim (\gamma^*, \sigma^*, \nu^*)$, where

$$\sigma^* = \sigma \quad \text{and} \quad \nu^*(\mathrm{d}x) = \left(1 - \frac{\gamma_S(\mathrm{e}^x - 1)}{\|(\sigma,\nu)\|}\right)\nu(\mathrm{d}x).$$

Moreover, if W^* and $\widetilde{N}^*$ are the standard Brownian motion and the compensated Poisson random measure of X under $\mathbb{P}^*$, then

$$W_t^* = W_t + \frac{\gamma_S\sigma}{\|(\sigma,\nu)\|}t, \tag{7.1.5}$$

$$\widetilde{N}^*(\mathrm{d}t,\mathrm{d}x) = \widetilde{N}(\mathrm{d}t,\mathrm{d}x) + \frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\mathrm{d}t. \tag{7.1.6}$$

7.2. Explicit LRM strategies

There are many works interested in finding an explicit representation for the FS decomposition and for the LRM strategy in the semimartingale framework, for example, see Hubalek, Kallsen and Krawczyk [34], Goutte, Oudjane and Russo [32], Kallsen and Pauwels [38], Tankov [58].

In the exponential Lévy setting and in the case of a European type option $H = g(S_T)$, Hubalek, Kallsen and Krawczyk [34] assumed that the function g can be represented as an integral transform of finite complex measures from which one can determine a closed form expression for the LRM strategy ([34, Proposition 3.1]). The key idea of this approach is the separation of the function g and the underlying price process S by using a kind of inverse Fourier transform. An advantage of this method is that one gains much flexibility for choosing the underlying Lévy process where there is no extra regularity required for the driving process S except some mild integrability.

As our first main result, Theorem 7.2.1 below provides a closed form for the LRM strategy $\vartheta^{\mathrm{FS}}(H)$ of an $H = g(S_T)$. To obtain this result, except of some mild integrability conditions, we neither assume any regularity for the payoff function g nor require any extra condition for the small jump behavior of X. However, the price one has to pay is the condition that $\mathbb{P}^*$ exists as a true probability measure (see Assumption 7.1.3) which leads to a constraint for the characteristics of X. This result might be regarded as a counterpart of [34, Proposition 3.1] in which only the square integrability is required for S while the function g are supposed to be the integral transform of finite complex measures. The notation $\mathbb{E}^*$ below means the expectation with respect to $\mathbb{P}^*$.

Theorem 7.2.1 (Explicit LRM strategies)**.** *Under Assumption 7.1.3, if* $g\colon (0,\infty) \to \mathbb{R}$ *is a Borel function with* $\mathbb{E}^*|g(yS_t)| < \infty$ *for all* $(t,y) \in [0,T]\times(0,\infty)$ *and* $g(S_T) \in L_2(\mathbb{P}) \cap L_2(\mathbb{P}^*)$*, then the following assertions hold:*

(1) *The LRM strategy* $\vartheta^{\mathrm{FS}}(H)$ *corresponding to* $H = g(S_T)$ *is of the form*

$$\vartheta_t^{\mathrm{FS}}(H) = \frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2\partial_y G^*(t,S_{t-}) + \int_{\mathbb{R}} \frac{G^*(t,\mathrm{e}^x S_{t-}) - G^*(t,S_{t-})}{S_{t-}}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right) \tag{7.2.1}$$

for $\mathbb{P}\otimes\lambda$*-a.e.* $(\omega,t) \in \Omega\times[0,T]$*, where* $\|(\sigma,\nu)\|$ *is provided in* (7.1.2)*,* $G^*(t,y) := \mathbb{E}^* g(yS_{T-t})$*, and we set* $\partial_y G^* := 0$ *when* $\sigma = 0$ *by convention.*

(2) *There exists a process* $\tilde{\vartheta}^{\mathrm{FS}}(g) \in \mathrm{CL}([0,T))$ *such that* $\tilde{\vartheta}^{\mathrm{FS}}_{-}(g) = \vartheta^{\mathrm{FS}}(H)$ *for* $\mathbb{P}\otimes\lambda$*-a.e.* $(\omega,t) \in \Omega\times[0,T)$*, and* $\tilde{\vartheta}^{\mathrm{FS}}(g)S$ *is a* $\mathbb{P}^*$*-martingale.*

Before proving this theorem, let us comment on it. According to Theorem 7.2.1(2), $\tilde{\vartheta}^{\mathrm{FS}}_{-}(g)$ is also a LRM strategy of $H = g(S_T)$. Moreover, the càdlàg property of $\tilde{\vartheta}^{\mathrm{FS}}(g)$ is useful to design some Riemann-type approximations for $\int_0^T \tilde{\vartheta}^{\mathrm{FS}}_{t-}(g)\mathrm{d}S_t$. For example, an approximation scheme based on tracking jumps of $\tilde{\vartheta}^{\mathrm{FS}}(g)$ has been constructed in Rosenbaum and Tankov [48]. We also use this càdlàg version for the discrete-time hedging problem in the next subsection. Such a path regularity for the integrand in the martingale setting was also studied in Ma, Protter and Zhang [42].

The main tool to prove Theorem 7.2.1 is the following martingale representation:

Proposition 7.2.2. *Assume that $f:\mathbb{R}\to\mathbb{R}$ is a Borel function such that $\mathbb{E}|f(x+X_t)|<\infty$ for all $(t,x)\in[0,T]\times\mathbb{R}$. If $f(X_T)\in L_2(\mathbb{P})$, then*

$$\mathbb{E}\int_0^T|\sigma\partial_x F(t,X_{t-})|^2\mathrm{d}t+\mathbb{E}\int_0^T\int_{\mathbb{R}}|F(t,X_{t-}+x)-F(t,X_{t-})|^2\nu(\mathrm{d}x)\mathrm{d}t<\infty$$

and, a.s.,

$$f(X_T)=\mathbb{E}f(X_T)+\int_0^T\sigma\partial_x F(t,X_{t-})\mathrm{d}W_t+\int_0^T\int_{\mathbb{R}_0}(F(t,X_{t-}+x)-F(t,X_{t-}))\widetilde{N}(\mathrm{d}t,\mathrm{d}x), \tag{7.2.2}$$

where $F(t,x):=\mathbb{E}f(x+X_{T-t})$ for $(t,x)\in[0,T]\times\mathbb{R}$, and we set $\partial_x F:=0$ if $\sigma=0$.

PROOF. See the proof of [61, Proposition 1.2]. □

Proposition 7.2.2 extends [14, Proposition 7] in which the function f has a polynomial growth and X satisfies a certain condition. A similar representation to (7.2.2) in a general framework (with different assumptions from ours) can be found in the proof of [36, Theorem 2.4]. On the other hand, when $f(X_T)$ is Malliavin differentiable then one can use the Clark–Ocone formula (e.g., see Arai and Suzuki [4], Benth et al. [6], Løkka [41]) to obtain its explicit martingale representation. However, the Malliavin differentiability of $f(X_T)$ fails to hold in many contexts. For example, if $f(x)=\mathbb{1}_{[K,\infty)}(x)$ for some $K\in\mathbb{R}$, X is of infinite variation and X_T has a density satisfying a mild condition, then $f(X_T)$ is not Malliavin differentiable (see Laukkarinen [40, Theorem 6(b)]).

PROOF OF THEOREM 7.2.1. Let $f(x):=g(\mathrm{e}^x)$ and $F^*(t,x):=\mathbb{E}^*f(x+X_{T-t})$. Then we have $G^*(t,\mathrm{e}^x)=F^*(t,x)$ for $(t,x)\in[0,T]\times\mathbb{R}$. Define

$$\Delta_J G^*(t,x):=G^*(t,\mathrm{e}^x S_{t-})-G^*(t,S_{t-}),\quad (t,x)\in[0,T]\times\mathbb{R}.$$

(1) We present here a direct proof for this assertion, an alternative argument for more general settings can be found in [12, Proof of Theorem 4.3]. By assumption, $f(X_T)=g(S_T)\in L_2(\mathbb{P}^*)$ and $\mathbb{E}^*|f(x+X_t)|=\mathbb{E}^*|g(\mathrm{e}^x S_t)|<\infty$ for any $(t,x)\in[0,T]\times\mathbb{R}$, we apply Proposition 7.2.2 to obtain

$$K^*=\mathbb{E}^*g(S_T)+\int_0^\cdot\sigma S_{t-}\partial_y G^*(t,S_{t-})\mathrm{d}W_t^*+\int_0^\cdot\int_{\mathbb{R}_0}\Delta_J G^*(t,x)\widetilde{N}^*(\mathrm{d}t,\mathrm{d}x), \tag{7.2.3}$$

where $K^*=(K_t^*)_{t\in[0,T]}$ is the càdlàg version of the $L_2(\mathbb{P}^*)$-martingale $(\mathbb{E}^*_{\mathcal{F}_t}[g(S_T)])_{t\in[0,T]}$, and where W^* and $\widetilde{N}^*$ are introduced in (7.1.5) and (7.1.6). Then it holds that $\mathcal{E}(U)K^*$ is a martingale under $\mathbb{P}$. Since the $\mathbb{P}$-martingale U given in (7.1.4) satisfies that

$$\|\langle U,U\rangle_T\|_{L_\infty(\mathbb{P})}=\frac{\gamma_S^2 T}{\|(\sigma,\nu)\|^2}\left(\sigma^2+\int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x)\right)<\infty,$$

it implies that $\mathcal{E}(U)$ is regular and satisfies (R_2) in the sense of Choulli, Krawczyk and Stricker [11, Proposition 3.7]. Since $K_T^*=g(S_T)\in L_2(\mathbb{P})$ by assumption, we apply [11, Theorem 4.9((i)⇔(ii))] to obtain

$$\mathbb{E}[K^*,K^*]_T<\infty.$$

Combining this with (7.2.3) yields

$$\mathbb{E}\int_0^T\sigma^2|S_{t-}\partial_y G^*(t,S_{t-})|^2\mathrm{d}t+\mathbb{E}\int_0^T\int_{\mathbb{R}_0}|\Delta_J G^*(t,x)|^2N(\mathrm{d}t,\mathrm{d}x)=\mathbb{E}[K^*,K^*]_T<\infty.$$

Since $\mathrm{d}t\,\nu(\mathrm{d}x)$ is the predictable $\mathbb{P}$-compensator of $N(\mathrm{d}t,\mathrm{d}x)$, it implies that

$$\mathbb{E}\int_0^T \sigma^2|S_{t-}\partial_y G^*(t,S_{t-})|^2\mathrm{d}t + \mathbb{E}\int_0^T\int_{\mathbb{R}} |\Delta_J G^*(t,x)|^2\nu(\mathrm{d}x)\mathrm{d}t < \infty. \tag{7.2.4}$$

Using Cauchy–Schwarz's inequality yields

$$\begin{aligned}
&\mathbb{E}\int_0^T \sigma^2 S_{t-}^2|\partial_y G^*(t,S_{t-})|\mathrm{d}t + \mathbb{E}\int_0^T\int_{\mathbb{R}} |\Delta_J G^*(t,x)S_{t-}(\mathrm{e}^x-1)|\nu(\mathrm{d}x)\mathrm{d}t \\
&\leqslant \sqrt{\mathbb{E}\int_0^T S_{t-}^2\mathrm{d}t}\sqrt{\mathbb{E}\int_0^T |\sigma^2 S_{t-}\partial_y G^*(t,S_{t-})|^2\mathrm{d}t} \\
&\qquad + \sqrt{\int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x)}\sqrt{\mathbb{E}\int_0^T S_{t-}^2\mathrm{d}t}\sqrt{\mathbb{E}\int_0^T\int_{\mathbb{R}}|\Delta_J G^*(t,x)|^2\nu(\mathrm{d}x)\mathrm{d}t} \\
&< \infty.
\end{aligned} \tag{7.2.5}$$

On the other hand, the FS decomposition of $H = g(S_T)$ is

$$g(S_T) = H_0 + \int_0^T \vartheta_t^{\mathrm{FS}}(H)\mathrm{d}S_t + L_T^{\mathrm{FS}} \tag{7.2.6}$$

where $H_0 \in \mathbb{R}$, $\vartheta^{\mathrm{FS}}(H) \in \mathcal{AS}(\mathbb{P})$ and $L^{\mathrm{FS}} \in \mathcal{M}_2^0(\mathbb{P})$ is $\mathbb{P}$-orthogonal to the martingale component S^{m} of S. According to [56, Eq. (3.10)], it holds that L^{FS} is a local $\mathbb{P}^*$-martingale. We remark that $\int_0^\cdot \vartheta_t^{\mathrm{FS}}(H)\mathrm{d}S_t$ is also a local $\mathbb{P}^*$-martingale. Using Cauchy–Schwarz's inequality and (7.1.3), we obtain

$$\mathbb{E}^*\sqrt{[L^{\mathrm{FS}},L^{\mathrm{FS}}]_T} \leqslant \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{\mathbb{E}[L^{\mathrm{FS}},L^{\mathrm{FS}}]_T} < \infty,$$
$$\mathbb{E}^*\sqrt{\int_0^T |\vartheta_t^{\mathrm{FS}}(H)|^2\mathrm{d}[S,S]_t} \leqslant \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{\mathbb{E}\int_0^T |\vartheta_t^{\mathrm{FS}}(H)|^2\mathrm{d}[S,S]_t} < \infty.$$

Hence, the Burkholder–Davis–Gundy inequality verifies that both L^{FS} and $\int_0^\cdot \vartheta_t^{\mathrm{FS}}(H)\mathrm{d}S_t$ are $\mathbb{P}^*$-martingales. Combining (7.2.3) with (7.2.6), we derive $H_0 = \mathbb{E}^* g(S_T)$ and

$$\int_0^\cdot \vartheta_t^{\mathrm{FS}}(H)\mathrm{d}S_t + L^{\mathrm{FS}} = \int_0^\cdot \sigma S_{t-}\partial_y G^*(t,S_{t-})\mathrm{d}W_t^* + \int_0^\cdot\int_{\mathbb{R}_0} \Delta_J G^*(t,x)\widetilde{N}^*(\mathrm{d}t,\mathrm{d}x). \tag{7.2.7}$$

Recall that the martingale part of S is $S^{\mathrm{m}} = \int_0^\cdot \sigma S_{t-}\mathrm{d}W_t + \int_0^\cdot\int_{\mathbb{R}_0} S_{t-}(\mathrm{e}^x-1)\widetilde{N}(\mathrm{d}t,\mathrm{d}x)$. Since $\langle L^{\mathrm{FS}}, S^{\mathrm{m}}\rangle^{\mathbb{P}} = 0$ by the definition of the FS decomposition, we take the predictable quadratic covariation on both sides of (7.2.7) with S^{m} under $\mathbb{P}$ and notice that the integrability condition (7.2.5) holds to obtain

$$\|(\sigma,\nu)\|\int_0^\cdot \vartheta_t^{\mathrm{FS}}(H)S_{t-}^2\mathrm{d}t = \int_0^\cdot \sigma^2 S_{t-}^2\partial_y G^*(t,S_{t-})\mathrm{d}t + \int_0^\cdot\int_{\mathbb{R}} \Delta_J G^*(t,x)S_{t-}(\mathrm{e}^x-1)\nu(\mathrm{d}x)\mathrm{d}t,$$

which yields (7.2.1) as desired.

(2) It follows from Cauchy–Schwarz's inequality and (7.2.4) that

$$\mathbb{E}^*\int_0^T |\sigma^2 S_{t-}\partial_y G^*(t,S_{t-})|\mathrm{d}t + \mathbb{E}^*\int_0^T\int_{\mathbb{R}} |\Delta_J G^*(t,x)(\mathrm{e}^x-1)|\nu(\mathrm{d}x)\mathrm{d}t$$

$$
\begin{aligned}
&\leqslant \sqrt{T}\|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{\mathbb{E}\int_0^T |\sigma^2 S_{t-}\partial_y G^*(t,S_{t-})|^2 \mathrm{d}t} \\
&\qquad + \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{T\int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x)}\sqrt{\mathbb{E}\int_0^T\int_{\mathbb{R}}|\Delta_J G^*(t,x)|^2\nu(\mathrm{d}x)\mathrm{d}t} \\
&< \infty. \qquad (7.2.8)
\end{aligned}
$$

By assumption, it is clear that $(G^*(t,\mathrm{e}^x S_t)-G^*(t,S_t))_{t\in[0,T]}$ is a $\mathbb{P}^*$-martingale for each $x\in\mathbb{R}$. In the case $\sigma>0$, due to $g(S_T)\in L_2(\mathbb{P}^*)$ and [61, Lemma 3.1], $(S_t\partial_y G^*(t,S_t))_{t\in[0,T)}$ is also a $\mathbb{P}^*$-martingale. Hence, the function

$$
[0,T)\ni t\mapsto \mathbb{E}^*|\sigma^2 S_t\partial_y G^*(t,S_t)|+\mathbb{E}^*\int_{\mathbb{R}}|G^*(t,\mathrm{e}^x S_t)-G^*(t,S_t)||\mathrm{e}^x-1|\nu(\mathrm{d}x)
$$

is non-decreasing by the martingale property. In addition, noticing that $S_{t-}=S_t$ a.s. for each $t\in[0,T]$, we infer from (7.2.8) and Fubini's theorem that

$$
\mathbb{E}^*|\sigma^2 S_t\partial_y G^*(t,S_t)|+\mathbb{E}^*\int_{\mathbb{R}}|G^*(t,\mathrm{e}^x S_t)-G^*(t,S_t)||\mathrm{e}^x-1|\nu(\mathrm{d}x)<\infty
$$

for all $t\in[0,T)$. Therefore,

$$
\left(\frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 S_t\partial_y G^*(t,S_t)+\int_{\mathbb{R}}(G^*(t,\mathrm{e}^x S_t)-G^*(t,S_t))(\mathrm{e}^x-1)\nu(\mathrm{d}x)\right)\right)_{t\in[0,T)}
$$

is a $\mathbb{P}^*$-martingale for which one can find a càdlàg modification, denoted by φ^g. Then the process $\tilde{\vartheta}^{\mathrm{FS}}(g)$ defined by

$$
\tilde{\vartheta}^{\mathrm{FS}}(g):=\frac{\varphi^g}{S} \qquad (7.2.9)
$$

satisfies the desired requirements. □

It turns out that any càdlàg version of the LRM strategy $\vartheta^{\mathrm{FS}}(H)$ of $H=g(S_T)$ can be determined as follows:

Remark 7.2.3. Let $\tilde{\vartheta}\in \mathrm{CL}([0,T))$ be such that $\tilde{\vartheta}=\tilde{\vartheta}^{\mathrm{FS}}(g)$ for $\mathbb{P}\otimes\lambda$-a.e. $(\omega,t)\in\Omega\times[0,T)$, where $\tilde{\vartheta}^{\mathrm{FS}}(g)$ given in (7.2.9). Then $\mathbb{P}(\tilde{\vartheta}_t=\tilde{\vartheta}^{\mathrm{FS}}_t(g),\ \forall t\in[0,T))=1$ due to the càdlàg property. Hence, $\tilde{\vartheta}_-$ is also a LRM strategy of $H=g(S_T)$, and it holds that, for any $t\in[0,T)$,

$$
\tilde{\vartheta}_t=\frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 S_t\partial_y G^*(t,S_t)+\int_{\mathbb{R}}(G^*(t,\mathrm{e}^x S_t)-G^*(t,S_t))(\mathrm{e}^x-1)\nu(\mathrm{d}x)\right) \quad \text{a.s.}
$$

7.3. Discretisation of LRM strategies

The results about approximation using the LRM strategy $\tilde{\vartheta}^{\mathrm{FS}}_-(g)$ are given in items (4)–(6) of Theorem 7.3.1 below. In fact, $\tilde{\vartheta}^{\mathrm{FS}}_-(g)$ is quite difficult to investigate directly under the original measure $\mathbb{P}$ but it fits well the main assumption Assumption 4.3.3 under the minimal martingale measure $\mathbb{P}^*$. Therefore, our idea is to switch between the original measure $\mathbb{P}$ and the minimal martingale measure $\mathbb{P}^*$ and use the fact that weighted BMO-norms allow a change of measure as given in Proposition 2.2.4(3). Here, we focus on the case $\gamma_S\neq 0$ (γ_S given in (7.1.1)) since the case $\gamma_S=0$, which corresponds to the martingale setting, is investigated in Chapter 6.

For $\eta\in[0,1]$, we recall the processes $\Theta(\eta)$, $\Phi(\eta)$ from (6.3.1), and define $\overline{\Phi}(\eta)$ as follows:

$$
\overline{\Phi}(\eta)_t:=\Phi(\eta)_t+\sup\nolimits_{u\in[0,t]}|\Delta\Phi(\eta)_u|, \quad t\in[0,T].
$$

We also recall $\mathring{W}^{1,q}(\mathbb{R}_+)$ from Definition 5.3.1, and $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ from Definition 5.3.2.

Theorem 7.3.1. *Assume Assumption 7.1.3, $\gamma_S \neq 0$ and $\int_{|x|>1} \mathrm{e}^{3x}\nu(\mathrm{d}x) < \infty$. Let $g \in \mathrm{Höl}_\eta(\mathbb{R}_+)$ with $\eta \in [0,1]$. Then the following assertions hold:*

(1) *Both $\Phi(\eta)$ and $\overline{\Phi}(\eta)$ belong to $\mathcal{SM}_3(\mathbb{P}) \cap \mathcal{SM}_2(\mathbb{P}^*)$.*

(2) *$\mathbb{P}^* \in \mathcal{RH}_3(\mathbb{P})$ and there is a constant $c > 0$ such that*
$$\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} \leqslant c\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})}.$$

(3) *Set $M := \tilde{\vartheta}^{\mathrm{FS}}(g)S$. Then Assumption 4.3.3 is satisfied under $\mathbb{P}^*$ for the selection*
$$\sigma(x) \equiv x, \quad \vartheta = \tilde{\vartheta}^{\mathrm{FS}}(g), \quad \Upsilon(\cdot,\mathrm{d}t) = \mathrm{d}\langle M,M\rangle_t^{\mathbb{P}^*} + M_t^2\mathrm{d}t, \quad \Theta = \Theta(\eta)$$
and for the parameter θ provided in Table 7.1.

(4) *With the adapted time-nets τ_n^θ given in (2.5.1), one has*
$$\sup_{n\geqslant 1} n^{\frac{1}{2r}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}^{\mathrm{FS}}(g), \tau_n^\theta \,\middle|\, n^{-\frac{1}{2r}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} < \infty, \tag{7.3.1}$$
where the parameters r and θ are provided in Table 7.1.

(5) *Let $s \in (1,\infty)$. Assume in addition when $\frac{\|(\sigma,\nu)\|}{\gamma_S} \in [-1,\infty)$ that $\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu(\mathrm{d}x) < \infty$. Then there is a constant $c \geqslant 1$ such that*
$$\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} \sim_c \|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})},$$
and hence
$$\sup_{n\geqslant 1} n^{\frac{1}{2r}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}^{\mathrm{FS}}(g), \tau_n^\theta \,\middle|\, n^{-\frac{1}{2r}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})} < \infty, \tag{7.3.2}$$
where the parameters r and θ are provided in Table 7.1. Moreover, (7.3.2) *holds true for the $L_3(\mathbb{P})$-norm in place of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm.*

(6) *If in addition $\int_{|x|>1} \mathrm{e}^{px}\nu(\mathrm{d}x) < \infty$ for some $p \in (3,\infty)$, then* (7.3.1) (*resp.* (7.3.2)) *is satisfied for the $L_{p-1}(\mathbb{P}^*)$-norm* (*resp. $L_p(\mathbb{P})$-norm*) *in place of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)$-norm* (*resp. $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm*).

Table 7.1: Values of parameters r and θ

	σ and η	*Small jump condition*	*Regularity of g*	*Conclusions for r and θ*
C1	$\sigma > 0$ $\eta \in (0,1]$	$\int_{|x|\leqslant 1} |x|^\alpha \nu(\mathrm{d}x) < \infty$ *for some* $\alpha \in [1,2]$	$g \in \mathrm{Höl}_\eta(\mathbb{R}_+)$	$\forall r \in [\alpha,2]$ $\forall \theta \in (0,\eta)$ *if* $\eta \in (0,1)$ $\theta = 1$ *if* $\eta = 1$
C2	$\sigma = 0$ $\eta \in [0,1]$	$\int_{|x|\leqslant 1} |x|^\alpha \nu(\mathrm{d}x) < \infty$ *for some* $\alpha \in [1,\eta+1]$	$g \in \mathrm{Höl}_\eta(\mathbb{R}_+)$	$\forall r \in [\alpha,2]$ $\theta = 1$
C3	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_1(\alpha)$ *for some* $\alpha \in [1+\eta,2)$	$g \in \mathrm{Höl}_\eta(\mathbb{R}_+)$	$\forall r \in (\alpha,2]$ $\forall \theta \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$

C4	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_2(\alpha)$ *for some* $\alpha \in [1+\eta, 2)$	$g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$	$\forall r \in (\alpha, 2]$ $\forall \theta \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$

PROOF. See the proof of [61, Theorem 5.12]. $\square$

Remark 7.3.2. (1) Let us comment on the parameters r and θ in Table 7.1. First, since we use the adapted time-net τ_n^θ which leads to better estimates, it implies that the parameter r only depends on the behavior of ν around zero. Moreover, the smaller r is, the better approximation accuracy one achieves. The parameter θ is the outcome of the interplay between the behavior of ν around zero and the Hölder regularity of the payoff function.

(2) Since X is a Lévy process under both measures $\mathbb{P}$ and $\mathbb{P}^*$, we apply [60, Proposition 5.3] (with $\alpha = 2$ and $\kappa = \frac{1-\theta}{2}$, $\varepsilon = n^{-\frac{1}{2r}}$) to conclude that the parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)$-norm in (7.3.1) can be regarded as the $L_2(\mathbb{P})$-norm and also the $L_2(\mathbb{P}^*)$-norm of the cardinality of the combined time-net $\tau_n^\theta \sqcup \rho(n^{-\frac{1}{2r}}, \frac{1-\theta}{2})$. The parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm in (7.3.2) can be interpreted in a similar manner.

Bibliography

1. D. Applebaum. *Lévy processes and stochastic calculus (2nd ed.)*. Cambridge University Press, Cambridge, 2009.
2. D. Applebaum, R. Bañuelos. Probabilistic approach to fractional integrals and the Hardy-Littlewood-Sobolev inequality. *Analytic methods in interdisciplinary applications*, Springer Proc. Math. Stat., 116, Cham, 2015, 17–40.
3. R. Arai, E. Nakai, G. Sadasue. Fractional integrals and their commutators on martingale Orlicz spaces. *J. Math. Anal. Appl.* **487** (2020), no. 2, 123991.
4. T. Arai, R. Suzuki. Local risk-minimization for Lévy markets. *Int. J. Financ. Eng.* **2** (2015), no. 2, 1550015.
5. F. Baumgartner, S. Geiss. Permutation invariant functionals of Lévy processes. *Trans. Amer. Math. Soc.* **369** (2017) 8607–8641.
6. F. Benth, G. Di Nunno, A. Løkka, B. Øksendal, F. Proske. Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. *Math. Finance* **13** (2003), no. 1, 55–72.
7. J. Bergh, J. Löfström, *Interpolation spaces*, Springer, 1976.
8. R. Blumenthal, R. Getoor. Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10** (1961) 493–516.
9. M. Brodén, P. Tankov. Tracking errors from discrete hedging in exponential Lévy model. *Int. J. Theor. Appl. Finance* **14** (2011), no. 6, 803–837.
10. Z.-Q. Chen, Z. Hao, X. Zhang. Hölder regularity and gradient estimates Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes. (2020) arXiv:2001.03873.
11. T. Choulli, L. Krawczyk, C. Stricker. $\mathcal{E}$-martingales and their applications in mathematical finance. *Ann. Probab.* **26** (1998), no. 2, 853–876.
12. T. Choulli, N. Vandaele, M. Vanmaele. The Föllmer–Schweizer decomposition. Comparison and description. *Stochastic Process. Appl.* **120** (2010) 853–872.
13. R. Cont, P. Tankov. *Financial modeling with jump processes*. Chapman & Hall/CRC Press, 2003.
14. R. Cont, P. Tankov, E. Voltchkova, *Hedging with options in models with jumps*, In: Stochastic Analysis and Applications – The Abel Symposium 2005, Springer, Berlin, 2007.
15. F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, C. Stricker. Weighted norm inequalities and hedging in incomplete markets. *Finance Stoch.* **1** (1997) 181–227.
16. C. Dellacherie, P.-A. Meyer. *Probabilities and potential B: Theory of martingales*. North-Holland Publishing Co., 1982.
17. F. Esche, M. Schweizer. Minimal entropy preserves the Lévy property: how and why. *Stochastic Process. Appl.* **115** (2005) 299–327.
18. S. Dereich, F. Heidenreich. A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. *Stochastic Process. Appl.* **121** (2011) 1565–1587.
19. M. Fukasawa, Asymptotically efficient discrete hedging. In: *Stochastic Analysis with Financial Applications. Progress in Probability*, **65**, 331–346, Springer, Basel, 2011.
20. M. Fukasawa. Efficient discretization of stochastic integrals. *Finance Stoch.* **18** (2014) 175–208.
21. C. Geiss, S. Geiss. On approximation of a class of stochastic integrals and interpolation. *Stochastics Stochastics Rep*. **76** (2004), no. 4, 339–362.
22. C. Geiss, S. Geiss, E. Laukkarinen. A note on Malliavin fractional smoothness for Lévy processes and approximation. *Potential Anal.* **39** (2013) 203–230.
23. C. Geiss, A. Steinicke. Malliavin derivative of random functions and applications to Lévy driven BSDEs. *Electron. J. Probab.* **21** (2016), no. 10, 28 pp.
24. S. Geiss. Quantitative approximation of certain stochastic integrals. *Stochastics Stochastics Rep*. **73** (2002) 241–270.
25. S. Geiss. Weighted BMO and discrete time hedging within the Black-Scholes model. *Probab. Theory Related Fields* **132** (2005) 13–38.
26. S. Geiss, M. Hujo. Interpolation and approximation in $L_2(\gamma)$. *J. Approx. Theory* **144** (2007) 213–232.
27. S. Geiss, A. Toivola. On fractional smoothness and L_p-approximation on the Gaussian space. *Ann. Probab.* **43** (2015), no. 2, 605–638.
28. S. Geiss, A. Toivola. Weak convergence of error processes in discretization of stochastic integrals and Besov spaces. *Bernoulli* **15** (2009), no. 4, 925–954.

29. S. Geiss, N. T. Thuan. On Riemann–Liouville operators, BMO, gradient estimates in the Lévy–Itô space, and approximation. (2020) arXiv:2009.00899v1.
30. E. Gobet, N. Landon. Almost sure optimal hedging strategy. *Ann. Appl. Probab.* **24** (2014), no. 4, 1652–1690.
31. E. Gobet, E. Temam. Discrete time hedging errors for options with irregular pay-offs. *Finance Stoch.* **5** (2001), no. 3, 357–367.
32. S. Goutte, N. Oudjane, F. Russo. Variance optimal hedging for continuous time additive processes and applications. *Stochastics* **86** (2014), no. 1, 147–185.
33. Y. Hu, D. Nualart, J. Song. Fractional martingales and characterization of the fractional Brownian motion. *Ann. Probab.* **37** (2009) 2404–2430.
34. F. Hubalek, J. Kallsen, L. Krawczyk. Variance-optimal hedging for processes with stationary independent increments. *Ann. Appl. Probab.* **16** (2006), no. 2, 853–885.
35. K. Itô. Spectral type of the shift transformation of differential processes with stationary increments. *Trans. Amer. Math. Soc.* **81** (1956) 253–263.
36. J. Jacob, S. Méléard, P. Protter. Explicit form and robustness of martingale representations. *Ann. Probab.* **28** (2000), no. 4, 1747–1780.
37. J. Jacod, A. Shiryaev. *Limit theorems for stochastic processes (2nd ed.)*, Springer, Berlin, 2003.
38. J. Kallsen, A. Pauwels. Variance-optimal hedging in general affine stochastic volatility models. *Adv. Appl. Prob.* **42** (2010) 83–105.
39. E. Laukkarinen. *On Malliavin calculus and approximation of stochastic integrals for Lévy process*. PhD thesis, University of Jyväskylä, 2013.
40. E. Laukkarinen. Malliavin smoothness on the Lévy space with Hölder continuous or BV functionals. *Stochastic Process. Appl.* **130** (2020), no. 8, 4766–4792.
41. A. Løkka. Martingale representation of functionals of Lévy processes. *Stoch. Anal. Appl.* **22** (2004), no. 4, 867–892.
42. J. Ma, P. Protter, J. Zhang. Explicit form and path regularity of martingale representations, In: *Lévy processes: Theory and applications*, Springer, New York, 2001.
43. P. Malliavin, H. Airault, L. Kay, G. Letac. *Integration and probability*, Springer, New York, 1995.
44. P. Monat, C. Stricker. Föllmer–Schweizer decomposition and mean-variance hedging for general claims, *Ann. Probab.* **23** (1995), no. 2, 605–628.
45. D. Nualart, E. Nualart. *Introduction to Malliavin calculus*, Cambridge University Press, Cambridge, 2018.
46. G. Di Nunno, B. Øksendal, F. Proske. *Malliavin calculus for Lévy processes with applications to finance*, Springer-Verlag Berlin Heidelberg, 2009.
47. P. Protter. *Stochastic integration and differential equations (2nd ed., ver. 2.1)*, Springer-Verlag, Berlin, 2005.
48. M. Rosenbaum, P. Tankov. Asymptotically optimal discretization of hedging strategies with jumps. *Ann. Appl. Probab.* **24** (2014), no. 3, 1002–1048.
49. W. Rudin. *Functional Analysis*, McGraw-Hill, 1973.
50. W. Rudin. *Real and complex analysis (3rd. ed.)*, McGraw-Hill, 1987.
51. K. Sato. *Lévy processes and infinitely divisible distributions (2nd ed.)*, Cambridge University Press, Cambridge, 2013.
52. R.L. Schilling, P. Sztonyk, J. Wang. Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18** (2012) 1128–1149.
53. W. Schoutens. *Lévy processes in finance: Pricing financial derivatives*, Wiley, 2003.
54. M. Schweizer. Approximating random variables by stochastic integrals. *Ann. Probab.* **22** (1994), no. 3, 1536–1575.
55. M. Schweizer. On the minimal martingale measure and the Föllmer–Schweizer decomposition. *Stoch. Anal. Appl.*, **13** (1995), no. 5, 573–599.
56. M. Schweizer, A guided tour through quadratic hedging approaches, In: *Option pricing, interest rates and risk management*, Cambridge Univ. Press, Cambridge, 2001.
57. J. Solé, F. Utzet, J. Vives. *Chaos expansion and Malliavin calculus for Lévy processes*, In: Stochastic Analysis and Applications – The Abel Symposium 2005, Springer, Berlin, 2007.
58. P. Tankov. *Pricing and hedging in exponential Lévy models: review of recent results*, Paris-Princeton Lecture Notes in Mathematical Finance, Springer, 2010.
59. P. Tankov, E. Voltchkova. Asymptotic analysis of hedging errors in models with jumps. *Stochastic Process. Appl.* **119** (2009) 2004–2027.

60. N. T. Thuan. Approximation of stochastic integrals with jumps in weighted bounded mean oscillation spaces. (2020) arXiv:2009.02116v2.
61. N. T. Thuan. Explicit Föllmer–Schweizer decomposition and discrete-time hedging in exponential Lévy models. (2020) arXiv:2009.04328v1.
62. R. Zhang. *Couverture approchée des options Européennes*, Ph.D. thesis, École Nationale des Ponts et Chaussées, Paris, 1998.

Appendices

This extra part includes three manuscripts which correspond to the references [29, 60, 61], and they contain detailed proofs of results formulated in the thesis:

[29] S. Geiss, N. T. Thuan. On Riemann–Liouville operators, BMO, gradient estimates in the Lévy–Itô space, and approximation. (2020) arXiv:2009.00899v1.

[60] N. T. Thuan. Approximation of stochastic integrals with jumps in weighted bounded mean oscillation spaces. (2020) arXiv:2009.02116v2.

[61] N. T. Thuan. Explicit Föllmer–Schweizer decomposition and discrete-time hedging in exponential Lévy models. (2020) arXiv:2009.04328v1.

ON RIEMANN-LIOUVILLE OPERATORS, BMO, GRADIENT ESTIMATES IN THE LÉVY-ITÔ SPACE, AND APPROXIMATION

STEFAN GEISS AND NGUYEN TRAN THUAN

Abstract. In this article we discuss in a stochastic framework the interplay between Riemann-Liouville operators applied to càdlàg processes, real interpolation, weighted bounded mean oscillation, estimates for gradient processes on the Lévy-Itô space, and the connection to an approximation problem for stochastic integrals. We prove upper and lower bounds for gradient processes appearing in a Brownian setting within the Feynman-Kac theory for parabolic PDEs and in the setting of Lévy processes. The upper bounds are formulated by BMO-conditions on the fractional integrated gradient, the lower bounds are formulated in terms of oscillatory quantities. In the case of Lévy processes we are concerned with a gradient process with values in a Hilbert space where the regularity of this process depends on the direction within this Hilbert space. Moreover, it turns out that certain Hölder properties of terminal functions transfer into a singularity in time that can be compensated by Riemann-Liouville operators.

Contents

1. Introduction

This article investigates the interplay between Riemann-Liouville operators applied to càdlàg processes, gradient estimates for functionals on the Lévy-Itô space, bounded mean oscillation (BMO), approximation theory, and the real interpolation method from Banach space theory.

To explain this, let us assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ with finite time-horizon $T > 0$. There are various applications in which stochastic processes $\varphi = (\varphi_t)_{t\in[0,T)}$ appear that have a singularity when $t \uparrow T$, for example in L_p for some $p \in [1, \infty]$. Examples are gradient processes obtained from (semi-linear) parabolic backward PDEs within the Feynman-Kac theory, where these processes appear as integrands in stochastic integral representations (see Section 6) or in backward stochastic differential equations as gradient processes. The same type of processes appear also as gradient processes originating from convolution semi-groups based on Lévy processes and that are used, for example, in Galtchouk-Kunita-Watanabe projections (see Section 8).

If one analyzes these examples, then one realizes the following:

Date: Tuesday 17th November, 2020.

2010 *Mathematics Subject Classification.* Primary 26A33, 46B70, 60G44, 60H10, 60G51, Secondary 60G07, 60Hxx.

Key words and phrases. Riemann-Liouville operator, real interpolation, bounded mean oscillation, diffusion process, Lévy process, gradient estimate, Hölder space.

The authors were supported by the Project 298641 *'Stochastic Analysis and Nonlinear Partial Differential Equations, Interactions and Applications'* of the Academy of Finland.

– Self-similarity: There is a Markovian structure behind that generates a self-similarity in the sense that, given $a \in (0,T)$ and $B \in \mathcal{F}_a$ of positive measure, then $(\varphi_t)_{t\in[a,T)}$ restricted to B has similar properties as $(\varphi_t)_{t\in[0,T)}$. If one is interested in good distributional estimates of $(\varphi_t)_{t\in[0,T)}$ or functionals of it, then it is useful to consider the BMO-setting: the particular feature of BMO-estimates is that one uses conditional L_2-estimates, where one might exploit conditional orthogonality, in order to deduce L_p-estimates for $p > 2$ or exponential estimates by John-Nirenberg type theorems.
– Polynomial blow-up: In the problems mentioned above the size of the singularity of φ (or, again, a functional of it) increases polynomially in time with a rate $(T-t)^{-\alpha}$ for some $\alpha > 0$. In particular, this often occurs in the presence of Hölder functionals as terminal conditions in backward problems.

The above observations lead to an interplay between Riemann-Liouville operators, BMO, and the real interpolation method. These components interact as follows: We realized that the Riemann-Liouville operators allow for a transformation of a stochastic process with a certain singularity when $t \uparrow T$ into a stochastic process without this singularity (but without loosing any information about the process one is starting from). In particular, this is of interest for martingales. By the obtained formulas this opens a link to real interpolation theory, which has a natural explanation as we interpolate with a two-parametric scale between, for example, martingales without singularity and martingales with a singularity. As a consequence of the self-similarity of the singular process one is starting from, it is natural to think that the Riemann-Liouville operator turns this process into a BMO-process by removing the singularity but keeping the self-similarity. Therefore we consider the stochastic processes transformed by the Riemann-Liouville operator in the BMO-setting. One starting point to investigate the connections between Riemann-Liouville operators, BMO, and real interpolation is an approximation problem for stochastic integrals, so that we will deal with four objects that interact with each other.

In the second part of this article we give two applications of the above methodology in Section 6 and Section 8. To explain this, let $C_b^0(\mathbb{R})$ be the bounded continuous functions and $\mathrm{H\ddot{o}l}_1^0(\mathbb{R})$ be the Lipschitz functions, both defined on $\mathbb{R}$ and vanishing at zero. We define the two-parametric scale of Hölder functions by the real interpolation method as

$$\mathrm{H\ddot{o}l}_{\eta,q}^0(\mathbb{R}) := (C_b^0(\mathbb{R}), \mathrm{H\ddot{o}l}_1^0(\mathbb{R}))_{\eta,q} \quad \text{for} \quad (\eta,q) \in (0,1)\times[1,\infty].$$

Section 6: Let $W = (W_t)_{t\in[0,T]}$ be a standard Brownian motion and $Y = (Y_t)_{t\in[0,T]}$ be the geometric Brownian motion

$$Y_t = \mathrm{e}^{W_t - \frac{t}{2}},$$

and consider a Borel function $g : (0,\infty) \to \mathbb{R}$ with $g(Y_T) \in L_2$ and

$$g(Y_T) = \mathbb{E}g(Y_T) + \int_{(0,T)} \varphi_t \mathrm{d}Y_t \quad \text{a.s.} \tag{1.1}$$

Here, for $t \in [0,T)$ we use

$$G(t,y) := \mathbb{E}g(yY_{T-t}), \quad \varphi_t := (\partial G/\partial y)(t,Y_t), \quad \text{and} \quad Z_t := \varphi_t Y_t,$$

so that $g(Y_T) = \mathbb{E}g(Y_T) + \int_{(0,T)} Z_t \mathrm{d}W_t$. For a deterministic time-net τ, $0 = t_0 < t_1 < \cdots < t_n = T$, we define the approximation error for the Riemann approximation of the stochastic integral as

$$E_t(g;\tau) := \int_0^t \varphi_s \mathrm{d}Y_s - \sum_{i=1}^n \varphi_{t_{i-1}}(Y_{t_i\wedge t} - Y_{t_{i-1}\wedge t}).$$

One has $\|E_T(g;\tau)\|_{L_2} \geqslant \frac{c}{\sqrt{n}}$ for some $c > 0$ for all time-nets τ, $0 = t_0 < \cdots < t_n = T$, provided that there are no $a, b \in \mathbb{R}$ such that $g(Y_T) = a + bY_T$ a.s. (see [16, Theorem 2.5]). To estimate $E_T(g;\tau)$ from above usually the L_2-setting is used to exploit orthogonality (see, for example, [19, 16, 22] for the Wiener space and [17] for the corresponding problem on the Lévy-Itô space). The approximation in L_p for $p \in [2,\infty)$ is considered in [24] on the Wiener space. A different route is taken in [20] where it is shown by [20, Theorems 7 and 8] that

$$\|(E_t(g;\tau))_{t\in[0,T]}\|_{\mathrm{BMO}_2^Y([0,T])} \leqslant c\sqrt{\sup_{i=1,\dots,n}|t_i-t_{i-1}|} \text{ for all } \tau=\{t_i\}_{i=0}^n\in\mathcal{T}$$
$$\Longleftrightarrow g \text{ is (equivalent to) a Lipschitz function,} \qquad (1.2)$$

where $\mathcal{T}$ stands for the set of all deterministic time-nets $\tau=\{t_i\}_{i=0}^n$, $0=t_0<\cdots<t_n=T$, and the weighted BMO-spaces $\mathrm{BMO}_2^Y([0,T])$ are introduced in Definition 2.1. Note that $\|(E_t(g;\tau))_{t\in[0,T]}\|_{\mathrm{BMO}_2^Y([0,T])}$ only requires *local* L_2*-estimates* that are more feasible than L_p-estimates for $p>2$. The importance of the BMO_q^Φ-spaces, $q\in(0,\infty)$, comes from the fact that, for example,

$$\mathbb{P}\left(\sup_{t\in[0,T]}|A_t|>a\mu\nu\|A\|_{\mathrm{BMO}_q^\Phi([0,T])}\right)\le \mathrm{e}^{1-\mu}+\alpha\mathbb{P}\left(\sup_{t\in[0,T]}\Phi_t>\nu\right)$$

for $\mu,\nu>0$, where $a,\alpha>0$ are constants depending at most on q (this follows from [20, equation (5), part (a) of the proof of Corollary 1]). Therefore, the moments of the weight Φ determine the moments of A. This BMO-approach is also used in the context of BSDEs in [25].

Our **first main result** is the extension of the equivalence (1.2). Firstly we show in Theorem 6.4 that the geometric Brownian motion in (1.2) can be replaced by a more general diffusion while keeping the equivalence. However, this is still in the Lipschitz framework and gives the impression that this approach is tight to Lipschitz functionals $g(Y_T)$. But our second contribution is to move away from the Lipschitz framework, which is done in Theorem 6.5, where we prove for $\theta\in(0,1)$ that

$$\forall\tau\in\mathcal{T}\quad \|(E_t(g;\tau))_{t\in[0,T)}\|_{\mathrm{BMO}_2^\Phi([0,T))}\leqslant c\sqrt{\|\tau\|_\theta}\iff \begin{cases}\mathcal{I}^{\frac{1-\theta}{2}}Z-Z_0\in\mathrm{BMO}_2^\Phi([0,T))\\ (T-t)^{\frac{1-\theta}{2}}|Z_t|\leqslant c\Phi_t \quad \text{a.s.}\end{cases} \qquad (1.3)$$

under mild conditions on the weight-process Φ and an a-priori condition on φ, where the *Riemann-Liouville operator* $\mathcal{I}^\alpha$ is defined in (1.10) and

$$\|\tau\|_\theta := \sup_{i=1,\dots,n}\frac{t_i-t_{i-1}}{(T-t_{i-1})^{1-\theta}} \quad \text{for} \quad \tau=\{t_i\}_{i=0}^n\in\mathcal{T}. \qquad (1.4)$$

In particular, we consider the time-nets $\tau_n^\theta := (T-T(1-(i/n))^{\frac{1}{\theta}})_{i=0}^n$ which concentrate more around $t=T$ the smaller θ is and which are therefore suitable to handle singularities at $t\uparrow T$. Since we have

$$\|\tau_n^\theta\|_\theta\leqslant\frac{T^\theta}{\theta n},$$

one obtains the optimal rate $1/\sqrt{n}$ on the left-hand side in (1.3). The right-hand side in (1.3) is a statement about fractional smoothness in the following sense: After *removing a singularity of order* $\frac{1-\theta}{2}$ from the process Z by applying the Riemann-Liouville operator of order $\frac{1-\theta}{2}$ we obtain an object in $\mathrm{BMO}_2^\Phi([0,T))$. So one might think about a fractional smoothness of order $1-\frac{1-\theta}{2}=\frac{1+\theta}{2}$ in $\mathrm{BMO}_2^\Phi([0,T))$. The next step is to investigate the right-hand side of (1.3) which is of independent interest. For $g\in\mathrm{Höl}_{\theta,2}^0(\mathbb{R})$ and the weight process $\Phi_t:=Y_t\sup_{s\in[0,t]}(Y_s^{\theta-1})$ we show in Theorem 6.5 and Theorem 6.6 for all $q\in(0,\infty)$ that

$$\mathcal{I}^{\frac{1-\theta}{2}}Z-Z_0\in\mathrm{BMO}_q^\Phi([0,T)) \quad\text{and}\quad \mathcal{I}_T^{\frac{1-\theta}{2}}Z:=\lim_{t\uparrow T}\mathcal{I}_t^{\frac{1-\theta}{2}}Z \text{ in } L_q \text{ and } a.s.,$$

and that Φ satisfies a generalized reverse Hölder inequality (denoted by $\Phi\in\mathcal{SM}_q([0,T))$ in Definition 2.2).

Section 8: The second application concerns gradient estimates for functionals of Lévy processes. Let us assume a pure jump Lévy process $X=(X_t)_{t\in[0,T]}$ which is an L_2-martingale with a non-zero *Lévy measure* ν. Given a functional $f(X_T)\in L_2$, where $f:\mathbb{R}\to\mathbb{R}$ is a Borel function, we consider a *gradient process*

$$M_t:\Omega\to H\cong L_2(\mathbb{R},z^2\nu(\mathrm{d}z)),\quad t\in[0,T),$$

associated to $f(X_T)$, which naturally replaces the process $Z=(Z_t)_{t\in[0,T)}$, obtained from (1.1) by $Z=\varphi Y$, and satisfies (for the precise interpretation see Appendix D.2 and Appendix D.3)

$$M_t=\frac{1}{t}\int_0^t D_{s,\cdot}\mathbb{E}^{\mathcal{F}_t}[f(X_T)]\,\mathrm{d}s \quad \text{for} \quad t\in(0,T), \qquad (1.5)$$

where $D_{s,z}$ is the Malliavin derivative. Assume that the Lévy process satisfies

$$\sup_{s\in(0,T]} \sup_{z\in\operatorname{supp}(\nu)\setminus\{0\}} s^{\frac{1}{\beta}} \frac{\|\mathbb{P}_{z+X_s}-\mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} < \infty \tag{1.6}$$

for some $\beta \in (1,2)$, where $\mathbb{P}_{z+X_s}$ and $\mathbb{P}_{X_s}$ are the laws of $z+X_s$ and X_s, respectively, and $\|\cdot\|_{\mathrm{TV}}$ stands for the total variation. Upper bounds for $\|\mathbb{P}_{z+X_s}-\mathbb{P}_{X_s}\|_{\mathrm{TV}}$ are already investigated in the literature (see [39, Theorem 3.1]). We measure the fractional smoothness of $f(X_T)$ in the direction $D \in L_2(\mathbb{R}, z^2\nu(\mathrm{d}z))$, $D \geqslant 0$, by determining the regularity of the "directional" martingales $(\langle M, D\rangle_H(t))_{t\in[0,T)}$ in dependence on D. It turns out that, for $\varepsilon \in (0,1)$, the ε-*small ball condition*,

$$\sup_{n\geqslant 1} 2^{\varepsilon n} \int_{\{2^{-n}\leqslant |z|<2^{-n+1}\}} D(z) z^2 \nu(\mathrm{d}z) < \infty, \tag{1.7}$$

plays a central role. A **second main result** of the article is

$$\mathcal{I}^\alpha\left(\Big(\langle M,D\rangle_H(t)-\langle M,D\rangle_H(0)\Big)_{t\in[0,T)}\right) \in \mathrm{BMO}_2([0,T)) \quad \text{with} \quad \alpha := \frac{1-(\varepsilon+\eta)}{\beta} \tag{1.8}$$

for $(\varepsilon,\beta)\in(0,1)\times(1,2)$, $\eta\in(0,1-\varepsilon)$, and $f\in \mathrm{Höl}^0_{\eta,2}(\mathbb{R})$. To check this, we define the measure

$$\rho(\mathrm{d}z) := D(z)z^2\nu(\mathrm{d}z)/\int_{\mathbb{R}} D(z)z^2\nu(\mathrm{d}z) \quad \text{on} \quad \mathcal{B}(\mathbb{R})$$

so that (1.7) turns into $\sup_{n\geqslant 1} 2^{\varepsilon n}\rho(\{2^{-n}\leqslant |z| < 2^{-n+1}\}) < \infty$, use equation (D.4), and apply to $\overline{D}_\rho F(t, X_t)$ from (D.4) the statements Theorem 8.11 and Corollary 8.13. The relation (1.8) is the counterpart to (1.3), however α depends on the direction D via the small ball condition (1.7). An application, we discuss, is the approximation of the stochastic integral appearing in the Galtchouk-Kunita-Watanabe projection of $f(X_T)$ if one projects on the space of stochastic integrals driven by

$$X_t^D := \int_{(0,t]\times\mathbb{R}} D(x)x\widetilde{N}(\mathrm{d}s,\mathrm{d}x),$$

where $\widetilde{N}$ is the compensated Poisson random measure of X. By Proposition 8.2 and equation (D.4) we have for $f\in\mathcal{D}_X\cap L_2(\mathbb{R},\mathbb{P}_{X_T})$ ($\mathcal{D}_X$ is given in Definition 8.1) the explicit representation of the Galtchouk-Kunita-Watanabe projection

$$\frac{1}{\langle D,D\rangle_H}\int_{(0,T)} \langle M,D\rangle_H(t-)\mathrm{d}X_t^D.$$

In our later notation we will have $\varphi_t(f,\rho) = \langle M,D\rangle_H(t)/\int_{\mathbb{R}} D(z)z^2\nu(\mathrm{d}z)$ and define the corresponding error process of the Riemann approximation of the stochastic integral with respect to the time-net $\tau=\{t_i\}_{i=0}^n \in\mathcal{T}$ as

$$E_t(f;\tau,D) := \int_0^t \varphi_{s-}(f,\rho)\mathrm{d}X_s^D - \sum_{i=1}^n \varphi_{t_{i-1}-}(f,\rho)(X^D_{t_i\wedge t}-X^D_{t_{i-1}\wedge t}), \quad t\in[0,T).$$

Let us additionally assume that the Lévy measure satisfies $\nu(\mathrm{d}z) = p_\nu(z)\mathrm{d}z$, where p_ν is symmetric and

$$0 < \liminf_{|z|\to 0} |z|^{1+\beta}p_\nu(z) \leqslant \limsup_{|z|\to 0}|z|^{1+\beta}p_\nu(z) < \infty$$

which ensures that (1.6) is satisfied. Assume also that the functional D satisfies the ε-small ball condition (1.7). Then, in Theorem 8.21 we prove that for $p\in[2,\infty)$, $\theta := 1-2\alpha$ (α is given by (1.8)), and $f\in\mathrm{Höl}_{\eta,2}(\mathbb{R})$ one has

$$\mathbb{P}_{\mathcal{F}_a}\left(|E_t(f;\tau_n^\theta,D)-E_a(f;\tau_n^\theta,D)|>\lambda\right) \leqslant c\min\left\{\frac{1}{n\lambda^2}, \frac{\mathbb{E}^{\mathcal{F}_a}\Big[\sup_{u\in[a,t]}\Phi_u^p\Big]}{\lambda^p(T-t)^{p\alpha}}\right\} \text{ a.s.} \tag{1.9}$$

for $0\leqslant a<t<T$ and $\lambda>0$ and any non-negative adapted càdlàg process $(\Phi_u)_{u\in[0,t]}$ with $1\vee|\Delta X_s|\leqslant\Phi_s$ for all $s\in[0,t]$ and $\sup_{u\in[0,t]}\Phi_u\in L_p$. Inequality (1.9) corresponds to the left-hand side of (1.3). Here $1/(n\lambda^2)$ is achieved by using the adapted time-nets τ_n^θ. If $p>2$, then we have a higher integrability by the term $1/\lambda^p$ that goes back to the self-improving properties of the BMO-spaces. For example, this term leads to the large deviation inequality

$$\mathbb{P}\left(|E_t(f;\tau_n^\theta,D)|>\lambda\right)\leqslant c\frac{1}{\lambda^p}\frac{\mathbb{E}\sup_{u\in[a,t]}\Phi_u^p}{(T-t)^{p\alpha}}$$

that gives a better upper bound than $\frac{1}{n\lambda^2}$ for large λ.

In order to treat the applications described so far we deal with some general results about the interaction of Riemann-Liouville operators, interpolation, BMO, and approximation theory:

Section 3: We study general properties of Riemann-Liouville operators applied to martingales and the relation to real interpolation and an integrated square function. In Definition 3.1 for $\alpha>0$ and a càdlàg function $K:[0,T)\to\mathbb{R}$ we define $\mathcal{I}^\alpha K:=(\mathcal{I}_t^\alpha K)_{t\in[0,T)}$ by

$$\mathcal{I}_t^\alpha K:=\frac{\alpha}{T^\alpha}\int_0^T(T-u)^{\alpha-1}K_{u\wedge t}\mathrm{d}u\quad\text{and}\quad\mathcal{I}_t^0K:=K_t. \tag{1.10}$$

Furthermore, in Definition 3.5 we define for a càdlàg process $\varphi=(\varphi_t)_{t\in[0,T)}$, $a\in[0,T]$, and a deterministic time-net $\tau=\{t_i\}_{i=0}^n$, $0=t_0<\cdots<t_n=T$, the following integrated square-function

$$[\varphi;\tau]_a:=\int_0^a\left|\varphi_u-\sum_{i=1}^n\varphi_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2\mathrm{d}u.$$

In Theorem 3.6 we prove for $\theta\in(0,1)$ and a càdlàg martingale $\varphi=(\varphi_t)_{t\in[0,T)}\subseteq L_2$ the equivalence

$$(\mathcal{I}_t^{\frac{1-\theta}{2}}\varphi)_{t\in[0,T)}\text{ is a martingale closable in }L_2\iff\exists c>0\;\forall\tau\in\mathcal{T}\;\mathbb{E}[\varphi;\tau]_T\leqslant c\|\tau\|_\theta \tag{1.11}$$

where $\|\cdot\|_\theta$ is defined in (1.4). Theorem 3.6 also includes an *equivalence* to interpolation spaces of type $(E_0,E_1)_{\theta,2}$. Theorem 3.6 enables us to connect the Riemann-Liouville operators and the Hölder spaces $\mathrm{Höl}_{\eta,2}^0(\mathbb{R})$ to our approximation problem. Independently from the above connections, the functional $[\varphi;\tau]$ can be interpreted as a square-function adapted to non-closable martingales.

Section 4 transfers (1.11) to the setting of weighted BMO, the setting we exploit for our estimates later. A special case of Theorem 4.8 is the following:

Theorem 1.1. *Assume a càdlàg martingale $\varphi=(\varphi_t)_{t\in[0,T)}\subseteq L_2$. Then for $\theta\in(0,1]$ and $\alpha:=\frac{1-\theta}{2}$ the following assertions are equivalent:*

(1) *One has $\mathcal{I}^\alpha\varphi-\varphi_0\in\mathrm{bmo}_2([0,T))$ and there is a $c_{(1.12)}>0$ such that one has*

$$|\varphi_a-\varphi_s|\leqslant c_{(1.12)}\frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac12}}\quad\textit{for}\quad 0\leqslant s<a<T\textit{ a.s.} \tag{1.12}$$

(2) *There is a constant $c_{(1.13)}>0$ such that, for all time-nets $\tau\in\mathcal{T}$,*

$$\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))}\leqslant c_{(1.13)}\|\tau\|_\theta. \tag{1.13}$$

In (1) the $\mathrm{bmo}_2([0,T))$-spaces are defined in Definition 2.1. Moreover, in no direction the conditions $\mathcal{I}^\alpha\varphi-\varphi_0\in\mathrm{bmo}_2([0,T))$ and (1.12) imply each other in general (see [21]).

Section 5: We find lower bounds for (1.13) by using a concept of lower *oscillation* of stochastic processes. In Definition 5.1 we define for a stochastic process $\varphi=(\varphi_t)_{t\in[0,T)}$ and $t\in(0,T)$ the oscillatory quantity

$$\underline{\mathrm{Osc}}_t(\varphi):=\inf_{s\in[0,t)}\|\varphi_t-\varphi_s\|_{L_\infty}$$

and call φ of *maximal oscillation* with constant $c\geqslant1$ if for all $t\in(0,T)$ one has

$$\underline{\mathrm{Osc}}_t(\varphi)\geqslant\frac1c\|\varphi_t-\varphi_0\|_{L_\infty}.$$

For us the maximal oscillation is of interest for martingales as it says that for all $0<s<t$ one has $\|\varphi_t-\varphi_s\|_{L_\infty}\sim\|\varphi_t-\varphi_0\|_{L_\infty}$ up to some factor. The corresponding lower bounds for (1.13) are summarized in the following statement (see Theorem 5.7):

Theorem 1.2. *Assume $\theta\in(0,1]$ and a martingale $(\varphi_t)_{t\in[0,T)}\subseteq L_2$. Assume that one has $\infty>\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))}\to0$ whenever $\|\tau\|_1\to0$. Then the following assertions are equivalent:*

(1) $\inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\underline{\mathrm{Osc}}_t(\varphi)>0$.
(2) *There is a $c>0$ such that for all time-nets $\tau=\{t_i\}_{i=0}^n\in\mathcal{T}$ one has* $\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))}\geqslant c\|\tau\|_\theta$.

Section 7: We provide with Theorem 7.1 an interpolation theorem adapted to gradient estimates in the Lévy setting which is formulated in a general context and for this reason of possible independent interest.

Section 8: We return to a Lévy processes $X=(X_t)_{t\in[0,T]}$ that is a pure jump L_2-martingale with a non-zero *Lévy measure* ν and fix a probability measure ρ on $\mathcal{B}(\mathbb{R})$. In Definition 8.5 we introduce a linear space $\mathrm{Dom}(\Gamma^0_\rho)$ of Borel functions $f:\mathbb{R}\to\mathbb{R}$ and the operator

$$\Gamma^0_{t,\rho}:\mathrm{Dom}(\Gamma^0_\rho)\to\mathbb{R}\quad\text{with}\quad\langle f,\Gamma^0_{t,\rho}\rangle:=\int_{\mathbb{R}\setminus\{0\}}\frac{\mathbb{E}f(z+X_{T-t})-\mathbb{E}f(X_{T-t})}{z}\rho(\mathrm{d}z).$$

In the special case $\rho\ll\nu$ with $\rho(\mathrm{d}z)=D(z)z^2\nu(\mathrm{d}z)/\int_{\mathbb{R}}D(z)z^2\nu(\mathrm{d}z)$, where $D\in L_2(\mathbb{R},z^2\nu(\mathrm{d}z))$ is non-negative with $\int_{\mathbb{R}}D(z)z^2\nu(\mathrm{d}z)>0$, these operators satisfy (formally)

$$\langle f(\cdot+x),\Gamma^0_{t,\rho}\rangle|_{x=X_t}=\int_{\mathbb{R}}\left[\frac{1}{t}\int_0^t D_{s,z}\mathbb{E}^{\mathcal{F}_t}[f(X_T)]\,\mathrm{d}s\right]D(z)z^2\nu(\mathrm{d}z)/\int_{\mathbb{R}}D(z)z^2\nu(\mathrm{d}z)$$

for $t\in(0,T)$, see (D.3), that takes us back to (1.5). The deterministic operators $\Gamma^0_{t,\rho}$ will be the main tool to obtain estimates on stochastic gradients where we use that the operators $\Gamma^0_{t,\rho}$ are linear and deterministic and allow therefore for the application of interpolation techniques from Banach space theory. To understand $\Gamma^0_{t,\rho}$ as mathematical object we associate to the probability measure ρ (that was arbitrary) and to the process X a probability density $\gamma_{t,\rho}\in L_1(\mathbb{R})$ for which it follows from Theorem 8.10 that in a distributional sense

$$\Gamma^0_{t,\rho}=-D\gamma_{t,\rho},$$

i.e. $\Gamma^0_{t,\rho}$ can be seen as a derivative of a distribution of L_1-type. Because $(\langle f(\cdot+x),\Gamma^0_{t,\rho}\rangle|_{x=X_t})_{t\in[0,T)}$ will be a martingale under our assumptions, we let

$$(\varphi_t(f,\rho))_{t\in[0,T)}\quad\text{be a càdlàg version of}\quad(\langle f(\cdot+x),\Gamma^0_{t,\rho}\rangle|_{x=X_t})_{t\in[0,T)}.\tag{1.14}$$

Section 8.3 (UPPER BOUNDS FOR GRADIENTS): We introduce

$$|||f|||^2_{\rho,\alpha}:=\frac{2\alpha}{T^{2\alpha}}\int_0^T(T-t)^{2\alpha-1}\sup_{x\in\mathbb{R}}|\langle f(\cdot+x),\Gamma^0_{t,\rho}\rangle|^2\mathrm{d}t$$

for $\alpha>0$ and obtain as a corollary of Theorem 8.11:

Theorem 1.3. *For $\alpha>0$ and $f\in\mathrm{Dom}(\Gamma^0_\rho)$ one has*

$$\sup_{a\in[0,T)}\left(\frac{T-a}{T}\right)^\alpha\|\varphi_a(f,\rho)\|_{L_\infty}+\left\|\mathcal{I}^\alpha\varphi(f,\rho)-\varphi_0(f,\rho)\right\|_{\mathrm{BMO}_2([0,T))}\leqslant 4|||f|||_{\rho,\alpha}.\tag{1.15}$$

To estimate $|||f|||_{\rho,\alpha}$ in the next step, for $\eta\in[0,1]$ and $s\in[0,T]$ we introduce

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)}:=\inf_P\left\{\int_{\mathbb{R}\setminus\{0\}}P(z)^{1-\eta}\rho(\mathrm{d}z)\right\}\in[0,\infty],$$

where the infimum is taken over all measurable $P:\mathbb{R}\setminus\{0\}\to[0,\infty)$ such that

$$\frac{\|\mathbb{P}_{z+X_s}-\mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|}\leqslant P(z)\quad\text{for}\quad z\in\mathbb{R}\setminus\{0\}.$$

Then Theorem 8.9 verifies

$$|\langle f,\Gamma^0_{t,\rho}\rangle|\leqslant\|f\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})}\|X_{T-t}\|_{\mathrm{TV}(\rho,\eta)}$$

which serves as end-point estimates in the interpolation Theorem 7.1 to get in Corollary 8.13 that

$$|||f|||_{\rho,\alpha}\leqslant c_{(1.16)}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}\quad\text{for}\quad\alpha:=\frac{1-(\varepsilon+\eta)}{\beta}\tag{1.16}$$

under the assumptions

$$\sup_{n\geqslant 1} 2^{\varepsilon n}\rho(\{2^{-n}\leqslant |z|<2^{-n+1}\})<\infty \quad \text{and} \quad \sup_{s\in(0,T]}\ \sup_{z\in\operatorname{supp}(\nu)\setminus\{0\}} s^{\frac{1}{\beta}}\frac{\|\mathbb{P}_{z+X_s}-\mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|}<\infty \tag{1.17}$$

for $(\varepsilon,\beta)\in(0,1)\times(1,2)$ and $\eta\in(0,1-\varepsilon)$. Combining (1.16) and (1.15) gives

$$\sup_{a\in[0,T)}\left(\frac{T-a}{T}\right)^{\alpha}\|\varphi_a(f,\rho)\|_{L_\infty}+\left\|\mathcal{I}^{\alpha}\varphi(f,\rho)-\varphi_0(f,\rho)\right\|_{\mathrm{BMO}_2([0,T))}\leqslant 4c_{(1.16)}\|f\|_{\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})}. \tag{1.18}$$

This estimate develops further the result [39, Theorem 1.3] as explained in Remark 8.22. One application of (1.18) is that we are now in a position to apply Theorem 4.7 (which corresponds to (1)⇒ (2) in Theorem 1.1).

Section 8.4 (LOWER BOUNDS FOR GRADIENTS): We consider the case

$$\rho(\mathrm{d}z):=z^2\nu(\mathrm{d}z)\Big/\int_{\mathbb{R}} z^2\nu(\mathrm{d}z) \tag{1.19}$$

where ν is the Lévy measure. This case yields to the gradients appearing in the classical Galtchouk-Kunita-Watanabe projection. We assume the following bounds that are the counterpart to the upper bounds in (1.17):

$$\rho([-d,d])\geqslant cd^{2-\beta} \quad \text{and} \quad \inf_{|v|\vee|z|\leqslant c' s^{\frac{1}{\beta}},z\neq 0}\frac{\mathbb{P}(X_s\in v+[-z^+,z^-))}{|z|}\geqslant c'' s^{-\frac{1}{\beta}} \tag{1.20}$$

for $d\in(0,d_0]$ and $s\in(0,T]$, respectively, where $c,c',c'',d_0>0$ are constants and z^+ and z^- are the positive and negative part of z. In the case of β-stable like processes as in Section 8.5 we have that (1.17) is satisfied with $\varepsilon:=2-\beta$ and (1.20) is satisfied. For the fractional smoothness α in the upper bound we get then

$$\alpha=\frac{1-(\varepsilon+\eta)}{\beta}=\frac{1-(2-\beta+\eta)}{\beta}=1-\frac{1+\eta}{\beta}.$$

This coincides with the lower bound we get for η-Hölder continuous functions from Theorem 8.20:

Theorem 1.4. *Let $\eta\in(0,1)$ and $\beta\in[1+\eta,2)$ and assume $\|X_s\|_{\mathrm{TV}(\rho,\eta)}<\infty$ for $s\in(0,T]$. Suppose that (1.19) and (1.20) are satisfied. If $f(x):=(x\vee 0)^{\eta}\in \mathrm{H\ddot{o}l}_{\eta}(\mathbb{R})$, then*

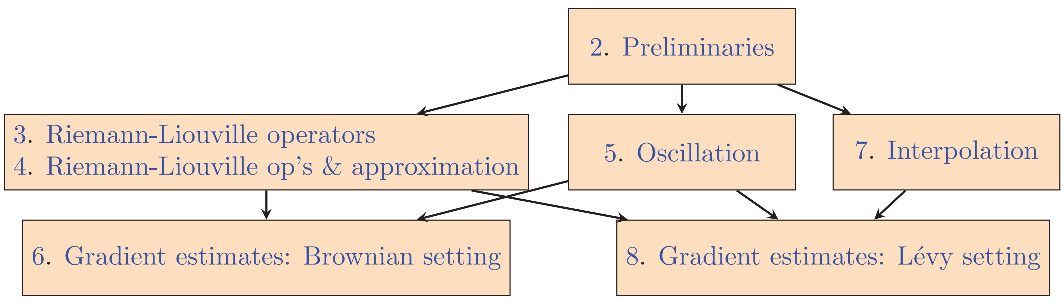
$$c_{(1.4)}:=\inf_{t\in[0,T)}(T-t)^{1-\frac{1+\eta}{\beta}}\langle f,\Gamma^0_{t,\rho}\rangle>0.$$

Now we combine the maximal oscillation of $(\varphi_t)_{t\in[0,T)}$ (Theorem 8.19(2)) and Theorem 1.4 to deduce that, for $\varphi=\varphi(f,\rho)$ given in (1.14) ,

$$\underline{\mathrm{Osc}}_t(\varphi)\geqslant\frac{1}{2}\|\varphi_t-\varphi_0\|_{L_\infty}\geqslant\frac{1}{2}\left[\|\varphi_t\|_{L_\infty}-|\varphi_0|\right]\geqslant\frac{1}{2}\left[\frac{c_{(1.4)}}{(T-t)^{1-\frac{1+\eta}{\beta}}}-|\varphi_0|\right].$$

Section 8.5 discusses the application of the results to β-stable like processes.

The sections of the article interact as follows:



2. Preliminaries

2.1. **General notation.** We let $\mathbb{N} := \{1, 2, \ldots\}$ and $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$. For $a, b \in \mathbb{R}$ we use $a \vee b := \max\{a, b\}$, $a \wedge b := \min\{a, b\}$, $a^+ := a \vee 0$, $a^- := (-a) \vee 0$, and for $A, B \geqslant 0$ and $c \geqslant 1$ the notation $A \sim_c B$ for $\frac{1}{c} B \leqslant A \leqslant cB$. The corresponding one-sided inequalities are abbreviated by $A \succeq_c B$ and $A \preceq_c B$. Given $x \in \mathbb{R}$, $\operatorname{sign}(x) := 1$ for $x \geqslant 0$ and $\operatorname{sign}(x) := -1$ for $x < 0$ is the standard sign function, and we agree about $0^0 := 1$. For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable map $X : \Omega \to \mathbb{R}^d$, where $\mathbb{R}^d$ is equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ generated by the open sets, the law of X is denoted by $\mathbb{P}_X$. Given $p \in (0, \infty]$ and a measure space $(\Omega, \mathcal{F}, \mu)$, we use the standard Lebesgue spaces $L_p(\Omega, \mathcal{F}, \mu)$ and denote by $L_0(\Omega, \mathcal{G})$ the space of all $\mathcal{G}$-measurable maps $X : \Omega \to \mathbb{R}$. We drop the corresponding measure space in the notation if there is no risk of confusion. Given a (finite) signed measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we denote by $|\mu| := \mu^+ + \mu^-$ its *variation* and by $\|\mu\|_{\mathrm{TV}} := |\mu|(\mathbb{R})$ its *total variation*. The Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ will be denoted by λ. For two measures μ and ν on a measurable space $(\Omega, \mathcal{F})$ we write $\nu \ll \mu$ if ν is absolutely continuous with respect to μ. For a set $A \in \mathcal{F}$ with $\mu(A) \in (0, \infty)$ we let μ_A be the normalized restriction of μ to the trace σ-algebra $\mathcal{F}|_A$. For $0 < p \leqslant q \leqslant \infty$, σ-finite measure spaces (M, Σ, μ) and $(N, \mathcal{N}, \nu)$, and a measurable map $f : M \times N \to [0, \infty)$ we use the inequality

$$\left\| \|f\|_{L_p(\mu)} \right\|_{L_q(\nu)} \leqslant \left\| \|f\|_{L_q(\nu)} \right\|_{L_p(\mu)}. \tag{2.1}$$

2.2. **Support of a measure.** Let μ be a measure on $\mathcal{B}(\mathbb{R}^d)$, then $\operatorname{supp}(\mu)$ denotes the closed set $\{x \in \mathbb{R}^d : \mu(U_\varepsilon(x)) > 0 \text{ for all } \varepsilon > 0\}$, where $U_\varepsilon(x)$ is the open euclidean ball centered at x with radius $\varepsilon > 0$. Given a random variable $X : \Omega \to \mathbb{R}^d$, we let $\operatorname{supp}(X) := \operatorname{supp}(\mathbb{P}_X)$. One knows that $\mathbb{P}(\{X \in \operatorname{supp}(X)\}) = 1$ and that for independent random variables $X : \Omega \to \mathbb{R}^m$ and $Y : \Omega \to \mathbb{R}^n$ it holds $\operatorname{supp}((X, Y)) = \operatorname{supp}(X) \times \operatorname{supp}(Y)$. Finally, for a random variable $X : \Omega \to \mathbb{R}^d$ and a Borel measurable $H : \mathbb{R}^d \to \mathbb{R}$ that is continuous on $\operatorname{supp}(X)$ (with respect to the induced topology) it holds that $\|H(X)\|_{L_\infty(\Omega, \mathcal{F}, \mathbb{P})} = \sup_{x \in \operatorname{supp}(X)} |H(x)|$.

2.3. **Interpolation spaces.** We will only consider Banach spaces over $\mathbb{R}$. Let (E_0, E_1) be a couple of Banach spaces such that E_0 and E_1 are continuously embedding into some topological Hausdorff space X ((E_0, E_1) is called an interpolation couple). We equip $E_0 + E_1 := \{x = x_0 + x_1 : x_i \in E_i\}$ with the norm $\|x\|_{E_0 + E_1} := \inf\{\|x_0\|_{E_0} + \|x_1\|_{E_1} : x_i \in E_i, x = x_0 + x_1\}$ and $E_0 \cap E_1$ with the norm $\|x\|_{E_0 \cap E_1} := \max\{\|x\|_{E_0}, \|x\|_{E_1}\}$ to get Banach spaces $E_0 \cap E_1 \subseteq E_0 + E_1$. For $x \in E_0 + E_1$ and $v \in (0, \infty)$ we define the K-functional

$$K(v, x; E_0, E_1) := \inf\{\|x_0\|_{E_0} + v\|x_1\|_{E_1} : x = x_0 + x_1\}.$$

Given $(\theta, q) \in (0, 1) \times [1, \infty]$ we set

$$(E_0, E_1)_{\theta, q} := \left\{ x \in E_0 + E_1 : \|x\|_{(E_0, E_1)_{\theta, q}} := \left\| v \mapsto v^{-\theta} K(v, x; E_0, E_1) \right\|_{L_q\left((0, \infty), \frac{\mathrm{d}v}{v}\right)} < \infty \right\}.$$

We obtain a family of Banach spaces $\left((E_0, E_1)_{\theta, q}, \|\cdot\|_{(E_0, E_1)_{\theta, q}}\right)$ with the lexicographical ordering

$$(E_0, E_1)_{\theta, q_0} \subseteq (E_0, E_1)_{\theta, q_1} \quad \text{for all} \quad \theta \in (0, 1) \text{ and } 1 \leqslant q_0 < q_1 \leqslant \infty,$$

and, under the additional assumption that $E_1 \subseteq E_0$ with $\|x\|_{E_0} \leqslant c\|x\|_{E_1}$ for some $c > 0$,

$$(E_0, E_1)_{\theta_0, q_0} \subseteq (E_0, E_1)_{\theta_1, q_1} \quad \text{for all} \quad 0 < \theta_1 < \theta_0 < 1 \text{ and } q_0, q_1 \in [1, \infty].$$

Given a linear operator $T : E_0 + E_1 \to F_0 + F_1$ with $T : E_i \to F_i$ for $i = 0, 1$, we use that the real interpolation method is an exact interpolation functor, i.e.

$$\|T : (E_0, E_1)_{\theta, q} \to (F_0, F_1)_{\theta, q}\| \leqslant \|T : E_0 \to F_0\|^\theta \|T : E_1 \to F_1\|^{1-\theta}. \tag{2.2}$$

For more information about the real interpolation method the reader is referred to [5, 7, 43]. Given a Banach space E and $(q, s) \in [1, \infty] \times \mathbb{R}$, we will use the Banach spaces

$$\ell_q^s(E) := \{(x_k)_{k=0}^\infty \subseteq E : \|(x_k)_{k=0}^\infty\|_{\ell_q^s(E)} := \|(2^{ks}\|x_k\|_E)_{k=0}^\infty\|_{\ell_q} < \infty\}$$

and the notation $\ell_q(E) := \ell_q^0(E)$. For $q_0, q_1, q \in [1, \infty]$ and $s_0, s_1 \in \mathbb{R}$ with $s_0 \neq s_1$, and $\theta \in (0, 1)$, one has according to [7, Theorem 5.6.1] that

$$(\ell_{q_0}^{s_0}(E), \ell_{q_1}^{s_1}(E))_{\theta, q} = \ell_q^s(E) \quad \text{where} \quad s := (1 - \theta)s_0 + \theta s_1 \tag{2.3}$$

and there is a $c_{(2.4)} \geqslant 1$ that depends at most on $(s_0, s_1, q_0, q_1, \theta, q)$ such that

$$\|\cdot\|_{\ell_q^s(E)} \sim_{c_{(2.4)}} \|\cdot\|_{(\ell_{q_0}^{s_0}(E),\ell_{q_1}^{s_1}(E))_{\theta,q}}. \tag{2.4}$$

2.4. **Function spaces.** We let $B_b(\mathbb{R})$ be the Banach space of bounded Borel functions $f\colon \mathbb{R} \to \mathbb{R}$ with $\|f\|_{B_b(\mathbb{R})} := \sup_{x\in\mathbb{R}} |f(x)|$, $C_b^0(\mathbb{R})$ be the closed subspace of $B_b(\mathbb{R})$ of continuous functions vanishing at zero, and $C_b^\infty(\mathbb{R}) \subseteq B_b(\mathbb{R})$ the infinitely often differentiable functions such that the derivatives satisfy $f^{(k)} \in B_b(\mathbb{R})$, $k \geqslant 1$. The space $C^1(\mathbb{R})$ consists of differentiable functions with continuous derivative and $C^\infty(\mathbb{R})$ of the functions that are infinitely often differentiable. For $\eta \in [0,1]$ we use the Hölder spaces

$$\begin{aligned} \mathrm{Höl}_\eta(\mathbb{R}) &:= \left\{ f\colon \mathbb{R}\to\mathbb{R} \text{ Borel} \; ; \; |f|_\eta := \sup_{-\infty<x<y<\infty} \frac{|f(x)-f(y)|}{|x-y|^\eta} < \infty \right\}, \\ \mathrm{Höl}_\eta^0(\mathbb{R}) &:= \{ f \in \mathrm{Höl}_\eta(\mathbb{R}) : f(0) = 0\}, \\ \mathrm{Höl}_{\eta,q}^0(\mathbb{R}) &:= (C_b^0(\mathbb{R}), \mathrm{Höl}_1^0(\mathbb{R}))_{\eta,q} \quad \text{for} \quad (\eta,q) \in (0,1)\times[1,\infty]. \end{aligned}$$

Note that we can define the Banach space $C_b^0(\mathbb{R}) + \mathrm{Höl}_1^0(\mathbb{R})$, so that $(C_b^0(\mathbb{R}), \mathrm{Höl}_1^0(\mathbb{R}))$ forms an interpolation pair. If we use on $C_b^0(\mathbb{R})$ the equivalent norm $\|f\|_{C_b^0(\mathbb{R})}^0 := \sup\{|f(x)-f(y)| : x,y \in \mathbb{R}\}$, then $\frac{1}{2}\|f\|_{C_b^0(\mathbb{R})}^0 \leqslant \|f\|_{C_b^0(\mathbb{R})} \leqslant \|f\|_{C_b^0(\mathbb{R})}^0$ and build with this norm the interpolation spaces $\mathrm{Höl}_{\eta,q}^0(\mathbb{R})$ and denote the norms by $\|f\|_{\mathrm{Höl}_{\eta,q}^0(\mathbb{R})}^0$, then we get the 'translation invariance' (useful later for us)

$$\|f\|_{\mathrm{Höl}_{\eta,q}^0(\mathbb{R})}^0 = \|f(x+\cdot) - f(x)\|_{\mathrm{Höl}_{\eta,q}^0(\mathbb{R})}^0 \quad \text{for all} \quad x \in \mathbb{R}.$$

By the reiteration theorem (see [7, Theorem 3.5.3] or [5, Theorem 5.2.4]) it follows

$$(\mathrm{Höl}_{\eta_0,q_0}^0(\mathbb{R}), \mathrm{Höl}_{\eta_1,q_1}^0(\mathbb{R}))_{\theta,q} = \mathrm{Höl}_{\eta,q}^0(\mathbb{R}) \tag{2.5}$$

for $\theta, \eta_0, \eta_1 \in (0,1)$ with $\eta_0 \neq \eta_1$, $q, q_0, q_1 \in [1,\infty]$, $\eta := (1-\theta)\eta_0 + \theta\eta_1$, where the norms are equivalent up to a multiplicative constant. By the above definitions we obtain Banach space by $(\mathrm{Höl}_\eta^0(\mathbb{R}), |\cdot|_\eta)$ and for $\eta \in (0,1)$ we have that $\mathrm{Höl}_{\eta,\infty}^0(\mathbb{R}) = \mathrm{Höl}_\eta^0(\mathbb{R})$ with equivalent norms up to a multiplicative constant (a direct proof can be obtained by an adaptation of [33, Lemma A.3], see also [43, Theorem 2.7.2/1]).

2.5. **Stochastic basis.** We fix a time horizon $T \in (0,\infty)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a right continuous filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ such that $\mathcal{F}_0$ is generated by the $\mathbb{P}$-null sets and $\mathcal{F} = \mathcal{F}_T$. For

$$\mathbb{I} = [0,T] \quad \text{or} \quad \mathbb{I} = [0,T)$$

we denote by $\mathrm{CL}(\mathbb{I})$ the set of $\mathbb{F}$-adapted *càdlàg* (right continuous with left limits) processes $Y = (Y_t)_{t\in\mathbb{I}}$, by $\mathrm{CL}^+(\mathbb{I})$ the sub-set of $Y \in \mathrm{CL}(\mathbb{I})$ with $Y_t(\omega) \geqslant 0$ on $\mathbb{I}\times\Omega$, and by $\mathrm{CL}_0(\mathbb{I})$ the sub-set of $Y \in \mathrm{CL}(\mathbb{I})$ with $Y_0 \equiv 0$. For $Y \in \mathrm{CL}(\mathbb{I})$ we use

(1) $Y^* = (Y_t^*)_{t\in\mathbb{I}}$ with $Y_t^* = \sup_{s\in[0,t]} |Y_s|$,
(2) $\Delta Y = (\Delta Y_t)_{t\in\mathbb{I}}$ with $\Delta Y_t := Y_t - Y_{t-}$, where $Y_{0-} := Y_0$ and $Y_{t-} := \lim_{s<t,\, s\uparrow t} Y_s$ for $t > 0$.

The collection of all stopping times $\rho\colon \Omega \to [0,t]$ is denoted by $\mathcal{S}_t$. We write $\mathbb{E}^{\mathcal{G}}[X]$ for the conditional expectation of X given $\mathcal{G}$. The usual conditions imposed on $\mathbb{F}$ allow us to assume that every martingale adapted to this filtration is càdlàg. Given a càdlàg L_2-martingale $X = (X_t)_{t\in\mathbb{I}}$, the sharp bracket process is denoted by $\langle X\rangle = (\langle X\rangle_t)_{t\in\mathbb{I}}$ and the square bracket process by $[X] = ([X]_t)_{t\in\mathbb{I}}$ (see [14, Chapter VII]). In particular, the process $\langle X\rangle = (\langle X\rangle_t)_{t\in\mathbb{I}}$ is the unique (up to indistinguishability) non-decreasing, predictable, càdlàg process with $\langle X\rangle_0 \equiv 0$ such that $(X_t^2 - \langle X\rangle_t)_{t\in\mathbb{I}}$ is a martingale.

2.6. **Bounded mean oscillation and regular weights.** We use the following weighted BMO spaces, where we agree about $\inf\emptyset := \infty$ in this subsection.

Definition 2.1. Let $p \in (0,\infty)$.

(1) For $Y \in \mathrm{CL}_0(\mathbb{I})$ and $\Phi \in \mathrm{CL}^+(\mathbb{I})$ we let $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} := \inf c$, where the infimum is taken over all $c \in [0,\infty)$ such that, for all $t \in \mathbb{I}$ and $\rho \in \mathcal{S}_t$,

$$\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_{\rho-}|^p] \leqslant c^p \Phi_\rho^p \text{ a.s.} \tag{2.6}$$

(2) For $Y \in \mathrm{CL}_0(\mathbb{I})$ and $\Phi \in \mathrm{CL}^+(\mathbb{I})$ we let $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} := \inf c$, where the infimum is taken over all $c \in [0,\infty)$ such that, for all $t \in \mathbb{I}$ and $\rho \in \mathcal{S}_t$,

$$\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_\rho|^p] \leqslant c^p \Phi_\rho^p \quad \text{a.s.} \tag{2.7}$$

For $\|Y\|_\Theta < \infty$ we write $Y \in \Theta$ with $\Theta \in \{\mathrm{BMO}_p^\Phi(\mathbb{I}), \mathrm{bmo}_p^\Phi(\mathbb{I})\}$. If $\Phi \equiv 1$, then we use the notation $\mathrm{BMO}_p(\mathbb{I})$ and $\mathrm{bmo}_p(\mathbb{I})$, respectively.

If $Y_0 \equiv 0$ is not necessarily satisfied, then we use the notation $\|Y - Y_0\|_{\mathrm{BMO}^\Phi(\mathbb{I})}$ for $\|(Y_t - Y_0)_{t\in\mathbb{I}}\|_{\mathrm{BMO}^\Phi(\mathbb{I})}$. If $Y \in \mathrm{CL}_0(\mathbb{I})$ has continuous paths a.s., then $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} = \|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})}$. The theory of classical non-weighted BMO-martingales can be found in [14, Ch.VII] or [35, Ch.IV]; non-weighted bmo-martingales were mentioned in [14, Ch.VII, Remark 87] and used after that in [11, 13]. The BMO_p^Φ space was introduced and discussed in [20]. Some relations between bmo_p^Φ and BMO_p^Φ, that are necessary for us, are discussed in the appendix below. Next we recall (and adapt) the class $\mathcal{SM}_p$, introduced in [20, Definition 3]:

Definition 2.2. For $p \in (0,\infty)$ and $\Phi \in \mathrm{CL}^+(\mathbb{I})$ we let $\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})} := \inf c$, where the infimum is taken over all $c \in [1,\infty)$ such that for all stopping times $\rho : \Omega \to \mathbb{I}$ one has

$$\mathbb{E}^{\mathcal{F}_\rho}\left[\sup_{\rho\leqslant t\in\mathbb{I}} \Phi_t^p\right] \leqslant c^p \Phi_\rho^p \quad \text{a.s.}$$

If $\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})} < \infty$, then we write $\Phi \in \mathcal{SM}_p(\mathbb{I})$.

By choosing $\rho \equiv 0$, $\Phi \in \mathcal{SM}_p(\mathbb{I})$ implies that $\mathbb{E}\sup_{t\in\mathbb{I}} \Phi_t^p < \infty$. Moreover, it follows directly from the definition that $\mathcal{SM}_p(\mathbb{I}) \subseteq \mathcal{SM}_r(\mathbb{I})$ whenever $0 < r \leqslant p < \infty$. Simplifications in Definition 2.1 and Definition 2.2 and relations between the BMO- and bmo-spaces are recalled in Appendix A. If $p \in (1,\infty)$ and Φ is a martingale, then $\Phi \in \mathrm{CL}^+(\mathbb{I})$ is equivalent to the standard reverse Hölder condition $\mathbb{E}^{\mathcal{F}_a}[\Phi_t^p] \leqslant d^p \Phi_a^p$ a.s. for $0 \leqslant a \leqslant t < T$.

2.7. **Uniform quantization and time-nets.** For $\theta \in (0,1]$ we introduce the non-uniform time-nets $\tau_n^\theta = \{t_{i,n}^\theta\}_{i=0}^n$ with

$$t_{i,n}^\theta := T - T\left(1 - \tfrac{i}{n}\right)^{\frac{1}{\theta}} \tag{2.8}$$

for $i = 0,\dots,n$, that are characterized by the uniform quantization property

$$\frac{\theta}{T^\theta}\int_{t_{i-1,n}^\theta}^{t_{i,n}^\theta} (T-u)^{\theta-1}\mathrm{d}u = \frac{1}{n} \quad \text{for} \quad i = 1,\dots,n.$$

We define the set of all *deterministic* time-nets

$$\mathcal{T} := \{\tau = \{t_i\}_{i=0}^n : 0 = t_0 < t_1 < \cdots < t_n = T, \, n \in \mathbb{N}\}$$

and, for $\theta \in (0,1]$ and $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$,

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

Note that

$$\|\tau_n^\theta\|_1 \leqslant \frac{T}{\theta n} \quad \text{and} \quad \|\tau_n^\theta\|_\theta \leqslant \frac{T^\theta}{\theta n}, \tag{2.9}$$

and

$$\frac{t_i - u}{(T-u)^{1-\theta}} \leqslant \frac{t_i - t_{i-1}}{(T - t_{i-1})^{1-\theta}} \quad \text{for} \quad u \in [t_{i-1}, t_i] \cap [0,T). \tag{2.10}$$

3. Riemann-Liouville type operators

Riemann-Liouville operators are a central object and tool in fractional calculus. It is natural and useful to extend them to random frameworks. There are two principal approaches: Directly translating the approach from fractional calculus, that uses Volterra kernels, leads to the notion of fractional processes, in particular fractional martingales. In our setting one would take a càdlàg process K and would consider

$$t \mapsto \int_0^t (t-u)^{\alpha-1} K_u \mathrm{d}u.$$

This yields to an approach natural for path-wise fractional calculus of stochastic processes and is used, for example, for Gaussian processes [26]. For our purpose we use the different approach

$$t \mapsto \int_0^T (T-u)^{\alpha-1} K_{u\wedge t} \mathrm{d}u$$

to define $\mathcal{I}_t^\alpha K$ in Definition 3.1 below. The idea behind the operator $\mathcal{I}^\alpha$ is to remove or reduce singularities of a càdlàg process $(K_t)_{t\in[0,T)}$ when $t \uparrow T$. As we see in Theorem 3.6 below, this approach is the right one to handle fractional smoothness in the Malliavin sense and in the sense of interpolation theory. One basic difference to the Volterra-kernel approach is that, starting with a (sub-, super-) martingale φ, we again obtain a (sub-, super-) martingale $\mathcal{I}_t^\alpha \varphi$. This second approach was used in [23, Definition 4.2], [24, Section 4], and [2], and relates to fractional integral transforms of martingales (see, for example, [3]). This corresponds to equation (3.3) of our Proposition 3.8.

Definition 3.1. For $\alpha > 0$ and a càdlàg function $K : [0,T) \to \mathbb{R}$ we define $\mathcal{I}^\alpha K := (\mathcal{I}_t^\alpha K)_{t\in[0,T)}$ by

$$\mathcal{I}_t^\alpha K := \frac{\alpha}{T^\alpha} \int_0^T (T-u)^{\alpha-1} K_{u\wedge t} \mathrm{d}u.$$

Moreover, for $\alpha = 0$ we let $\mathcal{I}_t^0 K := K_t$.

The càdlàg property implies the boundedness of K on any compact interval of $[0,T)$. Therefore, $\mathcal{I}^\alpha K$ is well-defined and càdlàg on $[0,T)$. The above definition can be re-formulated in terms of the classical Riemann-Liouville operator $\mathcal{R}_a^\alpha(f) := \frac{1}{\Gamma(\alpha)} \int_0^a (a-u)^{\alpha-1} f(u)\mathrm{d}u$ by

$$\mathcal{R}_T^\alpha(K^{(t)}) = \frac{T^\alpha}{\Gamma(\alpha+1)} \mathcal{I}_t^\alpha K \quad \text{with} \quad K_u^{(t)} := K_{u\wedge t}$$

where we compute the Riemann-Liouville operator, applied to the function $u \mapsto K_u^{(t)}$, at $a = T$. We use a different normalisation as we want to interpret the kernel in the Riemann-Liouville integral as density of a probability measure. It follows directly from the definition that we have, for $\alpha \geqslant 0$,

$$\mathcal{I}_t^\alpha K = \frac{\alpha}{T^\alpha} \int_0^t (T-u)^{\alpha-1} K_u \mathrm{d}u + \left(\frac{T-t}{T}\right)^\alpha K_t. \tag{3.1}$$

In the following we only need $\mathcal{I}^\alpha K$ for $\alpha \geqslant 0$. However, to derive an inversion formula we extend the definition to the case $\alpha < 0$ and prove that there is a group structure behind:

Proposition 3.2. *Define for $\alpha < 0$, a càdlàg function $K : [0,T) \to \mathbb{R}$, and $t \in [0,T)$, $\mathcal{I}_t^\alpha K$ by formula (3.1). Then*

(1) *$\mathcal{I}_t^\alpha(\mathcal{I}_\cdot^\beta K) = \mathcal{I}_t^{\alpha+\beta} K$ for all $\alpha, \beta \in \mathbb{R}$,*
(2) *$\mathcal{I}_t^{-\alpha}(\mathcal{I}_\cdot^\alpha K) = K_t$ for all $\alpha \in \mathbb{R}$.*

Proof. As (2) follows from (1), we only need to check (1). Here we get that

$$\begin{aligned}
\mathcal{I}_t^\alpha(\mathcal{I}^\beta K) &= \frac{\alpha}{T^\alpha} \int_0^t (T-u)^{\alpha-1} \mathcal{I}_u^\beta K \mathrm{d}u + \left(\frac{T-t}{T}\right)^\alpha \mathcal{I}_t^\beta K \\
&= \frac{\alpha}{T^\alpha} \int_0^t (T-u)^{\alpha-1} \left(\frac{\beta}{T^\beta} \int_0^u (T-v)^{\beta-1} K_v \mathrm{d}v + \left(\frac{T-u}{T}\right)^\beta K_u \right) \mathrm{d}u \\
&\quad + \left(\frac{T-t}{T}\right)^\alpha \left(\frac{\beta}{T^\beta} \int_0^t (T-u)^{\beta-1} K_u \mathrm{d}u + \left(\frac{T-t}{T}\right)^\beta K_t \right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\alpha\beta}{T^{\alpha+\beta}} \int_0^t (T-u)^{\alpha-1} \int_0^u (T-v)^{\beta-1} K_v \mathrm{d}v\mathrm{d}u + \frac{\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\alpha+\beta-1} K_u \mathrm{d}u \\
&\quad + \frac{\beta(T-t)^\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\beta-1} K_u \mathrm{d}u + \left(\frac{T-t}{T}\right)^{\alpha+\beta} K_t \\
&= \frac{\beta}{T^{\alpha+\beta}} \int_0^t (T-v)^{\beta-1} K_v \left((T-v)^\alpha - (T-t)^\alpha\right) \mathrm{d}v + \frac{\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\alpha+\beta-1} K_u \mathrm{d}u \\
&\quad + \frac{\beta(T-t)^\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\beta-1} K_u \mathrm{d}u + \left(\frac{T-t}{T}\right)^{\alpha+\beta} K_t \\
&= \frac{\beta}{T^{\alpha+\beta}} \int_0^t (T-v)^{\alpha+\beta-1} K_v \mathrm{d}v - \frac{\beta(T-t)^\alpha}{T^{\alpha+\beta}} \int_0^t (T-v)^{\beta-1} K_v \mathrm{d}v \\
&\quad + \frac{\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\alpha+\beta-1} K_u \mathrm{d}u + \frac{\beta(T-t)^\alpha}{T^{\alpha+\beta}} \int_0^t (T-u)^{\beta-1} K_u \mathrm{d}u + \left(\frac{T-t}{T}\right)^{\alpha+\beta} K_t \\
&= \mathcal{I}_t^{\alpha+\beta} K. \qquad \square
\end{aligned}
$$

We continue with some more structural properties:

Proposition 3.3. *For a càdlàg function $K : [0,T) \to \mathbb{R}$ and $t \in [0,T)$ one has:*
(1) $\lim_{\alpha\downarrow 0} \mathcal{I}_t^\alpha K = K_t$.
(2) $\lim_{\alpha\uparrow\infty} \mathcal{I}_t^\alpha K = K_0$.
(3) $\Delta \mathcal{I}_t^\alpha K = \left(\frac{T-t}{T}\right)^\alpha \Delta K_t$ *for* $\alpha \in \mathbb{R}$.

Proof. (1) and (3) follow from (3.1), and (2) from the càdlàg property of K. $\square$

The particular case that the function K is a path of a càdlàg martingale φ is of our interest. The following statement is obvious, but useful:

Proposition 3.4. *If $\alpha \geqslant 0$ and $\varphi = (\varphi)_{t\in[0,T)}$ is a càdlàg martingale (càdlàg super-, or sub-martingale), then $(\mathcal{I}_t^\alpha \varphi)_{t\in[0,T)}$ is a càdlàg martingale (càdlàg super-, or sub-martingale).*

The following functional $[\varphi;\tau]$ measures the oscillation of a martingale along a time-net in terms of an area and can be considered as a square function. Besides this functional occurs in various approximation problems for stochastic integrals, the functional is particularly designed to deal with martingales non-closable in a certain sense. In Theorem 3.6 we characterize by the behaviour of this functional the degree of singularity of a martingale not closable in L_2. Moreover, under a certain regularity of the martingale we prove in Proposition 3.9 that this functional converges to a classical square function as the time-nets refine.

Definition 3.5. For a deterministic time-net $\tau = \{t_i\}_{i=0}^n$, $0 = t_0 < \cdots < t_n = T$, $a \in [0,T)$, and a càdlàg process $\varphi = (\varphi_t)_{t\in[0,T)}$ we let

$$
[\varphi;\tau]_a := \int_0^a \left| \varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u) \right|^2 \mathrm{d}u \in [0,\infty).
$$

Moreover, we define $[\varphi;\tau]_T := \lim_{a\uparrow T} [\varphi;\tau]_a \in [0,\infty]$.

Now we give in Theorem 3.6 a first link between the Riemann-Liouville type operators $\mathcal{I}_t^\alpha$, real interpolation, and the square function $[\varphi;\tau]$. To do this as simple as possible, we replace a martingale $\varphi = (\varphi_t)_{t\in[0,T)}$ by its discrete time version

$$
\varphi^d := (\varphi_{t_k})_{k=0}^\infty \quad \text{with} \quad t_k := T\left(1 - \frac{1}{2^k}\right).
$$

For the vector-valued interpolation we use $H := L_2(\Omega, \mathcal{F}, \mathbb{P})$ and the end-point spaces

$$
\begin{aligned}
\varphi^d \in \ell_2^{-\frac{1}{2}}(H) &\iff \int_0^T \|\varphi_t\|_{L_2}^2 \mathrm{d}t < \infty, \\
\varphi^d \in \ell_\infty(H) &\iff \|\varphi^d\|_{\ell_\infty(H)} = \sup_{t\in[0,T)} \|\varphi_t\|_{L_2} < \infty,
\end{aligned}
$$

where the first equivalence follows from (3.6) below and the spaces $\ell_q^s(H)$ and $\ell_\infty(H)$ were introduced in Section 2.3. The first condition, $\int_0^T \|\varphi_t\|_{L_2}^2 \mathrm{d}t < \infty$, is a typical condition on martingales that appear as gradient processes. The other end-point, $\sup_{t\in[0,T)} \|\varphi_t\|_{L_2} < \infty$, consists of the martingales φ that are closable in L_2. We will interpolate between these two end-points by the real interpolation method:

Theorem 3.6. *For $\theta \in (0,1)$, $\alpha := \frac{1-\theta}{2}$, and a càdlàg martingale $\varphi = (\varphi_t)_{t\in[0,T)} \subseteq L_2$ the following assertions are equivalent:*

(1) $\varphi^d \in (\ell_2^{-\frac{1}{2}}(H), \ell_\infty(H))_{\theta,2}$.
(2) $(\mathcal{I}_t^\alpha \varphi)_{t\in[0,T)}$ *is closable in* L_2.
(3) *There is a* $c > 0$ *such that* $\mathbb{E}[\varphi;\tau]_T \leqslant c\|\tau\|_\theta$ *for all* $\tau \in \mathcal{T}$.

Before we prove Theorem 3.6 let us comment on it:

Remark 3.7. From Item (2) we get for all $\varepsilon > 0$ a $t(\varepsilon) \in [0,T)$ such that for $s \in [t(\varepsilon),T)$ one has

$$\mathbb{E} \sup_{t\in[s,T)} \left| \int_s^T (\varphi_{u\wedge t} - \varphi_s)(T-u)^{\alpha-1} \frac{\alpha}{T^\alpha} \mathrm{d}u \right|^2 < \varepsilon. \tag{3.2}$$

Without the supremum the left-hand side is equal to $\mathbb{E}|\mathcal{I}_t^\alpha\varphi - \mathcal{I}_s^\alpha\varphi|^2$, the statement including the supremum follows from Doob's maximal inequality. The convergence in (3.2) is the replacement of the L_2- and a.s. convergence of φ in the case φ would be closable in L_2.

For the proof of Theorem 3.6 and later in the article we need the following Proposition 3.8. We remark that Proposition 3.2(1) for $\alpha, \beta \geqslant 0$ can be also understood from equation (3.3) of Proposition 3.8 in the martingale setting.

Proposition 3.8. *For $\alpha > 0$, a càdlàg martingale $\varphi = (\varphi_t)_{t\in[0,T)} \subseteq L_2$ and $0 \leqslant a < t < T$ one has, a.s.,*

$$\mathcal{I}_t^\alpha \varphi = \varphi_0 + \int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha \mathrm{d}\varphi_u, \tag{3.3}$$

$$\mathbb{E}^{\mathcal{F}_a}\Big[|\mathcal{I}_t^\alpha\varphi - \mathcal{I}_a^\alpha\varphi|^2\Big] = 2\alpha \mathbb{E}^{\mathcal{F}_a}\left[\int_a^T |\varphi_{u\wedge t} - \varphi_a|^2 \left(\frac{T-u}{T}\right)^{2\alpha-1} \frac{\mathrm{d}u}{T}\right], \tag{3.4}$$

$$\mathbb{E}^{\mathcal{F}_a}\Big[|\mathcal{I}_t^\alpha\varphi - \mathcal{I}_a^\alpha\varphi|^2\Big] + \left(\frac{T-a}{T}\right)^{2\alpha} |\varphi_a|^2 = 2\alpha \mathbb{E}^{\mathcal{F}_a}\left[\int_a^T |\varphi_{u\wedge t}|^2 \left(\frac{T-u}{T}\right)^{2\alpha-1} \frac{\mathrm{d}u}{T}\right]. \tag{3.5}$$

Proof. (3.3) We apply partial integration to $\left(\left(\frac{T-t}{T}\right)^\alpha \varphi_t\right)_{t\in[0,T)}$ and obtain, for $t \in [0,T)$, that

$$\left(\frac{T-t}{T}\right)^\alpha \varphi_t = \left(\frac{T-0}{T}\right)^\alpha \varphi_0 + \int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha \mathrm{d}\varphi_u - \frac{\alpha}{T^\alpha}\int_{(0,t]} (T-u)^{\alpha-1}\varphi_u \mathrm{d}u \text{ a.s.}$$

Taking the last term to the left side, we obtain (3.3). For (3.4) we use Itô's isometry to get, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_a}\Big[|\mathcal{I}_t^\alpha\varphi - \mathcal{I}_a^\alpha\varphi|^2\Big] &= \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} \left(\frac{T-u}{T}\right)^{2\alpha} \mathrm{d}[\varphi]_u\right] \\
&= \frac{1}{2\alpha T^{2\alpha}} \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} \int_{[u,T)} (T-v)^{2\alpha-1} \mathrm{d}v \mathrm{d}[\varphi]_u\right] \\
&= \frac{1}{2\alpha T^{2\alpha}} \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} \int_{(a,v\wedge t]} \mathrm{d}[\varphi]_u (T-v)^{2\alpha-1} \mathrm{d}v\right] \\
&= \frac{1}{2\alpha T^{2\alpha}} \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} |\varphi_{v\wedge t} - \varphi_a|^2 (T-v)^{2\alpha-1} \mathrm{d}v\right].
\end{aligned}$$

(3.5) follows directly from (3.4) and the orthogonality of $\varphi_{u\wedge t} - \varphi_a$ and φ_a. □

Proof of Theorem 3.6. Because $(\|\varphi_{t_k}\|_H)_{k=0}^\infty$ is non-decreasing we observe for $s \in \mathbb{R}$ that

$$\frac{\|(\varphi_{t_k})_{k=0}^\infty\|^2_{\ell_2^s(H)}}{2T^{2s}} = \sum_{k=0}^\infty (T-t_k)^{-1-2s}(t_{k+1}-t_k)\|\varphi_{t_k}\|_H^2 \sim_{c_{T,s}} \int_0^T (T-t)^{-1-2s}\|\varphi_t\|_H^2 \mathrm{d}t \tag{3.6}$$

for some $c_{T,s} \geqslant 1$. For $s := (1-\theta)\left(-\frac{1}{2}\right) + \theta 0$ (so that $-1-2s = -\theta$) we use Proposition 3.8 (equation (3.5)) with $a = 0$ to get

$$\int_0^T (T-t)^{-\theta}\|\varphi_t\|_H^2 \mathrm{d}t = \sup_{t\in[0,T)} \frac{T^{2\alpha}}{2\alpha}\mathbb{E}[|\mathcal{I}_t^\alpha \varphi - \varphi_0|^2 + |\varphi_0|^2] = \sup_{t\in[0,T)} \frac{T^{2\alpha}}{2\alpha}\mathbb{E}|\mathcal{I}_t^\alpha\varphi|^2.$$

Now the equivalence (1) $\Leftrightarrow$ (2) follows from (2.3) and (3.6). The equivalence (2) $\Leftrightarrow$ (3) follows from Theorem 4.7, equation (4.5), applied to $M := \varphi$, $\sigma \equiv 1$, $a := 0$, and $\mathcal{G} := \{\emptyset, \Omega\}$. □

We close this section with the connection between the square function $[\mathcal{I}^{\frac{1-\theta}{2}}\varphi]$ and $[\varphi; \tau_n^\theta]$:

Proposition 3.9. *Let $\theta \in (0,1]$ and $\varphi = (\varphi_t)_{t\in[0,T)}$ be a path-wise continuous martingale such that we can choose $\mathrm{d}[\varphi]_t = K_t \mathrm{d}t$ on $[0,T)\times\Omega$, where $(K_t)_{t\in[0,T)}$ is continuous, adapted, and satisfies $C_a := \sup_{(\omega,t)\in\Omega\times[0,a]} |K_t(\omega)| < \infty$ for all $a \in [0,T)$. Then, with the time-nets τ_n^θ from (2.8), one has*

$$\frac{2\theta}{T}\lim_n \left(n\,[\varphi;\tau_n^\theta]_a\right) = \left[\mathcal{I}^{\frac{1-\theta}{2}}\varphi\right]_a \quad \textit{in } L_p \quad \textit{for all} \quad (p,a) \in (0,\infty)\times[0,T).$$

Proof. Let $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$, $a \in [0,T)$, $s_i := t_i \wedge a$, and define $M_0 \equiv 0$ and, for $i = 1,\dots n$,

$$dM_i := \int_{(s_{i-1},s_i]} |\varphi_u - \varphi_{s_{i-1}}|^2 \mathrm{d}u - \int_{(s_{i-1},s_i]} (s_i - u)\mathrm{d}[\varphi]_u.$$

We obtain a martingale difference sequence $(dM_i)_{i=1}^n \subseteq L_2$ with respect to $(\mathcal{F}_{s_i})_{i=0}^n$. It is sufficient to consider $p \in [2,\infty)$. By the Burkholder-Davis-Gundy inequalities (with constant $\beta_p > 0$) and exploiting (2.1) we get

$$\begin{aligned}
\frac{1}{\beta_p}\left\|\sum_{i=1}^n dM_i\right\|_{L_p} &\leqslant \left(\sum_{i=1}^n \|dM_i\|_{L_p}^2\right)^{\frac{1}{2}} \\
&\leqslant \left(\sum_{i=1}^n \left\|\int_{(s_{i-1},s_i]} |\varphi_u - \varphi_{s_{i-1}}|^2 \mathrm{d}u\right\|_{L_p}^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^n \left\|\int_{(s_{i-1},s_i]} (s_i-u)\mathrm{d}[\varphi]_u\right\|_{L_p}^2\right)^{\frac{1}{2}} \\
&\leqslant \left(\sum_{i=1}^n \left|\int_{(s_{i-1},s_i]} \|\varphi_u - \varphi_{s_{i-1}}\|_{L_{2p}}^2 \mathrm{d}u\right|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^n \left\|\int_{(s_{i-1},s_i]} (s_i-u)\mathrm{d}[\varphi]_u\right\|_{L_p}^2\right)^{\frac{1}{2}} \\
&\leqslant \left(\sum_{i=1}^n \|\varphi_{s_i} - \varphi_{s_{i-1}}\|_{L_{2p}}^4 |s_i - s_{i-1}|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^n \left\|[\varphi]_{s_i} - [\varphi]_{s_{i-1}}\right\|_{L_p}^2 |s_i - s_{i-1}|^2\right)^{\frac{1}{2}} \\
&\leqslant (\beta_{2p}^2 + 1)\left(\sum_{i=1}^n \left\|[\varphi]_{s_i} - [\varphi]_{s_{i-1}}\right\|_{L_p}^2 |s_i - s_{i-1}|^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Using $\|[\varphi]_{s_i} - [\varphi]_{s_{i-1}}\|_{L_p} \leqslant C_a|s_i - s_{i-1}|$, this implies

$$2\theta n \left\|\sum_{i=1}^n dM_i\right\|_{L_p} \leqslant 2\theta\beta_p(\beta_{2p}^2+1)C_a\sqrt{T}\left[n \sup_{i=1,\dots,n} |s_i - s_{i-1}|^{\frac{3}{2}}\right]$$

and

$$\lim_{n\to\infty} 2\theta n \left\|[\varphi;\tau_n^\theta]_a - \int_{(0,a]} (\overline{u}_{\tau_n^\theta} - u)\mathrm{d}[\varphi]_u\right\|_{L_p} = 0$$

with $\overline{u}_{\tau_n^\theta} := t_{i,n}^\theta$ for $u \in (t_{i-1,n}^\theta, t_{i,n}^\theta]$ (here we use the boundedness assumptions on $(K_t)_{t\in[0,T)}$ to replace $t_{i,n}^\theta \wedge a$ by $t_{i,n}^\theta$). At the same time we have

$$\lim_{n\to\infty} \left\| \frac{2\theta n}{T} \int_{(0,a]} (\overline{u}_{\tau_n^\theta} - u) \mathrm{d}[\varphi]_u - \int_{(0,a]} \left(\frac{T-u}{T} \right)^{1-\theta} \mathrm{d}[\varphi]_u \right\|_{L_p} = 0$$

which proves our statement. Regarding the last limit, we first observe that $\frac{2\theta n}{T} \int_{(0,a]} (\overline{u}_{\tau_n^\theta} - u) \mathrm{d}[\varphi]_u$ converges point-wise to $\int_{(0,a]} \left(\frac{T-u}{T}\right)^{1-\theta} \mathrm{d}[\varphi]_u$ as $\mathrm{d}[\varphi]_t(\omega) = K_t(\omega)\mathrm{d}t$ and $t \mapsto K_t(\omega)$ is continuous and because the measures $\mu_{n,\theta}$ on $\mathcal{B}([0,T))$ with $\mu_{n,\theta}(\mathrm{d}u) := \left[\frac{2\theta n}{T} \sum_{i=1}^n \mathbb{1}_{(t_{i-1,n}^\theta, t_{i,n}^\theta]}(u)(\overline{u}_{\tau_n^\theta} - u) \right] \mathrm{d}u$ converges weakly to $\mu_\theta(\mathrm{d}u) := \left(\frac{T-u}{T}\right)^{1-\theta} \mathrm{d}u$ on each interval $[0,a] \subset [0,T)$ (one has $\lim_n \mu_{n,\theta}([0,a])$ $= \mu_\theta([0,a])$ for $a \in [0,T)$ which follows from $\lim_n \mu_{n,\theta}([0,a]) = \lim_n \frac{\theta n}{T} \sum_{i\geqslant 1: t_{i,n}^\theta \leqslant a} (t_{i,n}^\theta - t_{i-1,n}^\theta)^2$ and $t_{i,n}^\theta - t_{i-1,n}^\theta = \frac{T^\theta}{n\theta}(T - \xi_{n,i}^\theta)^{1-\theta}$ for some $\xi_{n,i}^\theta \in [t_{i-1,n}^\theta, t_{i,n}^\theta]$). To apply dominated convergence in order to get the L_p-limit we use $n \int_{(0,a]} (\overline{u}_{\tau_n^\theta} - u) \mathrm{d}[\varphi]_u \leqslant a C_a n \|\tau_n^\theta\|_1 \leqslant a C_a T/\theta$ (see (2.9)). $\square$

4. Riemann-Liouville type operators and approximation

Various L_p-approximation problems in stochastic integration theory can be translated by the Burkholder-Davis-Gundy inequalities into problems about quadratic variation processes. In the special case of L_2-approximations this is particularly useful as there is a chance to turn the approximation problem into -in a sense- more deterministic problem by Fubini's theorem when the interchange of the integration in time and in ω is possible. When $p \neq 2$ this does not work (at least) in this straight way, see for example [24]. However, passing from global L_2-estimates to weighted local L_2-estimates, i.e. weighted bounded mean oscillation estimates, and exploiting a weighted John-Nirenberg type theorem, gives a natural approach to L_p- and exponential estimates.

Theorem 4.3 and Theorem 4.4 below are the key to exploit these local L_2-estimates in our article later. It turned out that one can naturally formulate these theorems in the general setting of random measures (Π, Υ). Later, the measure Π will describe the quadratic variation of the driving process of the stochastic integral to be approximated and Υ will describe some kind of *curvature* of the stochastic integral. For this one needs a replacement of *orthogonality*. For us, this replacement is the relation given in (4.1) below.

So let us start by introducing the random measures and the quasi-orthogonality where we use extended conditional expectations for non-negative random variables.

Assumption 4.1. We assume random measures

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

a progressively measurable process $(\varphi_t)_{t\in[0,T)}$, and a constant $\kappa \geqslant 1$, such that

$$\Pi(\omega, (0,b]) + \Upsilon(\omega, (0,b]) + \sup_{t\in[0,b]} |\varphi_t(\omega)| < \infty$$

for $(\omega, b) \in \Omega \times (0,T)$ and such that, for $0 \leqslant s \leqslant a < b < T$,

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |\varphi_u - \varphi_s|^2 \,\Pi(\cdot, \mathrm{d}u) \right] \sim_\kappa \mathbb{E}^{\mathcal{F}_a}\left[|\varphi_a - \varphi_s|^2 \,\Pi(\cdot, (a,b]) + \int_{(a,b]} (b-u) \Upsilon(\cdot, \mathrm{d}u) \right] \text{ a.s.} \quad (4.1)$$

When (4.1) holds with $\preceq_\kappa$, then we denote the inequality by $(4.1)^\leqslant$, in case of $\succeq_\kappa$, by $(4.1)^\geqslant$.

To simplify the notation in some situations we extend Π and Υ to $\Pi, \Upsilon \colon \Omega \times \mathcal{B}((0,T]) \to [0,\infty]$ by $\Pi(\omega, \{T\}) = \Upsilon(\omega, \{T\}) = 0$ for all $\omega \in \Omega$.

Definition 4.2. Under the Assumption 4.1 we define for $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$ the non-negative, non-decreasing, and càdlàg process $[\varphi;\tau]^\pi = ([\varphi;\tau]_a^\pi)_{a\in[0,T)}$ by

$$[\varphi;\tau]_a^\pi := \int_{(0,a]} \left| \varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u) \right|^2 \Pi(\cdot, \mathrm{d}u) \in [0,\infty)$$

and let $[\varphi;\tau]_T^\pi := \lim_{a\uparrow T} [\varphi;\tau]_a^\pi \in [0,\infty]$.

The next two statements, Theorems 4.3 and 4.4, develop further ideas from [22, Lemma 3.8] and [24, Lemma 5.6] to a general conditional setting using random measures we exploit in the sequel. For $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$ and $a \in [t_{k-1}, t_k)$ we let

$$\underline{a}(\tau) := t_{k-1} \quad \text{and} \quad \overline{a}(\tau) := t_k.$$

Theorem 4.3 (Upper bounds)**.** *Suppose Assumption 4.1 with (4.1)$^{\leqslant}$. If $(\theta, a) \in (0,1] \times [0,T)$, $\tau \in \mathcal{T}$, and $(\underline{a}, \overline{a}] := (\underline{a}(\tau), \overline{a}(\tau)]$, then*

$$\frac{\mathbb{E}^{\mathcal{F}_a}[[\varphi;\tau]^\pi_T - [\varphi;\tau]^\pi_a]}{\|\tau\|_\theta} \leqslant \kappa \begin{cases} \mathbb{E}^{\mathcal{F}_a}\Big[\int_{(a,T)}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u) + \frac{(T-\underline{a})^{1-\theta}}{\overline{a}-\underline{a}}|\varphi_a - \varphi_{\underline{a}}|^2\Pi(\cdot,(a,\overline{a}])\Big] & \\ \mathbb{E}^{\mathcal{F}_a}\Big[\int_{(a,T)}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\Big] \qquad \textit{if} \quad a \in \tau \end{cases} \quad a.s.$$

Theorem 4.4 (Lower bounds)**.** *Suppose Assumption 4.1 with (4.1)$^{\geqslant}$ and $(\theta, a) \in (0,1] \times [0,T)$.*

(1) *If $\tau \in \mathcal{T}$, $(\underline{a}, \overline{a}] := (\underline{a}(\tau), \overline{a}(\tau)]$, and $\|\tau\|_\theta = \frac{\overline{a}-\underline{a}}{(T-\underline{a})^{1-\theta}}$, then*

$$\frac{\mathbb{E}^{\mathcal{F}_a}[[\varphi;\tau]^\pi_{\overline{a}} - [\varphi;\tau]^\pi_a]}{\|\tau\|_\theta} \geqslant \frac{1}{\kappa}\mathbb{E}^{\mathcal{F}_a}\left[\frac{(T-\underline{a})^{1-\theta}}{\overline{a}-\underline{a}}|\varphi_a - \varphi_{\underline{a}}|^2\Pi(\cdot,(a,\overline{a}])\right] \quad a.s.$$

(2) *There exist $\tau_n \in \mathcal{T}$, $n \in \mathbb{N}$, with $a \in \tau_n$ and $\lim_n \|\tau_n\|_\theta = 0$ such that*

$$\liminf_n \frac{\mathbb{E}^{\mathcal{F}_a}[[\varphi;\tau_n]^\pi_T - [\varphi;\tau_n]^\pi_a]}{\|\tau_n\|_\theta} \geqslant \frac{1}{\kappa 2^{\frac{1}{\theta}+2}}\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\right] \quad a.s.$$

Proof of Theorem 4.3. To simplify the notation we set $\varphi_T := 0$. It is obvious that we only need to show the first inequality. For $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$, $(t_{k-1}, t_k] = (\underline{a}(\tau), \overline{a}(\tau)]$, and $s_i := t_i \vee a$ one has, a.s.,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T]}\left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \Pi(\cdot,\mathrm{d}u)\right] \\
&= \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t_k]}\left|\varphi_u - \varphi_{t_{k-1}}\right|^2\Pi(\cdot,\mathrm{d}u) + \sum_{i=k+1}^n \int_{(t_{i-1},t_i]}\left|\varphi_u - \varphi_{t_{i-1}}\right|^2\Pi(\cdot,\mathrm{d}u)\right] \\
&\leqslant \kappa\mathbb{E}^{\mathcal{F}_a}\left[\left|\varphi_a - \varphi_{t_{k-1}}\right|^2\Pi(\cdot,(a,t_k]) + \sum_{i=k}^n \int_{(s_{i-1},s_i]}(s_i-u)\Upsilon(\cdot,\mathrm{d}u)\right] \\
&\leqslant \kappa\mathbb{E}^{\mathcal{F}_a}\left[\frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}\left|\varphi_a - \varphi_{t_{k-1}}\right|^2 \frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}}\Pi(\cdot,(a,t_k])\right. \\
&\qquad\qquad \left. + \sum_{i=k}^n \int_{(s_{i-1},s_i]}\frac{s_i - u}{(T-u)^{1-\theta}}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\right] \\
&\leqslant \kappa\|\tau\|_\theta\mathbb{E}^{\mathcal{F}_a}\left[\left|\varphi_a - \varphi_{t_{k-1}}\right|^2\frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}}\Pi(\cdot,(a,t_k]) + \int_{(a,T)}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\right]
\end{aligned}$$

where we use (2.10). □

Proof of Theorem 4.4. (1) Beginning the proof as for Theorem 4.3 with $(t_{k-1}, t_k] = (\underline{a}(\tau), \overline{a}(\tau)]$, we get, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t_k]}\left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2\Pi(\cdot,\mathrm{d}u)\right] &= \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t_k]}\left|\varphi_u - \varphi_{t_{k-1}}\right|^2\Pi(\cdot,\mathrm{d}u)\right] \\
&\geqslant \frac{1}{\kappa}\mathbb{E}^{\mathcal{F}_a}\left[\left|\varphi_a - \varphi_{t_{k-1}}\right|^2\Pi(\cdot,(a,t_k])\right].
\end{aligned}$$

Dividing by $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$ we obtain the desired statement.

(2) We partition the interval $[a,T]$ with

$$u_{i,n}^{\theta,a} := a + (T-a)\left[1 - \left(1 - \tfrac{i}{n}\right)^{\frac{1}{\theta}}\right], \quad i = 0, \dots, n,$$

$$r_{i,n}^{\theta,a} := a + (T-a)\left[1 - \left(1 - \tfrac{2i-1}{2n}\right)^{\frac{1}{\theta}}\right], \quad i = 1, \ldots, n,$$

and add $r_{0,n}^{\theta,a} := a$ and $r_{n+1,n}^{\theta,a} := T$. Choosing for both nets the remaining time-knots on $[0,a]$ fine enough, we obtain nets $\tau_n^{\theta,a}$ and $\widetilde{\tau}_n^{\theta,a}$ satisfying

$$\|\tau_n^{\theta,a}\|_\theta = \sup_{i=1,\ldots,n} \frac{u_{i,n}^{\theta,a} - u_{i-1,n}^{\theta,a}}{(T - u_{i-1,n}^{\theta,a})^{1-\theta}} \quad \text{and} \quad \|\widetilde{\tau}_n^{\theta,a}\|_\theta = \sup_{i=0,1,\ldots,n} \frac{r_{i+1,n}^{\theta,a} - r_{i,n}^{\theta,a}}{(T - r_{i,n}^{\theta,a})^{1-\theta}}.$$

By a computation, we have for $i = 1, \ldots, n$ and $u \in (u_{i-1,n}^{\theta,a}, r_{i,n}^{\theta,a}]$ that

$$\frac{(T-a)^\theta}{\theta 2^{\frac{1}{\theta}+1} n} \leqslant \frac{u_{i,n}^{\theta,a} - r_{i,n}^{\theta,a}}{(T - r_{i,n}^{\theta,a})^{1-\theta}} \leqslant \frac{u_{i,n}^{\theta,a} - u}{(T-u)^{1-\theta}} \leqslant \frac{u_{i,n}^{\theta,a} - u_{i-1,n}^{\theta,a}}{(T - u_{i-1,n}^{\theta,a})^{1-\theta}} \leqslant \frac{(T-a)^\theta}{\theta n}, \tag{4.2}$$

and for $i = 1, \ldots, n-1$ and $u \in (r_{i,n}^{\theta,a}, u_{i,n}^{\theta,a}]$ that

$$\frac{(T-a)^\theta}{\theta 2^{\frac{1}{\theta}+1} n} \leqslant \frac{r_{i+1,n}^{\theta,a} - u_{i,n}^{\theta,a}}{(T - u_{i,n}^{\theta,a})^{1-\theta}} \leqslant \frac{r_{i+1,n}^{\theta,a} - u}{(T-u)^{1-\theta}} \leqslant \frac{r_{i+1,n}^{\theta,a} - r_{i,n}^{\theta,a}}{(T - r_{i,n}^{\theta,a})^{1-\theta}} \leqslant \frac{(T-a)^\theta}{\theta n}, \tag{4.3}$$

where the last inequality holds for $i \in \{0, n\}$ as well. By the above relations we obtain, a.s.,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a, r_{n,n}^{\theta,a}]} (T-u)^{1-\theta} \Upsilon(\cdot, \mathrm{d}u)\right] \\
&\qquad = \sum_{i=1}^n \mathbb{E}^{\mathcal{F}_a}\left[\int_{(u_{i-1,n}^{\theta,a}, r_{i,n}^{\theta,a}]} (T-u)^{1-\theta} \Upsilon(\cdot, \mathrm{d}u)\right] \\
&\qquad\quad + \sum_{i=1}^{n-1} \mathbb{E}^{\mathcal{F}_a}\left[\int_{(r_{i,n}^{\theta,a}, u_{i,n}^{\theta,a}]} (T-u)^{1-\theta} \Upsilon(\cdot, \mathrm{d}u)\right] \\
&\qquad \leqslant \frac{\theta 2^{\frac{1}{\theta}+1} n}{(T-a)^\theta} \left[\sum_{i=1}^n \mathbb{E}^{\mathcal{F}_a}\left[\int_{(u_{i-1,n}^{\theta,a}, r_{i,n}^{\theta,a}]} (u_{i,n}^{\theta,a} - u) \Upsilon(\cdot, \mathrm{d}u)\right]\right. \\
&\qquad\qquad\qquad\qquad \left. + \sum_{i=1}^{n-1} \mathbb{E}^{\mathcal{F}_a}\left[\int_{(r_{i,n}^{\theta,a}, u_{i,n}^{\theta,a}]} (r_{i+1,n}^{\theta,a} - u) \Upsilon(\cdot, \mathrm{d}u)\right]\right] \\
&\qquad \leqslant (\kappa 2^{\frac{1}{\theta}+1}) \mathbb{E}^{\mathcal{F}_a}\left[\frac{[\varphi; \tau_n^{\theta,a}]_T^\pi - [\varphi; \tau_n^{\theta,a}]_a^\pi}{\|\tau_n^{\theta,a}\|_\theta} + \frac{[\varphi; \widetilde{\tau}_n^{\theta,a}]_T^\pi - [\varphi; \widetilde{\tau}_n^{\theta,a}]_a^\pi}{\|\widetilde{\tau}_n^{\theta,a}\|_\theta}\right],
\end{aligned}$$

where for the last inequality we first use (4.1), that gives the factor κ, and then (4.2) and (4.3) that give $\|\tau_n^{a,\theta}\|_\theta \leqslant ((T-a)^\theta)/(\theta n)$ and $\|\widetilde{\tau}_n^{a,\theta}\|_\theta \leqslant ((T-a)^\theta)/(\theta n)$. For each n we choose the time-net that gives the larger quotient and obtain the desired nets. To obtain the final statement we observe that $r_{n,n}^{\theta,a} \uparrow T$. □

Now we specialize Assumption 4.1 to the settings that will be used in Sections 6 and 8:

Assumption 4.5. We assume that there are

(1) a positive continuous and adapted process $(\sigma_t)_{t\in[0,T]}$ such that $\sigma_T^* \in L_2$ and such that there is a $c_\sigma \geqslant 1$ with

$$\mathbb{E}^{\mathcal{F}_a}\left[\frac{1}{b-a}\int_a^b \sigma_u^2 \mathrm{d}u\right] \sim_{c_\sigma} \sigma_a^2 \quad \text{a.s. for all} \quad 0 \leqslant a < b \leqslant T,$$

(2) a square integrable martingale $M = (M_t)_{t\in[0,T)}$ with $M_0 \equiv 0$,
(3) a $\varphi \in \mathrm{CL}([0,T))$ with $\mathbb{E}\sup_{u\in[a,T]} |\varphi_a \sigma_u|^2 < \infty$ for all $a \in [0,T)$,
(4) let $\Pi(\omega, \mathrm{d}u) := \sigma_u^2(\omega)\mathrm{d}u$ and $\Upsilon(\omega, \mathrm{d}u) := \mathrm{d}\langle M\rangle_u(\omega)$ for $u \in [0,T)$, where $\langle M\rangle$ is the conditional square-function (see Section 2.5),
(5) assume that (4.1) is satisfied, and
(6) let $[\varphi; \tau]^\sigma := [\varphi; \tau]^\pi$.

Remark 4.6. Assumption 4.5, the equality

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |M_u - M_a|^2 \mathrm{d}u\right] = \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} (b-u)\mathrm{d}\langle M\rangle_u\right] \text{ a.s.}$$

for $0 \leqslant a < b < T$ yield, for $0 \leqslant s \leqslant a < b < T$ and with $\kappa' := \kappa c_\sigma$, to

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |\varphi_u - \varphi_s|^2 \sigma_u^2 \mathrm{d}u\right] \sim_{\kappa'} (b-a)\,|\varphi_a - \varphi_s|^2 \sigma_a^2 + \mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |M_u - M_a|^2 \mathrm{d}u\right] \text{ a.s.} \tag{4.4}$$

From Theorem 4.3 and Theorem 4.4 we immediately deduce:

Theorem 4.7. *Assume Assumption 4.5, $(\theta, a) \in (0,1] \times [0,T)$, and a σ-algebra $\mathcal{G} \subseteq \mathcal{F}_a$. Then there are constants $c_{(4.5)}, c_{(4.6)} \geqslant 1$ depending at most on $(\theta, \kappa, c_\sigma)$ such that one has, a.s.,*

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{T},\tau\ni a} \frac{\mathbb{E}^{\mathcal{G}}[[\varphi;\tau]_T^\sigma - [\varphi;\tau]_a^\sigma]}{\|\tau\|_\theta} \sim_{c_{(4.5)}} \mathbb{E}^{\mathcal{G}}\Big[\sup_{t\in[a,T)} |\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M|^2\Big], \tag{4.5}$$

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{T}} \frac{\mathbb{E}^{\mathcal{F}_a}[[\varphi;\tau]_T^\sigma - [\varphi;\tau]_a^\sigma]}{\|\tau\|_\theta} \sim_{c_{(4.6)}} \mathbb{E}^{\mathcal{F}_a}\Big[\sup_{t\in[a,T)} |\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M|^2\Big] + \sup_{s\in[0,a]} \frac{T-a}{(T-s)^\theta} |\varphi_a - \varphi_s|^2 \sigma_a^2. \tag{4.6}$$

In order to prove Theorem 3.6 the inequality (4.5) is formulated for a more general σ-algebra $\mathcal{G}$. In (4.6) such a formulation is not necessary for us.

Proof of Theorem 4.7. Relation (4.6): For $0 \leqslant \underline{a} \leqslant a < \overline{a} \leqslant T$ Assumption 4.5 implies that

$$\mathbb{E}^{\mathcal{F}_a}\left[\frac{(T-\underline{a})^{1-\theta}}{\overline{a}-\underline{a}} |\varphi_a - \varphi_{\underline{a}}|^2 \Pi(\cdot,(a,\overline{a}])\right] \sim_{c_\sigma} |\varphi_a - \varphi_{\underline{a}}|^2 \frac{(T-\underline{a})^{1-\theta}}{\overline{a}-\underline{a}} \sigma_a^2 (\overline{a}-a) \text{ a.s.}$$

Maximizing the right-hand side over $\overline{a}$ gives $\frac{T-a}{(T-\underline{a})^\theta} |\varphi_a - \varphi_{\underline{a}}|^2 \sigma_a^2$ a.s. Moreover, by Proposition 3.8, equation (3.3), we have, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\left[\left|\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M\right|^2\right] = \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} \left(\frac{T-u}{T}\right)^{1-\theta} \mathrm{d}[M]_u\right] = \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} \left(\frac{T-u}{T}\right)^{1-\theta} \mathrm{d}\langle M\rangle_u\right]$$

for $0 \leqslant a < t < T$ so that

$$\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T)} \left|\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M\right|^2\right] \sim_4 \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} \left(\frac{T-u}{T}\right)^{1-\theta} \mathrm{d}\langle M\rangle_u\right]$$

by Doob's maximal inequality. Now we use Theorem 4.3 and Theorem 4.4.

Relation (4.5) for $\mathcal{G} = \mathcal{F}_a$ follows again from Theorem 4.3 and Theorem 4.4. In the case of $\mathcal{G} \subsetneq \mathcal{F}_a$ we argue as follows: let $c_{(4.5)} \geqslant 1$ be the constant in (4.5) for $\mathcal{F}_a$, then we get

$$\frac{\mathbb{E}^{\mathcal{G}}[[\varphi;\tau]_T^\sigma - [\varphi;\tau]_a^\sigma]}{\|\tau\|_\theta} \leqslant c_{(4.5)} \mathbb{E}^{\mathcal{G}}\Big[\sup_{t\in[a,T)} |\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M|^2\Big]$$

as well for all τ with $a \in \tau$ which implies the general inequality $\leqslant$ in (4.5). Regarding the remaining inequality we choose the time-nets from Theorem 4.4(2) to get by Fatou's lemma that, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{G}}\Big[\sup_{t\in[a,T)} |\mathcal{I}_t^{\frac{1-\theta}{2}} M - \mathcal{I}_a^{\frac{1-\theta}{2}} M|^2\Big] &\leqslant \kappa 2^{\frac{1}{\theta}+2} \mathbb{E}^{\mathcal{G}}\left[\liminf_n \mathbb{E}^{\mathcal{F}_a}\left[\frac{[\varphi;\tau_n]_T^\sigma - [\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right]\right] \\
&\leqslant \kappa 2^{\frac{1}{\theta}+2} \liminf_n \mathbb{E}^{\mathcal{G}}\left[\mathbb{E}^{\mathcal{F}_a}\left[\frac{[\varphi;\tau_n]_T^\sigma - [\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right]\right] \\
&= \kappa 2^{\frac{1}{\theta}+2} \liminf_n \mathbb{E}^{\mathcal{G}}\left[\frac{[\varphi;\tau_n]_T^\sigma - [\varphi;\tau_n]_a^\sigma}{\|\tau_n\|_\theta}\right].
\end{aligned}$$

□

The next theorem gives a complete characterization of $\|[\varphi;\tau]^\sigma\|_{\mathrm{BMO}_1^{\Phi^2}([0,T))} \leqslant c^2 \|\tau\|_\theta$:

Theorem 4.8. *Assume that Assumption 4.5 is satisfied. Then for $\theta \in (0,1]$ and $\Phi \in \mathrm{CL}^+([0,T))$ the following assertions are equivalent:*

(1) *One has $\mathcal{I}^{\frac{1-\theta}{2}} M \in \mathrm{bmo}_2^{\Phi}([0,T))$ and there is a $c_{(4.7)} > 0$ such that one has*

$$|\varphi_a - \varphi_s| \leqslant c_{(4.7)} \frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}} \frac{\Phi_a}{\sigma_a} \quad \text{for} \quad 0 \leqslant s < a < T \text{ a.s.} \tag{4.7}$$

(2) *There is a constant $c_{(4.8)} > 0$ such that, for all time-nets $\tau \in \mathcal{T}$,*

$$\|[\varphi;\tau]^{\sigma}\|_{\mathrm{BMO}_1^{\Phi^2}([0,T))} \leqslant c_{(4.8)}^2 \|\tau\|_{\theta}. \tag{4.8}$$

If $\Phi = (\sigma_t \Psi_t)_{t\in[0,T)}$, where $\Psi \in \mathrm{CL}^+([0,T))$ is non-decreasing, then (4.7) is equivalent to the existence of $c_{(4.9)}, c_{(4.10)} > 0$ such that

$$|\varphi_a - \varphi_0| \leqslant c_{(4.9)} (T-a)^{\frac{\theta-1}{2}} \Psi_a \qquad \text{for} \quad 0 \leqslant a < T \text{ a.s.} \qquad \text{if} \quad \theta \in (0,1), \tag{4.9}$$

$$|\varphi_a - \varphi_s| \leqslant c_{(4.10)} \left(1 + \log \frac{T-s}{T-a}\right) \Psi_a \quad \text{for} \quad 0 \leqslant s < a < T \text{ a.s.} \quad \text{if} \quad \theta = 1. \tag{4.10}$$

Proof. The equivalence between (1) and (2) follows directly from the second equivalence in Theorem 4.7 and Proposition A.4. The equivalence between (4.7) and (4.9)-(4.10) follows from Lemma C.1 below. □

5. Oscillation of stochastic processes and lower bounds

In this section we consider lower bounds for the oscillation of stochastic processes and use them in Section 6 (Case (C1)) and Section 8. As such, the approach is intended for stochastic processes $(\varphi_t)_{t\in[0,T)} \subseteq L_\infty$ with a blow-up of $\|\varphi_t\|_{L_\infty}$ if $t \uparrow T$. This is a typical case for the gradient processes we consider. The quantities, we are interested in, concern the degree of the oscillation of the process measured in L_∞, here denoted by $\underline{\mathrm{Osc}}_t(\varphi)$ and $\overline{\mathrm{Osc}}_t(\varphi)$. In order to get lower bounds for these oscillatory quantities, we use the concept of *maximal oscillation*. The above mentioned concepts are introduced in Definition 5.1 below. The maximal oscillation is verified in Example 5.5 and Example 5.6 below. The application to $[\varphi;\tau]$ is given in Theorem 5.7. Example 5.5 and Theorem 5.7 will be used in Section 6, and Example 5.6 and Theorem 5.7 will be used in the Lévy case in Section 8. Let us start to introduce our concept:

Definition 5.1. If $\varphi = (\varphi_t)_{t\in[0,T)}$ is a stochastic process and $t \in (0,T)$, then we let

$$\underline{\mathrm{Osc}}_t(\varphi) := \inf_{s\in[0,t)} \|\varphi_t - \varphi_s\|_{L_\infty} \in [0,\infty] \quad \text{and} \quad \overline{\mathrm{Osc}}_t(\varphi) := \inf_{s\in[0,t)} \sup_{u\in[s,t]} \|\varphi_t - \varphi_u\|_{L_\infty} \in [0,\infty].$$

The process is called of *maximal oscillation* with constant $c \geqslant 1$ if for all $t \in (0,T)$ one has

$$\underline{\mathrm{Osc}}_t(\varphi) \geqslant \frac{1}{c} \|\varphi_t - \varphi_0\|_{L_\infty}.$$

If both sides equal infinity, then we use $c = 1$ (however, this case is not of relevance for us).

Lemma 5.2. *For a stochastic process $\varphi = (\varphi_t)_{t\in[0,T)}$ the following holds:*

(1) *One has $\underline{\mathrm{Osc}}_t(\varphi) \leqslant \overline{\mathrm{Osc}}_t(\varphi)$ for $t \in (0,T)$.*
(2) *One has $\overline{\mathrm{Osc}}_t(\varphi) \leqslant 2\underline{\mathrm{Osc}}_t(\varphi)$ for $t \in (0,T)$ if φ is a martingale.*
(3) *If $\varphi_a \equiv \mathbb{1}_{\mathbb{Q}\cap[0,T)}(a)$ for $a \in [0,T)$, then $0 = \underline{\mathrm{Osc}}_t(\varphi) < \overline{\mathrm{Osc}}_t(\varphi) = 1$ for all $t \in (0,T)$.*

Proof. (1) follows from the definition. (2) If φ is a martingale and $0 \leqslant s < t < T$, then we have

$$\sup_{u\in[s,t]} \|\varphi_t - \varphi_u\|_{L_\infty} \leqslant \|\varphi_t - \varphi_s\|_{L_\infty} + \sup_{u\in[s,t]} \|\varphi_u - \varphi_s\|_{L_\infty} \leqslant 2\|\varphi_t - \varphi_s\|_{L_\infty}.$$

Taking the infimum on both sides over $s \in [0,t)$ yields the assertion. Item (3) is obvious. □

Remark 5.3. In the sequel we do not need the following two statements, so that we state them without proof:

(1) It is possible to construct examples such that for a given $c \in [1,\infty)$ the constant c is optimal in the definition of maximal oscillation.

(2) Again by examples one can see that the constant 2 in Lemma 5.2(2) is optimal.

To verify a maximal oscillation we make use of the following observation:

Lemma 5.4. *Assume two random variables $A, B : \Omega \to \mathbb{R}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. Assume a probability measure $\mathbb{Q} \ll \mathbb{P}$ such that $\mathbb{E}^{\mathbb{Q}}|B| < \infty$ and $\mathbb{E}^{\mathbb{Q}} B = 0$. Then*

$$\|B - A\|_{L_\infty(\mathbb{P})} \geqslant \inf_{a\in\mathbb{R}} \|B - a\|_{L_\infty(\mathbb{P})} \quad \textit{implies} \quad \|B - A\|_{L_\infty(\mathbb{P})} \geqslant \frac{1}{2}\|B\|_{L_\infty(\mathbb{P})}.$$

Proof. We may assume that $\|B - A\|_{L_\infty(\mathbb{P})} < \infty$, otherwise there is nothing to prove. Because of our assumption, for all $\varepsilon > 0$ there is an $a_\varepsilon \in \mathbb{R}$ such that we have

$$\|B-A\|_{L_\infty(\mathbb{P})} \geqslant \|B\|_{L_\infty(\mathbb{P})} - |a_\varepsilon| - \varepsilon \quad \text{and} \quad \|B-A\|_{L_\infty(\mathbb{P})} \geqslant \mathbb{E}^{\mathbb{Q}}|B-a_\varepsilon| - \varepsilon \geqslant |\mathbb{E}^{\mathbb{Q}} B - a_\varepsilon| - \varepsilon = |a_\varepsilon| - \varepsilon.$$

The combination of the inequalities implies

$$\|B - A\|_{L_\infty(\mathbb{P})} \geqslant \|B\|_{L_\infty(\mathbb{P})} - |a_\varepsilon| - \varepsilon \geqslant \|B\|_{L_\infty(\mathbb{P})} - \|B - A\|_{L_\infty(\mathbb{P})} - 2\varepsilon$$

so that $2\|B - A\|_{L_\infty(\mathbb{P})} \geqslant \|B\|_{L_\infty(\mathbb{P})} - 2\varepsilon$. By $\varepsilon \downarrow 0$ we get our statement. □

Now we consider two examples relevant for us:

Example 5.5 (Markov type processes, Section 6)**.** Let $(Y_t)_{t\in[0,T]}$ be a process with values in $\mathcal{R}_Y$, where $\mathcal{R}_Y = \mathbb{R}$ or $\mathcal{R}_Y = (0, \infty)$, and $Y_0 \equiv y_0 \in \mathcal{R}_Y$. Assume continuous transition densities $\Gamma_Y : \{(s,t) : 0 \leqslant s < t \leqslant T\} \times \mathcal{R}_Y \times \mathcal{R}_Y \to (0, \infty)$ such that

$$\mathbb{P}(Y_t \in B \,|\, Y_s) = \int_B \Gamma_Y(s, t; Y_s, y)\mathrm{d}y \quad \text{a.s.} \tag{5.1}$$

for $B \in \mathcal{B}(\mathcal{R}_Y)$ and $0 \leqslant s < t \leqslant T$. Then, for $0 < s < t \leqslant T$ and continuous $H, \tilde{H} : \mathcal{R}_Y \to \mathbb{R}$, one has

$$\|H(Y_t) - \tilde{H}(Y_s)\|_{L_\infty} \geqslant \|H(Y_t) - \tilde{H}(y_0)\|_{L_\infty}.$$

This follows from the fact that the density $D_{s,t} : \mathcal{R}_Y \times \mathcal{R}_Y \to [0, \infty)$ of $law(Y_s, Y_t)$ with respect to the Lebesgue measure $\lambda \otimes \lambda|_{\mathcal{R}_Y\times\mathcal{R}_Y}$ is the positive and continuous function

$$D_{s,t}(y_1, y_2) := \Gamma_Y(0, s; y_0, y_1)\Gamma_Y(s, t; y_1, y_2).$$

Consequently, if there is a probability measure $\mathbb{Q} \ll \mathbb{P}$ and if for all $t \in [0, T)$ one has that $H(t, \cdot) : \mathcal{R}_Y \to \mathbb{R}$ is continuous, $\mathbb{E}^{\mathbb{Q}}|H(t, Y_t)| < \infty$, and $\mathbb{E}^{\mathbb{Q}}(H(t, Y_t) - H(0, y_0)) = 0$, then $(H(t, Y_t) - H(0, y_0))_{t\in[0,T)}$ is of maximal oscillation with constant 2 according to Lemma 5.4.

Example 5.6 (Lévy processes, Section 8)**.** Let $(X_t)_{t\in[0,T]}$, $X_t : \Omega \to \mathbb{R}$, be a Lévy process. By [38, Theorem 61.2] there are $\ell \in \mathbb{R}$ and a closed non-empty $Q \subseteq \mathbb{R}$ such that $0 \in Q$, $Q + Q = Q$, and $\mathrm{supp}(X_t) = Q + \ell t$ for $t \in (0, T]$. Define

$$Y_t := (X_t - \ell t)\mathbb{1}_{\{X_t\in \mathrm{supp}(X_t)\}}$$

so that $Y_t(\Omega) \subseteq Q$ and $\mathrm{supp}(Y_t) = Q$ for all $t \in (0, T]$. Let $0 < s < t \leqslant T$ and $H, \tilde{H} : Q \to \mathbb{R}$ be continuous on Q . Then

$$\|H(Y_t) - \tilde{H}(Y_s)\|_{L_\infty} \geqslant \|H(Y_t) - \tilde{H}(0)\|_{L_\infty}.$$

This can be seen from

$$\begin{aligned} \|H(Y_t) - \tilde{H}(Y_s)\|_{L_\infty} = \|H(Y_s + (Y_t - Y_s)) - \tilde{H}(Y_s)\|_{L_\infty} &= \sup_{y,y'\in Q} |H(y' + y) - \tilde{H}(y')| \\ &\geqslant \sup_{y\in Q} |H(y) - \tilde{H}(0)| = \|H(Y_t) - \tilde{H}(0)\|_{L_\infty}. \end{aligned}$$

Consequently, if there is a probability measure $\mathbb{Q} \ll \mathbb{P}$ and if for all $t \in [0, T)$ one has that $H(t, \cdot) : Q \to \mathbb{R}$ is continuous, $\mathbb{E}^{\mathbb{Q}}|H(t, Y_t)| < \infty$, and $\mathbb{E}^{\mathbb{Q}}(H(t, Y_t) - H(0, 0)) = 0$, then $(H(t, Y_t) - H(0, 0))_{t\in[0,T)}$ is of maximal oscillation with constant 2 according to Lemma 5.4.

Now we connect the notion of oscillation to the behavior of $[\varphi; \tau]$, where we use extended conditional expectations for non-negative random variables.

Theorem 5.7. *Assume $\theta \in (0,1]$, $c_{(5.2)} > 0$, and an adapted càdlàg process $(\varphi_t)_{t\in[0,T)}$ such that*

$$\frac{1}{c_{(5.2)}}|\varphi_a - Z|^2 \leqslant \mathbb{E}^{\mathcal{F}_a}\left[\frac{1}{b-a}\int_a^b |\varphi_u - Z|^2 \mathrm{d}u\right] \quad a.s. \tag{5.2}$$

for all $0 \leqslant a < b < T$ and all $\mathcal{F}_a$-measurable $Z : \Omega \to \mathbb{R}$. Consider the following assertions:

(1) $\inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\underline{\mathrm{Osc}}_t(\varphi) > 0$.

(2) *There is a $c_{(5.3)} > 0$ such that for all $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$ with $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$ one has*

$$\inf_{\vartheta_{i-1}\in L_0(\mathcal{F}_{t_{i-1}})} \sup_{a\in[t_{k-1},t_k)} \left\|\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|_{L_\infty} \geqslant c^2_{(5.3)}\|\tau\|_\theta. \tag{5.3}$$

(3) *There is a constant $c_{(5.4)} > 0$ such that for all time-nets $\tau \in \mathcal{T}$ one has*

$$\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} \geqslant c^2_{(5.4)}\|\tau\|_\theta. \tag{5.4}$$

(4) $\inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\overline{\mathrm{Osc}}_t(\varphi) > 0$.

Then we have (1) ⇒ (2) ⇒ (3). If $\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} < \infty$ for all $\tau \in \mathcal{T}$ and $\|[\varphi;\tau]\|_{\mathrm{BMO}_1([0,T))} \to 0$ for $\|\tau\|_1 \to 0$, then (3) ⇒ (4).

Proof. (1) ⇒ (2) If $\delta := \inf_{t\in(0,T)}(T-t)^{\frac{1-\theta}{2}}\underline{\mathrm{Osc}}_t(\varphi) > 0$ and $a \in [t_{k-1}, t_k)$, then, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right] &\geqslant \mathbb{E}^{\mathcal{F}_a}\left[\int_a^{t_k} \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right] \\ &= \mathbb{E}^{\mathcal{F}_a}\left[\int_a^{t_k} |\varphi_u - \vartheta_{k-1}|^2 \mathrm{d}u\right] \\ &\geqslant \frac{1}{\kappa}(t_k - a)|\varphi_a - \vartheta_{k-1}|^2.\end{aligned}$$

We apply this inequality to $a = t_{k-1}$ and $a = a_0 := \frac{1}{2}(t_{k-1} + t_k)$, observe that

$$\begin{aligned}\frac{1}{2}\left[(t_k - t_{k-1})\|\varphi_{t_{k-1}} - \vartheta_{k-1}\|^2_{L_\infty} + (t_k - a_0)\|\varphi_{a_0} - \vartheta_{k-1}\|^2_{L_\infty}\right] &\geqslant \frac{t_k - a_0}{4}\|\varphi_{a_0} - \varphi_{t_{k-1}}\|^2_{L_\infty} \\ &\geqslant \frac{t_k - t_{k-1}}{8}\underline{\mathrm{Osc}}^2_{a_0}(\varphi),\end{aligned}$$

and deduce

$$\sup_{a\in[t_{k-1},t_k)} \left\|\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|^2_{L_\infty} \geqslant \frac{1}{\kappa}\frac{\delta^2}{8}\frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}} = \frac{1}{\kappa}\frac{\delta^2}{8}\|\tau\|_\theta.$$

(2) ⇒ (3) with $c_{(5.4)} := c_{(5.3)}$ is obvious because we can choose $\vartheta_{i-1} := \varphi_{t_{i-1}}$.

(3) ⇒ (4) For $a \in [0,T)$ and $0 \leqslant s < t < T$, a time-net $\tau = \{t_i\}_{i=0}^n$ such that $s = t_{k-1} < t_k = t$ and

$$\frac{t-s}{(T-s)^{1-\theta}} = \|\tau\|_\theta \tag{5.5}$$

we get

$$\begin{aligned}\left\|\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)\cap(s,t]} \left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|_{L_\infty} &= \left\|\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)\cap(s,t]} |\varphi_u - \varphi_s|^2 \mathrm{d}u\right]\right\|_{L_\infty} \\ &\leqslant (t-s)\sup_{u\in(s,t]}\|\varphi_u - \varphi_s\|^2_{L_\infty}\end{aligned}$$

and

$$c_{(5.4)}\sqrt{\frac{t-s}{(T-s)^{1-\theta}}} = c_{(5.4)}\|\tau\|_\theta^{\frac{1}{2}} \leqslant \sup_{a\in[0,T)}\left\|\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|^{\frac{1}{2}}_{L_\infty}$$

$$
\begin{aligned}
&\leqslant \sup_{a\in[0,T)} \left[\left\| \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)\cap(s,t]} \left| \varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u) \right|^2 \mathrm{d}u \right] \right\|_{L_\infty}^{\frac12} \right.\\
&\qquad \left. + \left\| \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)\setminus(s,t]} \left| \varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u) \right|^2 \mathrm{d}u \right] \right\|_{L_\infty}^{\frac12} \right]\\
&\leqslant \sqrt{t-s} \sup_{u\in(s,t]} \|\varphi_u - \varphi_s\|_{L_\infty}\\
&\qquad + \sup_{a\in[0,T)} \left\| \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)\setminus(s,t]} \left| \varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u) \right|^2 \mathrm{d}u \right] \right\|_{L_\infty}^{\frac12}.
\end{aligned}
$$

Assume a time-net $\tilde\tau$ that coincides with τ outside the interval (s,t). Then

$$c_{(5.4)} \sqrt{\frac{t-s}{(T-s)^{1-\theta}}} \leqslant \sqrt{t-s} \sup_{u\in(s,t]} \|\varphi_u - \varphi_s\|_{L_\infty} + \|[\varphi;\tilde\tau]\|_{\mathrm{BMO}_1([0,T))}^{\frac12}.$$

Choosing a sequence $(\tau_n, \tilde\tau_n)$ of $(\tau,\tilde\tau)$ with (5.5) as above, such that $\|\tilde\tau_n\|_1 \to 0$, we conclude with

$$c_{(5.4)} \sqrt{\frac{t-s}{(T-s)^{1-\theta}}} \leqslant \sqrt{t-s} \sup_{u\in(s,t]} \|\varphi_u - \varphi_s\|_{L_\infty} \quad \text{and} \quad \sup_{u\in(s,t]} \|\varphi_u - \varphi_s\|_{L_\infty}^2 \geqslant c_{(5.4)}^2 (T-s)^{\theta-1}.$$

For $s \in [(2t-T)^+, t)$ this gives $\sup_{u\in(s,t]} \|\varphi_u - \varphi_s\|_{L_\infty}^2 \geqslant c_{(5.4)}^2 2^{\theta-1}(T-t)^{\theta-1}$ and therefore

$$c_{(5.4)} 2^{\frac{\theta-1}{2}} (T-t)^{\frac{\theta-1}{2}} \leqslant \|\varphi_t - \varphi_s\|_{L_\infty} + \sup_{u\in(s,t]} \|\varphi_t - \varphi_u\|_{L_\infty} \leqslant 2 \sup_{u\in[s,t]} \|\varphi_t - \varphi_u\|_{L_\infty}.$$

This implies $\overline{\mathrm{Osc}}_t(\varphi) \geqslant c_{(5.4)} 2^{\frac{\theta-3}{2}} (T-t)^{\frac{\theta-1}{2}}$. □

6. Brownian setting: Gradient estimates and approximation

We suppose additionally that $\mathcal{F} = \mathcal{F}_T$ and that $(\mathcal{F}_t)_{t\in[0,T]}$ is the augmentation of the natural filtration of a standard one-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ with continuous paths and starting in zero for all $\omega \in \Omega$. We recall the setting from [16] and start with the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \hat\sigma(X_t)\mathrm{d}W_t + \hat b(X_t)\mathrm{d}t \quad \text{with} \quad X_0 \equiv x_0 \in \mathbb{R} \tag{6.1}$$

where $0 < \varepsilon_0 \leqslant \hat\sigma \in C_b^\infty(\mathbb{R})$ for some constant ε_0 and $\hat b \in C_b^\infty(\mathbb{R})$ and where all paths of X are assumed to be continuous. From this equation we derive the SDE

$$\mathrm{d}Y_t = \sigma(Y_t)\mathrm{d}W_t \quad \text{with} \quad Y_0 \equiv y_0 \in \mathbb{R}$$

where two settings are used simultaneously:

Case (C1): $Y := X$ with $\sigma \equiv \hat\sigma$, $\hat b \equiv 0$, and $\mathcal{R}_Y := \mathbb{R}$.
Case (C2): $Y := \mathrm{e}^X$ with $\sigma(y) := y\hat\sigma(\log y)$, $\hat b(x) := -\frac12 \hat\sigma^2(x)$, and $\mathcal{R}_Y := (0,\infty)$.

In both cases, we let C_Y be the set of all Borel functions $g\colon \mathcal{R}_Y \to \mathbb{R}$ such that

$$\sup_{x\in\mathbb{R}} \mathrm{e}^{-m|x|} \int_{\mathbb{R}} |g(\alpha(x+ty))|^2 \, \mathrm{e}^{-y^2} \, \mathrm{d}y < \infty \quad \text{for all} \quad t > 0$$

for some $m > 0$, where $\alpha(x) = x$ in the case (C1) and $\alpha(x) = \mathrm{e}^x$ in the case (C2). Let us denote by $(Y_s^{t,y})_{s\in[t,T]}$ be the diffusion Y started at time $t \in [0,T]$ in $y \in \mathcal{R}_Y$ and let us define, for $g \in C_Y$,

$$G(t,y) := \mathbb{E} g(Y_T^{t,y}) \quad \text{for} \quad (t,y) \in [0,T] \times \mathcal{R}_Y.$$

Remark 6.1. We collect some facts we shall use and that hold in both cases, (C1) and (C2):

(A) $\|\sigma'\|_{B_b(\mathcal{R}_Y)} + \|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)} < \infty$.
(B) In the case (C2) we have $\sigma(y) \sim_c y$ for $y \in \mathcal{R}_Y$ and some $c \geqslant 1$.
(C) One has $G \in C^\infty([0,T) \times \mathcal{R}_Y)$ and $\frac{\partial G}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 G}{\partial y^2} = 0$ on $[0,T) \times \mathcal{R}_Y$.
(D) $\mathbb{E} \sup_{t\in[0,b]} \left| \left(\sigma \frac{\partial G}{\partial y}\right)(t, Y_t) \right|^2 < \infty$ for all $b \in [0,T)$.

(E) The process $\left(\left(\sigma^2 \frac{\partial^2 G}{\partial y^2}\right)(t, Y_t)\right)_{t\in[0,T)}$ is an L_2-martingale.

(F) The process X has a transition density Γ_X in the sense of Theorem B.1.

Items (A) and (B) are obvious, (C) is contained in [16, Preliminaries], (D) is [16, Lemma 5.2], and (E) is [16, Lemma 5.3].

This yields to the following setting:

Setting 6.2. In the notation of Assumption 4.5 we set

(1) $\sigma := (\sigma(Y_t))_{t\in[0,T]}$,

(2) $M := \left(\int_0^t \left(\sigma^2 \frac{\partial^2 G}{\partial y^2}\right)(u, Y_u) \mathrm{d}W_u\right)_{t\in[0,T)}$,

(3) $\varphi := \left(\frac{\partial G}{\partial y}(t, Y_t)\right)_{t\in[0,T)}$.

Lemma 6.8 and [19, Corollary 3.3] imply that Assumption 4.5 is fulfilled. To shorten the notation at some places we use

$$Z_t := \sigma_t \varphi_t, \quad \varphi(t,y) := \frac{\partial G}{\partial y}(t,y), \quad \text{and} \quad H_t := \sigma_t^2 \frac{\partial^2 G}{\partial y^2}(t, Y_t) \quad \text{for} \quad (t,y) \in [0,T) \times \mathcal{R}_Y.$$

Denote by $E(g;\tau) = (E_t(g;\tau))_{t\in[0,T]}$ the error process resulting from the difference between the stochastic integral and its Riemann approximation associated with the time-net $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$, i.e.

$$E_t(g;\tau) := \int_{(0,t]} \varphi_s \mathrm{d}Y_s - \sum_{i=1}^n \varphi_{t_{i-1}} (Y_{t_i\wedge t} - Y_{t_{i-1}\wedge t}) \quad \text{for } t \in [0,T].$$

For any $0 \leqslant a \leqslant t \leqslant T$, we apply the conditional Itô's isometry to obtain that, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\left[|E_t(g;\tau) - E_a(g;\tau)|^2\right] = \mathbb{E}^{\mathcal{F}_a}\left[\int_a^t \left|\varphi_u - \sum_{i=1}^n \varphi_{t_{i-1}} \mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma_u^2 \mathrm{d}u\right] = \mathbb{E}^{\mathcal{F}_a}[[\varphi;\tau]_t^\sigma - [\varphi;\tau]_a^\sigma]. \tag{6.2}$$

Using Proposition A.4 this implies, for $\Phi \in \mathrm{CL}^+([0,T))$, that

$$\left\|(E_t(g;\tau))_{t\in[0,T)}\right\|^2_{\mathrm{bmo}_2^\Phi([0,T))} = \left\|([\varphi;\tau]^\sigma)_{t\in[0,T)}\right\|_{\mathrm{bmo}_1^{\Phi^2}([0,T))}, \tag{6.3}$$

where $[\varphi;\tau]^\sigma$ is given in Assumption 4.5. Moreover, $\mathrm{bmo}_2^\Phi([0,T))$ and $\mathrm{bmo}_1^{\Phi^2}([0,T))$ above can be replaced by $\mathrm{BMO}_2^\Phi([0,T))$ and $\mathrm{BMO}_1^{\Phi^2}([0,T))$, respectively, due to the path continuity of $E(g;\tau)$ and $[\varphi;\tau]^\sigma$. To be in accordance with the previous sections we use in (6.3) the time interval $[0,T)$ instead of $[0,T]$.

6.1. **The results.** In this section we formulate the results, they are verified in Section 6.2. The first result shows that all gradient processes $(\varphi(t,Y_t))_{t\in[0,T)}$ have a large oscillation:

Theorem 6.3. *For $g \in C_Y$ the process $(\varphi(t,Y_t))_{t\in[0,T)}$ is of maximal oscillation with constant 2 in the sense of Definition 5.1.*

Now we discuss cases in which we get equivalences by choosing the weight Φ accordingly. For $\theta = 1$ we obtain a characterization in terms of Lipschitz functions that extends [20, Theorem 8]:

Theorem 6.4. *For $g \in C_Y$ and $\Phi = \sigma$ the following assertions are equivalent:*

(1) *There exists a Lipschitz function $\tilde{g}\colon \mathcal{R}_Y \to \mathbb{R}$ such that $g = \tilde{g}$ a.e. on $\mathcal{R}_Y$ with respect to the Lebesgue measure.*

(2) *There is a constant $c > 0$ such that $\|E(g;\tau)\|_{\mathrm{BMO}_2^\Phi([0,T))} \leqslant c\sqrt{\|\tau\|_1}$ for all $\tau \in \mathcal{T}$.*

In the case $\theta \in (0,1)$ we obtain an equivalence in terms of the Riemann–Liouville type integral (introduced in Section 3) of the gradient process:

Theorem 6.5. *Let $(\theta,q)\in(0,1)\times[2,\infty)$ and $\Phi=(\sigma_t\Psi_t)_{t\in[0,T)}$ where $\Psi\in \mathrm{CL}^+([0,T))$ is path-wise non-decreasing. If $g\in C_Y$ and if there is a constant $c_{(6.4)}>0$ such that, for $t\in[0,T)$,*

$$(T-t)^{\frac{1}{2}}|\varphi_t|\leqslant c_{(6.4)}\Psi_t \ \text{a.s.},\tag{6.4}$$

then the following assertions are equivalent:

(1) *One has $\mathcal{I}^{\frac{1-\theta}{2}}Z-Z_0\in \mathrm{BMO}_2^{\Phi}([0,T))$ and there is a constant $c>0$ such that, for all $t\in[0,T)$,*

$$(T-t)^{\frac{1-\theta}{2}}|\varphi_t|\leqslant c\Psi_t \quad \text{a.s.}\tag{6.5}$$

(2) *There is a constant $c>0$ such that $\|E(g;\tau)\|_{\mathrm{BMO}_2^{\Phi}([0,T))}\leqslant c\sqrt{\|\tau\|_\theta}$ for all $\tau\in\mathcal{T}$.*

If the conditions (1) and (2) are satisfied and $\Phi\in\mathcal{SM}_q([0,T))$, then $\mathcal{I}_T^{\frac{1-\theta}{2}}Z:=\lim_{t\uparrow T}\mathcal{I}_t^{\frac{1-\theta}{2}}Z$ exists in L_q and a.s.

Theorem 6.6. *Let $(\theta,q)\in(0,1)\times[2,\infty)$, $g\in \mathrm{Höl}^0_{\theta,2}(\mathbb{R})$, and $\Phi=(\sigma_t\Psi_t)_{t\in[0,T)}$ with $\Psi_t:=\sup_{s\in[0,t]}(\sigma_s^{\theta-1})$. Then one has $g|_{\mathcal{R}_Y}\in C_Y$ and the following holds:*

(1) *$\Phi\in\mathcal{SM}_q([0,T))$.*

(2) *There is a constant $c>0$ such that $(T-t)^{\frac{1-\theta}{2}}|\varphi_t|\leqslant c\Psi_t$ a.s. for all $t\in[0,T)$.*

(3) *$\mathcal{I}^{\frac{1-\theta}{2}}Z-Z_0\in\mathrm{BMO}_q^{\Phi}([0,T))$.*

6.2. **Preparations to prove the results of Section 6.1.** We collect some lemmas we need.

Lemma 6.7. *Assume that $\theta\in(0,1]$, $g\in C_Y$, that $\left(\mathcal{I}_t^{\frac{1-\theta}{2}}M\right)_{t\in[0,T)}$ is closable in L_2, and $\Phi\in \mathrm{CL}^+([0,T))$ such that*

$$\sup_{s\in[0,a]}\frac{T-a}{(T-s)^\theta}|\varphi_a-\varphi_s|^2\sigma_a^2+\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T)}\left|\mathcal{I}_t^{\frac{1-\theta}{2}}M-\mathcal{I}_a^{\frac{1-\theta}{2}}M\right|^2\right]\leqslant\Phi_a^2\quad \text{a.s.}\qquad \text{for } a\in[0,T).$$

Then there is a constant $c>0$ such that $\|E(g;\tau)\|_{\mathrm{BMO}_2^{\Phi}([0,T))}\leqslant c\sqrt{\|\tau\|_\theta}$ for all $\tau\in\mathcal{T}$.

Proof. The statement follows directly from the equivalence (4.6) in Theorem 4.7 and (6.3). □

Lemma 6.8. *The following assertions hold true:*

(1) *In the case (C2) one has $(Y_t^{\beta_0}(Y^{\beta_1})_t^*)_{t\in[0,T]}\in\mathcal{SM}_p([0,T])$ for $p\in(0,\infty)$ and $\beta_0,\beta_1\in\mathbb{R}$.*

(2) *There is a constant $c_{(6.6)}>0$ such that, for all $0\leqslant a<b\leqslant T$,*

$$\mathbb{E}^{\mathcal{F}_a}\left[\frac{1}{b-a}\int_a^b\sigma_u^2\mathrm{d}u\right]\sim_{c^2_{(6.6)}}\sigma_a^2\qquad \text{a.s.}\tag{6.6}$$

(3) *For $g\in C_Y$ one has $\mathbb{E}\sup_{u\in[a,T]}|\varphi_a\sigma_u|^2<\infty$ for $a\in[0,T)$.*

Proof. (1) Because $\hat\sigma\in B_b(\mathbb{R})$ for all $\alpha\in\mathbb{R}$ there is a constant $c_{(6.7)}=c_{(6.7)}(\alpha,T,\hat\sigma)>0$ such that

$$\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T]}\mathrm{e}^{\alpha\int_{(a,t]}\hat\sigma(X_s)\mathrm{d}W_s}\right]\le c_{(6.7)}\ \text{a.s.}\tag{6.7}$$

for $a\in[0,T]$. Because $\hat b$ is bounded this implies that $(Y_t^\beta)_{t\in[0,T]}\in\mathcal{SM}_p([0,T])$ for all $p\in(0,\infty)$ and $\beta\in\mathbb{R}$ by Proposition A.1. Therefore we may conclude by items (2) and (3) of Proposition A.2.

(2) We only need to check the case (C2) where we replace σ by Y by (B). As Y is a martingale we get $\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b Y_u^2\mathrm{d}u\right]\geqslant(b-a)Y_a^2$ a.s., otherwise $\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b Y_u^2\mathrm{d}u\right]\leqslant\|Y\|^2_{\mathcal{SM}_2([0,T])}(b-a)Y_a^2$ a.s.

(3) Because of (D) we only need to check (C2), use again (B) to replace σ by Y, and obtain

$$\mathbb{E}\sup_{u\in[a,T]}|\varphi_aY_u|^2=\mathbb{E}\left[|\varphi_a|^2\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,T]}Y_u^2\right]\right]\leqslant\|Y\|^2_{\mathcal{SM}_2([0,T]}\mathbb{E}|\varphi_aY_a|^2<\infty.$$ □

Lemma 6.9. *For $\theta \in (0,1]$, $\alpha := \frac{1-\theta}{2}$, and $t \in [0,T)$ one has, a.s.,*

$$\begin{aligned}(T-t)^\alpha Z_t =& T^\alpha Z_0 + \int_{(0,t]} (T-u)^\alpha H_u \mathrm{d}W_u + \int_{(0,t]} (T-u)^\alpha \sigma'(Y_u) Z_u \mathrm{d}W_u \\ & - \alpha \int_{(0,t]} (T-u)^{\alpha-1} Z_u \mathrm{d}u + \frac{1}{2}\int_{(0,t]} (T-u)^\alpha (\sigma\sigma'')(Y_u) Z_u \mathrm{d}u.\end{aligned}$$

Proof. The assertion follows by Itô's formula applied to the function $(t,y) \mapsto (T-t)^\alpha \left(\sigma \frac{\partial G}{\partial y}\right)(t,y)$ with Y_t inserted into the y-component, where we use the PDE from (C). □

Lemma 6.10. *For $\theta \in [0,1]$ there exists a constant $c_{(6.8)} > 0$ such that for all $g \in \mathrm{Höl}_\theta(\mathbb{R})$ one has*

$$\left|\frac{\partial G}{\partial y}(u,y)\right| \leqslant c_{(6.8)} \, |g|_\theta \, \sigma(y)^{\theta-1}(T-u)^{\frac{\theta-1}{2}} \quad \textit{for } (u,y) \in \mathcal{R}_Y \times [0,T). \tag{6.8}$$

Proof. Set $f := g$ and $F := G$ in case (C1) and $f(x) := g(\mathrm{e}^x)$ and $F(u,x) := G(u,\mathrm{e}^x)$ for $(u,x) \in [0,T) \times \mathbb{R}$ in case (C2), and let us fix $u \in [0,T)$. In both cases, (C1) and (C2), we have

$$\frac{\partial F}{\partial x}(u,x) = \int_\mathbb{R} \frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi) f(\xi)\mathrm{d}\xi = \int_\mathbb{R} \frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi)(f(\xi)-f(x))\mathrm{d}\xi$$

where we use (F) with the transition density Γ_X from Theorem B.1. For $t>0$ denote $\gamma_t(x) := \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{x^2}{2t}}$. In the case (C1) we derive that

$$\begin{aligned}\left|\frac{\partial G}{\partial y}(u,y)\right| = \left|\frac{\partial F}{\partial x}(u,x)\right| \quad &\leqslant \quad |g|_\theta \int_\mathbb{R} \left|\frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi)\right| |\xi - x|^\theta \mathrm{d}\xi \\ &\leqslant \quad |g|_\theta \int_\mathbb{R} c_{(B.1)}(T-u)^{-\frac{1}{2}} \gamma_{c_{(B.1)}(T-u)}(x-\xi)|\xi-x|^\theta \mathrm{d}\xi \\ &= \quad |g|_\theta (T-u)^{\frac{\theta-1}{2}} \int_\mathbb{R} c_{(B.1)} \gamma_{c_{(B.1)}}(\eta)|\eta|^\theta \mathrm{d}\eta\end{aligned}$$

where we use $\int_\mathbb{R} \frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi)\mathrm{d}\xi = \frac{\partial}{\partial x}\int_\mathbb{R} \Gamma_X(T-u,x,\xi)\mathrm{d}\xi = 0$. For $y = \mathrm{e}^x$ we get for (C2) that

$$\begin{aligned}\left|y\frac{\partial G}{\partial y}(u,y)\right| = \left|\frac{\partial F}{\partial x}(u,x)\right| \quad &\leqslant \quad |g|_\theta \int_\mathbb{R} \left|\frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi)\right| |\, \mathrm{e}^\xi - \mathrm{e}^x \,|^\theta \mathrm{d}\xi \\ &= \quad |g|_\theta \, \mathrm{e}^{x\theta} \int_\mathbb{R} \left|\frac{\partial \Gamma_X}{\partial x}(T-u,x,\xi)\right| |\, \mathrm{e}^{\xi-x} - 1|^\theta \mathrm{d}\xi \\ &\leqslant \quad |g|_\theta \, \mathrm{e}^{x\theta} \int_\mathbb{R} c_{(B.1)}(T-u)^{-\frac{1}{2}} \gamma_{c_{(B.1)}(T-u)}(x-\xi)|\, \mathrm{e}^{\xi-x} - 1|^\theta \mathrm{d}\xi.\end{aligned}$$

We conclude by

$$\begin{aligned}\int_\mathbb{R} \gamma_{c_{(B.1)}(T-u)}(x-\xi)|\, \mathrm{e}^{\xi-x} - 1|^\theta \mathrm{d}\xi &\leqslant \int_\mathbb{R} \gamma_{c_{(B.1)}(T-u)}(\xi)|\xi|^\theta \, \mathrm{e}^{\theta|\xi|} \, \mathrm{d}\xi \\ &\leqslant (T-u)^{\frac{\theta}{2}} \int_\mathbb{R} \gamma_{c_{(B.1)}}(\eta)|\eta|^\theta \, \mathrm{e}^{\theta\sqrt{T}|\eta|} \, \mathrm{d}\eta < \infty.\end{aligned}$$

□

Lemma 6.11. *Let $\mathrm{d}\hat{\mathbb{P}} := L\mathrm{d}\mathbb{P}$ with $L := \exp\left(\int_{(0,T]} \sigma'(Y_t)\mathrm{d}W_t - \frac{1}{2}\int_{(0,T]} |\sigma'(Y_t)|^2 \mathrm{d}t\right)$ and $g \in C_Y$. Then the process $(\varphi(t,Y_t))_{t\in[0,T)}$ is a $\hat{\mathbb{P}}$-martingale.*

Proof. Applying the PDE from (C) we get that

$$\frac{\partial \varphi}{\partial t}(t,y) + (\sigma\sigma')(y)\frac{\partial \varphi}{\partial y}(t,y) + \frac{\sigma^2(y)}{2}\frac{\partial^2 \varphi}{\partial y^2}(t,y) = \frac{\partial}{\partial y}\left[\frac{\partial G}{\partial t}(t,y) + \frac{\sigma^2(y)}{2}\frac{\partial^2 G}{\partial y^2}(t,y)\right] = 0$$

on $[0,T) \times \mathcal{R}_Y$. By Itô's formula this implies that

$$\varphi(t,Y_t) = \varphi(0,y_0) + \int_{(0,t]} \left(\sigma\frac{\partial \varphi}{\partial y}\right)(u,Y_u)\,[dW_u - \sigma'(Y_u)\mathrm{d}u] \;\text{ a.s.}$$

for $t \in [0,T)$. Because of (A) and Girsanov's theorem we obtain a $\hat{\mathbb{P}}$ standard Brownian motion $\hat{W}_t := W_t - \int_{(0,t]} \sigma'(Y_u)\mathrm{d}u$, $t \in [0,T]$. Moreover, for $t \in [0,T)$ we have that

$$\mathbb{E}^{\hat{\mathbb{P}}} \left| \int_0^t \left| \left(\sigma \frac{\partial \varphi}{\partial y}\right)(u,Y_u)\right|^2 \mathrm{d}u \right|^{\frac{1}{2}} \leqslant (\mathbb{E}^{\mathbb{P}} L^2)^{\frac{1}{2}} \left| \mathbb{E}^{\mathbb{P}} \int_0^t \left| \left(\sigma \frac{\partial \varphi}{\partial y}\right)(u,Y_u)\right|^2 \mathrm{d}u \right|^{\frac{1}{2}} < \infty.$$

As by the Burkholder-Davis-Gundy inequalities applied to continuous local martingales we also have

$$\mathbb{E}^{\hat{\mathbb{P}}} \left| \int_{(0,t]} \left| \left(\sigma \frac{\partial \varphi}{\partial y}\right)(u,Y_u)\right|^2 \mathrm{d}u \right|^{\frac{1}{2}} \sim_c \mathbb{E}^{\hat{\mathbb{P}}} \sup_{s\in[0,t]} \left| \int_{(0,s]} \left(\sigma \frac{\partial \varphi}{\partial y}\right)(u,Y_u) d\hat{W}_u \right|$$

for some absolute constant $c \geqslant 1$ and $t \in [0,T)$, we get that $(\varphi(t,Y_t))_{t\in[0,T)}$ is a $\hat{\mathbb{P}}$-martingale. □

6.3. **Proof of Theorem 6.3.** According to Lemma 6.11 there is an equivalent measure $\hat{\mathbb{P}} \sim \mathbb{P}$ such that $(\varphi(t,Y_t))_{t\in[0,T)}$ is a $\hat{\mathbb{P}}$-martingale. The transition density of Y under $\mathbb{P}$ computes as

$$\Gamma_Y(s,t;y_1,y_2) = \frac{1}{y_2}\Gamma_X(s,t;\log(y_1),\log(y_2)) \tag{6.9}$$

in the case (C2), otherwise $\Gamma_Y = \Gamma_X$, where Γ_X is taken from Theorem B.1 in both cases. We conclude by Example 5.5, where relation (5.1) follows from Theorem B.1, the uniqueness in law of the SDE (6.1), and the theory of Markov processes. □

6.4. **Proof of Theorem 6.4.** (1) ⇒ (2) We may assume that $g\colon \mathcal{R}_Y \to \mathbb{R}$ is Lipschitz. By Lemma 6.10 we have

$$\left|\frac{\partial G}{\partial y}(u,y)\right| \leqslant c_{(6.8)}|g|_1 \quad \text{and} \quad |Z_u| \leqslant c_{(6.8)}|g|_1\sigma_u.$$

Let $0 \leqslant a < t < T$. From Lemma 6.9 we get that

$$Z_t = Z_a + \int_{(a,t]} H_u \mathrm{d}W_u + \int_{(a,t]} \sigma'(Y_u)Z_u \mathrm{d}W_u + \frac{1}{2}\int_{(a,t]} (\sigma\sigma'')(Y_u)Z_u\mathrm{d}u \quad \text{a.s.}$$

Then one has, a.s.,

$$\begin{aligned}
&\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\int_a^t H_u^2\mathrm{d}u\right]} \\
&\leqslant \sqrt{\mathbb{E}^{\mathcal{F}_a}[|Z_t - Z_a|^2]} + \|\sigma'\|_{B_b(\mathcal{R}_Y)}\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\int_a^t Z_u^2\mathrm{d}u\right]} + \frac{1}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)}\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\left|\int_a^t |Z_u|\mathrm{d}u\right|^2\right]} \\
&\leqslant \sqrt{\mathbb{E}^{\mathcal{F}_a}[|Z_t - Z_a|^2]} + \left[\|\sigma'\|_{B_b(\mathcal{R}_Y)} + \frac{\sqrt{T}}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)}\right]\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\int_a^t Z_u^2\mathrm{d}u\right]} \\
&\leqslant c_{(6.8)}|g|_1\left[\sqrt{\mathbb{E}^{\mathcal{F}_a}[\sigma_t^2]} + \sigma_a\right] + c_{(6.8)}|g|_1\left[\|\sigma'\|_{B_b(\mathcal{R}_Y)} + \frac{\sqrt{T}}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)}\right]\sqrt{T}\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in(a,T]}\sigma_u^2\right]} \\
&\leqslant c_{(6.8)}|g|_1\left[2 + \sqrt{T}\|\sigma'\|_{B_b(\mathcal{R}_Y)} + \frac{T}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)}\right]\sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,T]}\sigma_u^2\right]} \\
&\leqslant c_{(6.8)}|g|_1\left[2 + \sqrt{T}\|\sigma'\|_{B_b(\mathcal{R}_Y)} + \frac{T}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)}\right]\|\sigma\|_{\mathcal{SM}_2([0,T])}\sigma_a
\end{aligned}$$

and hence

$$\sqrt{\mathbb{E}^{\mathcal{F}_a}[|M_t - M_a|^2]} = \sqrt{\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} H_u^2\mathrm{d}u\right]} \leqslant c_{(6.10)}|g|_1\|\sigma\|_{\mathcal{SM}_2([0,T])}\sigma_a \text{ a.s.,} \tag{6.10}$$

for some $c_{(6.10)} > 0$. Applying Lemma 6.10 for $\theta = 1$ and (6.10) (together with Doob's maximal inequality) to Lemma 6.7 for $\Phi_a = c\sigma_a$ for some appropriate $c > 0$ and $\theta = 1$, we derive (2).

(2) $\Rightarrow$ (1) Given $a \in (0,T)$, exploiting the last term in the relation (4.6) of Theorem 4.7 and (6.2) give

$$\sup_{s\in[0,a]} \frac{T-a}{T-s}|\varphi_a - \varphi_s|^2 \leqslant c_{(6.11)}^2 \text{ a.s.} \tag{6.11}$$

For $a \in \left(\frac{T}{2}, T\right)$ we choose $s \in [0,a)$ such that $\frac{T-a}{T-s} = \frac{1}{2}$. Therefore we may continue to

$$\left|\frac{\partial G}{\partial y}(a, y_a)\right| \leqslant \left|\frac{\partial G}{\partial y}(s, y_s)\right| + \sqrt{2}c_{(6.11)} \quad \text{for all} \quad y_a, y_s \in \mathcal{R}_Y$$

where we use the positivity and continuity of the transition density Γ_Y (for (C2) see (6.9)) and the continuity of $\frac{\partial G}{\partial y}(t,\cdot) : \mathcal{R}_Y \to \mathbb{R}$ for $t \in [0,T)$. Applying Lemma 6.11, we have $\mathbb{E}^{\hat{\mathbb{P}}}\varphi(s, Y_s) = \varphi(0, Y_0)$ for $s \in [0,T)$. Therefore, for each $s \in [0,T)$ there are $\omega_s^0, \omega_s^1 \in \Omega$ such that for $y_s^i := Y_s(\omega_s^i) \in \mathcal{R}_Y$ we have $\varphi(s, y_s^0) \leqslant \varphi(0, Y_0) \leqslant \varphi(s, y_s^1)$. Because $y \to \frac{\partial G}{\partial y}(s,y)$ is continuous on $\mathcal{R}_Y$ we find an $y_s \in \mathcal{R}_Y$ such that $\varphi(s, y_s) = \varphi(0, y_0)$. Therefore,

$$\left|\frac{\partial G}{\partial y}(a, y)\right| \leqslant \left|\frac{\partial G}{\partial y}(0, y_0)\right| + \sqrt{2}c_{(6.11)} =: c_{(6.12)} \quad \text{for all} \quad y \in \mathcal{R}_Y. \tag{6.12}$$

Let $\Omega_g \in \mathcal{F}$ be of measure one such that for all $\omega \in \Omega_g$ one has

$$\lim_{t\uparrow T} G(t, Y_t(\omega)) = g(Y_T(\omega)).$$

Let $I_g := Y_T(\Omega_g) \subseteq \mathcal{R}_Y$. Then g is Lipschitz on I_g with Lipschitz constant $c_{(6.12)}$, and since I_g is dense in $\mathcal{R}_Y$, the function $g|_{I_g}$ can be extended to $\tilde{g}\colon \mathcal{R}_Y \to \mathbb{R}$ to a Lipschitz function. Moreover, $\mathbb{P}(\{\omega \in \Omega : g(Y_T(\omega)) = \tilde{g}(Y_T(\omega))\}) \geqslant \mathbb{P}(\Omega_g) = 1$. $\square$

6.5. **Proof of Theorem 6.5.** Let $\alpha := \frac{1-\theta}{2}$. Observe that with

$$\alpha \int_0^T (T-u)^{\alpha-1} Z_{u\wedge t} \mathrm{d}u = \alpha \int_0^t (T-u)^{\alpha-1} Z_u \mathrm{d}u + (T-t)^\alpha Z_t$$

Lemma 6.9 implies that

$$\begin{aligned}\alpha \int_{(0,T]} (T-u)^{\alpha-1} Z_{u\wedge t} \mathrm{d}u =& T^\alpha Z_0 + \int_{(0,t]} (T-u)^\alpha H_u \mathrm{d}W_u + \int_{(0,t]} (T-u)^\alpha \sigma'(Y_u) Z_u \mathrm{d}W_u \\ &+ \frac{1}{2}\int_{(0,t]} (T-u)^\alpha (\sigma\sigma'')(Y_u) Z_u \mathrm{d}u \text{ a.s.}\end{aligned}$$

Denote $b_u(\omega) := \frac{1}{2}(\sigma\sigma'')(Y_u(\omega))$ and $B := \frac{1}{2}\|\sigma\sigma''\|_{B_b(\mathcal{R}_Y)} < \infty$. Dividing both sides of the equality above by T^α gives

$$\mathcal{I}_t^\alpha Z = Z_0 + \int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha H_u \mathrm{d}W_u + \int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha Z_u (\sigma'(Y_u)\mathrm{d}W_u + b_u \mathrm{d}u).$$

Next we observe that, for $0 \leqslant a < t < T$, a.s.,

$$\begin{aligned}&\left(\mathbb{E}^{\mathcal{F}_a}\left[\left|\int_{(a,t]} \left(\frac{T-u}{T}\right)^\alpha Z_u \sigma'(Y_u) \mathrm{d}W_u\right|^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}^{\mathcal{F}_a}\left[\left|\int_{(a,t]} \left(\frac{T-u}{T}\right)^\alpha |Z_u b_u| \mathrm{d}u\right|^2\right]\right)^{\frac{1}{2}} \\ &\leqslant (\|\sigma'\|_{B_p(\mathcal{R}_Y)} + B\sqrt{T}) \left(\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,t]} \left(\frac{T-u}{T}\right)^{2\alpha} |Z_u|^2 \mathrm{d}u\right]\right)^{\frac{1}{2}} \\ &\leqslant c_{(6.4)}(\|\sigma'\|_{B_b(\mathcal{R}_Y)} + B\sqrt{T}) \left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,T)} \Phi_u^2 \int_{(a,t]} \left(\frac{T-u}{T}\right)^{2\alpha} (T-u)^{-1} \mathrm{d}u\right]\right)^{\frac{1}{2}} \\ &\leqslant \frac{c_{(6.4)}(\|\sigma'\|_{B_b(\mathcal{R}_Y)} + B\sqrt{T})}{\sqrt{2\alpha}} \|\Phi\|_{\mathcal{SM}_2([0,T))} \left(\frac{T-a}{T}\right)^\alpha \Phi_a.\end{aligned}$$

We conclude that the martingale $(\int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha Z_u \sigma'(Y_u) \mathrm{d}W_u)_{t\in[0,T)}$ converges in L_q and a.s. because of $\Phi \in \mathcal{SM}_q([0,T))$ and Proposition A.6(2). Again by Proposition A.6(2), the non-negative and non-decreasing process $\left(\int_0^t \left(\frac{T-u}{T}\right)^\alpha |Z_u b_u| \mathrm{d}u\right)_{t\in[0,T)}$ converges in L_q and a.s. For this reason $\left(\int_0^t \left(\frac{T-u}{T}\right)^\alpha Z_u b_u \mathrm{d}u\right)_{t\in[0,T)}$ converges in L_q and a.s. as well. If we set

$$M^\alpha := \left(\int_{(0,t]} \left(\frac{T-u}{T}\right)^\alpha H_u \mathrm{d}W_u\right)_{t\in[0,T)},$$

then we can summarize as follows:

(a) $(\mathcal{I}_t^\alpha Z - Z_0)_{t\in[0,T)} \in \mathrm{BMO}_2^\Phi([0,T))$ if and only if $M^\alpha \in \mathrm{BMO}_2^\Phi([0,T))$.
(b) $\mathcal{I}^\alpha Z$ converges (is bounded) in L_q if and only if M^α does (is).
(c) $\mathcal{I}^\alpha Z$ converges a.s. if and only if M^α does.

(1) $\Rightarrow$ (2) By (a) we get $M^\alpha \in \mathrm{BMO}_2^\Phi([0,T))$. Because the Setting 6.2 and (6.5) hold we may use Theorem 4.8((1) $\Rightarrow$ (2)) and conclude by (6.3).

(2) $\Rightarrow$ (1) follows from (6.3), the validity of Setting 6.2, Theorem 4.8((2) $\Rightarrow$ (1)), and (a).

Regarding the final part we deduce from (1), $\Phi \in \mathcal{SM}_q([0,T))$, and Proposition A.6(2) that $\sup_{t\in[0,T)} |\mathcal{I}_t^\alpha Z| \in L_q$, conclude $\sup_{t\in[0,T)} \|M_t^\alpha\|_{L_q} < \infty$ by (b), and obtain from the martingale property the L_q- and a.s. convergence of M^α. We may finish by (b) and (c). $\square$

6.6. **Proof of Theorem 6.6.** (1) We only need to check the case (C2) and this case follows from Lemma 6.8(1). Item (2) follows directly from Lemma 6.10.

(3) We fix $a \in [0,T)$, a set $A \in \mathcal{F}_a$ of positive measure. First we observe that by (4.1) (applied to $s = a$ and with $b \uparrow T$), Lemma 6.10 for $\theta = 0$, and Lemma 6.8,

$$\begin{aligned}
\sqrt{\int_A \int_a^T (T-u) H_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}} &\sim_{\sqrt{\kappa}} \sqrt{\int_A \int_a^T |\varphi_u - \varphi_a|^2 \sigma_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}} \\
&\leqslant \sqrt{\int_A \int_a^T Z_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}} + \sqrt{\int_A \varphi_a^2 \int_a^T \sigma_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}} \\
&\leqslant \sqrt{\int_A g(Y_T)^2 \mathrm{d}\mathbb{P}} + \sqrt{\int_A \left[c_{(6.8)}^2 |g|_0^2 \sigma_a^{-2} (T-a)^{-1}\right]\left[c_{(6.6)}^2 (T-a)\sigma_a^2\right] \mathrm{d}u\mathrm{d}\mathbb{P}} \\
&\leqslant c_0 \|g\|_{B_b(\mathcal{R}_Y)} \sqrt{\mathbb{P}(A)}.
\end{aligned}$$

On the other hand (6.10) gives

$$\sqrt{\int_A \int_a^T H_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}} \leqslant c_{(6.10)} |g|_1 \|\sigma\|_{\mathcal{SM}_2([0,T])} \sqrt{\int_A \sigma_a^2 \mathrm{d}\mathbb{P}}.$$

For the linear map $T : g \mapsto (H_u)_{u\in[a,T)}$ we get

$$\left\|T : C_b^0(\mathbb{R}) \to L_2([a,T) \times A, ((T-\cdot)\lambda \otimes \mathbb{P}_A))\right\| \leqslant c_0, \tag{6.13}$$

$$\left\|T : \mathrm{H\ddot{o}l}_1^0(\mathbb{R}) \to L_2([a,T) \times A, \lambda \otimes \mathbb{P}_A)\right\| \leqslant c_1 \sqrt{\int_A \sigma_a^2 \mathrm{d}\mathbb{P}_A}, \tag{6.14}$$

where $\mathbb{P}_A$ is the normalized restriction of $\mathbb{P}$ to A. Applying the Stein-Weiss interpolation theorem [7, Theorem 5.4.1] to (6.13) and (6.14) yields

$$\left\|T : (C_b^0(\mathbb{R}), \mathrm{H\ddot{o}l}_1^0(\mathbb{R}))_{\theta,2} \to L_2([a,T) \times A, ((T-\cdot)^{1-\theta}\lambda \otimes \mathbb{P}_A))\right\| \leqslant c_{(6.15)} \left(\int_A \sigma_a^2 \mathrm{d}\mathbb{P}_A\right)^{\frac{\theta}{2}}, \tag{6.15}$$

with $c_{(6.15)} := C c_0^{1-\theta} c_1^\theta$. In other words, we did prove

$$\left(\int_A \int_a^t (T-u)^{1-\theta} H_u^2 \mathrm{d}u\mathrm{d}\mathbb{P}_A\right)^{\frac{1}{2}} \leqslant c_{(6.15)} \left(\int_A \sigma_a^2 \mathrm{d}\mathbb{P}_A\right)^{\frac{\theta}{2}} \|g\|_{\mathrm{H\ddot{o}l}_{\theta,2}^0(\mathbb{R})}.$$

For $\delta \in (0,1)$ and $l \in \mathbb{Z}$ define $A_l := \{\delta^{l+1} < \sigma_a^2 \leqslant \delta^l\}$. Then

$$\begin{aligned}
\int_A \int_a^t (T-u)^{1-\theta} H_u^2 \mathrm{d}u \mathrm{d}\mathbb{P}_A &= \sum_{\mathbb{P}(A\cap A_l)>0} \left(\int_{A\cap A_l} \int_a^t (T-u)^{1-\theta} H_u^2 \mathrm{d}u \mathrm{d}\mathbb{P}_{A\cap A_l} \right) \mathbb{P}_A(A\cap A_l) \\
&\leqslant c_{(6.15)}^2 \sum_{\mathbb{P}(A\cap A_l)>0} \left(\int_{A\cap A_l} \sigma_a^2 \mathrm{d}\mathbb{P}_{A\cap A_l} \right)^\theta \mathbb{P}_A(A\cap A_l) \, \|g\|^2_{\mathrm{H\ddot{o}l}^0_{\theta,2}(\mathbb{R})} \\
&\leqslant c_{(6.15)}^2 \sum_{\mathbb{P}(A\cap A_l)>0} \delta^{l\theta} \mathbb{P}_A(A\cap A_l) \, \|g\|^2_{\mathrm{H\ddot{o}l}^0_{\theta,2}(\mathbb{R})} \\
&\leqslant c_{(6.15)}^2 \delta^{-\theta} \int_A \sigma_a^{2\theta} \mathrm{d}\mathbb{P}_A \, \|g\|^2_{\mathrm{H\ddot{o}l}^0_{\theta,2}(\mathbb{R})}.
\end{aligned}$$

As $\delta \in (0,1)$ was arbitrary, we conclude

$$\int_A \int_a^t (T-u)^{1-\theta} H_u^2 \mathrm{d}u \mathrm{d}\mathbb{P}_A \leqslant c_{(6.15)}^2 \int_A \sigma_a^{2\theta} \mathrm{d}\mathbb{P}_A \, \|g\|^2_{\mathrm{H\ddot{o}l}^0_{\theta,2}(\mathbb{R})}.$$

and

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_a^t (T-u)^{1-\theta} H_u^2 \mathrm{d}u \right] \leqslant c_{(6.15)}^2 \sigma_a^{2\theta} \leqslant c_{(6.15)}^2 \Phi_a^2 \text{ a.s.}$$

We use item (a) from the proof of Theorem 6.5 to conclude that $\left(\mathcal{I}_t^{\frac{1-\theta}{2}} Z - Z_0\right)_{t\in[0,T)} \in \mathrm{BMO}_2^\Phi([0,T))$ and finish by $\Phi \in \mathcal{SM}_q([0,T))$ and Proposition A.6(1). □

7. An interpolation result

The following interpolation result is adapted to prove Corollary 8.13, but of independent interest. For this section we assume

(1) $\kappa_0, \kappa_1 \in (0,1)$ and $0 \leqslant \gamma_0 < \gamma_1 < \infty$,
(2) a probability space $(R, \mathcal{R}, \rho)$,
(3) an interpolation pair of Banach spaces (E_0, E_1) and a Banach space F,
(4) random variables $A_0, A_1 : R \to [0,\infty)$ with $\rho(\{2^n < A_i \leqslant 2^{n+1}\}) \leqslant c_{A_i} 2^{-\kappa_i n}$ for $n \in \mathbb{N}_0$,
(5) for $(t,r) \in [0,T) \times R$ linear operators $T_{t,\rho}, T_{t,r} : E_0 + E_1 \to F$ such that
   (a) $\|T_{t,r}x\|_F \leqslant c_i \min\{A_i(r), (T-t)^{-\gamma_i}\}\|f\|_{E_i}$ for $f \in E_i$ and $r \in R$,
   (b) if $\|T_{t,\cdot}x\|_F \leqslant P(\cdot)$ on R, where $P : R \to [0,\infty)$ is measurable and $x \in E_0 + E_1$, then $\|T_{t,\rho}x\|_F \leqslant \int_R P(r)\rho(\mathrm{d}r)$,
(6) $\|T_{s,\rho}x\|_F \leqslant \|T_{t,\rho}x\|_F$ for all $0 \leqslant s < t < T$ and $x \in E_0 + E_1$,

where $c_{A_0}, c_{A_1}, c_0, c_1 > 0$ are constants. Note that the map $[0,T) \ni t \mapsto \|T_{t,\rho}x\|_F$ is measurable by assumption (6). Under the above assumptions the following statement holds:

Theorem 7.1. *For all $(\delta, q) \in (0,1) \times [1,\infty]$ there is a $c_{(7.1)}(\delta, q, c_{A_0}, c_{A_1}, c_0, c_1, \kappa_0, \kappa_1, \gamma_0, \gamma_1, T) > 0$ such that, for $\alpha := (1-\delta)(1-\kappa_0)\gamma_0 + \delta(1-\kappa_1)\gamma_1$,*

$$\|(T-t)^\alpha \|T_{t,\rho}x\|_F\|_{L_q([0,T),\frac{\mathrm{d}t}{T-t})} \leqslant c_{(7.1)} \|f\|_{(E_0,E_1)_{\delta,q}} \quad \textit{for} \quad x \in (E_0,E_1)_{\delta,q}. \tag{7.1}$$

Proof. First we observe that

$$\begin{aligned}
\|T_{t,\rho}x\|_F &\leqslant c_i \int_R \min\{A_i(r), (T-t)^{-\gamma_i}\}\rho(\mathrm{d}r)\|x\|_{E_i} \\
&\leqslant c_i \left[1 + \sum_{n=0}^\infty \int_{\{2^n < A_i \leqslant 2^{n+1}\}} \min\{2^{n+1}, (T-t)^{-\gamma_i}\}\rho(\mathrm{d}r) \right] \|x\|_{E_i} \\
&\leqslant c_i \left[1 + c_{A_i} \sum_{n=0}^\infty 2^{-\kappa_i n} \min\{2^{n+1}, (T-t)^{-\gamma_i}\} \right] \|x\|_{E_i} \\
&\leqslant c_i(2c_{A_i} \vee 1) \left[1 + \sum_{n=0}^\infty 2^{-\kappa_i n} \min\{2^n, (T-t)^{-\gamma_i}\} \right] \|x\|_{E_i} \\
&\leqslant c_i(2c_{A_i} \vee 1) c_{\kappa_i,\gamma_i,T} (T-t)^{(\kappa_i-1)\gamma_i} \|x\|_{E_i} \\
&=: d_i (T-t)^{(\kappa_i-1)\gamma_i} \|x\|_{E_i}.
\end{aligned}$$

For $t_k := T(1-\frac{1}{2^k})$, $k \in \mathbb{N}_0$, and $\alpha_i := (1-\kappa_i)\gamma_i$ this gives

$$\|(T_{t_k,\rho}x)_{k\in\mathbb{N}_0}\|_{\ell_\infty^{-\alpha_i}(F)} \leqslant d_i T^{-\alpha_i}\|f\|_{E_i}.$$

Using real interpolation, (2.4), and (2.2), we derive

$$\|(T_{t_k,\rho}x)_{k\in\mathbb{N}_0}\|_{\ell_q^{-\alpha}(F)} \leqslant c_{\alpha_0,\alpha_1,\delta,q}\, d_0^{1-\delta} d_1^{\delta}\, T^{-\alpha}\|f\|_{(E_0,E_1)_{\theta,q}}.$$

The assertion follows, because assumption (6) implies that

$$\|(T-t)^\alpha \|T_{t,\rho}x\|_F\|_{L_q([0,T),\frac{\mathrm{d}t}{T-t})} \sim_c T^\alpha \|(T_{t_k,\rho}x)_{k\in\mathbb{N}_0}\|_{\ell_q^{-\alpha}(F)},$$

where $c \geqslant 1$ depends at most on (α, q). □

8. Lévy setting: Directional gradient estimates and applications

8.1. **Setting.** Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $X_0 \equiv 0$, X has stationary and independent increments, and càdlàg trajectories. Assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of X and $\mathcal{F} = \mathcal{F}_T$. The Poisson random measure N associated to X is defined by $N(E) := \#\{t \in (0,T] : (t, \Delta X_t) \in E\}$ for $E \in \mathcal{B}((0,T]\times\mathbb{R}\backslash\{0\})$, the Lévy measure ν is the unique σ-finite Borel measure on $\mathbb{R}$ such that $\nu(\{0\}) = 0$ and $\nu(B) := \frac{1}{T}\mathbb{E}N((0,T]\times B)$ for Borel sets B with $0 \notin \overline{B}$. Let $\sigma \geqslant 0$ be the coefficient of the standard Brownian motion W in the Lévy–Itô decomposition of X (see, e.g., [38, Theorem 19.2]). We define the σ-finite measure μ on $\mathcal{B}(\mathbb{R})$ by

$$\mu(\mathrm{d}x) := \sigma^2\delta_0(\mathrm{d}x) + x^2\nu(\mathrm{d}x). \tag{8.1}$$

To avoid degenerate settings we always assume that $\mu(\mathbb{R}) \in (0,\infty]$. The compensated random measure $\widetilde{N}$ of N is given by $\widetilde{N} := N - \lambda\otimes\nu$ on the ring of $E \in \mathcal{B}((0,T]\times\mathbb{R})$ such that $(\lambda\otimes\mu)(E) < \infty$. In the sequel we use of the following notation:

Definition 8.1. A Borel function $f\colon \mathbb{R}\to\mathbb{R}$ belongs to $\mathcal{D}_X$ if $\mathbb{E}|f(x+X_s)| < \infty$ for all $(s,x) \in [0,T]\times\mathbb{R}$. For $f \in \mathcal{D}_X$ we define $F\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ by

$$F(t,x) := \mathbb{E}f(x+X_{T-t}). \tag{8.2}$$

8.2. **Galtchouk-Kunita-Watanabe projection.** We additionally assume that $X = (X_t)_{t\in[0,T]}$ is an L_2-martingale so that $\mu(\mathbb{R}) \in (0,\infty)$ and assume that $f \in L_2(\mathbb{R}, \mathbb{P}_{X_T})$. Let $D \in L_2(\mathbb{R},\mu)$ such that $D \geqslant 0$ and $\int_\mathbb{R} D^2(z)\mu(\mathrm{d}z) > 0$ and define $\mathrm{d}\rho := D\mathrm{d}\mu / \int_\mathbb{R} D\mathrm{d}\mu$. If $\xi = f(X_T)$ has a chaos decomposition as in Lemma D.1(1) (the notion of the chaos decomposition is recalled in Appendix D as well), then we let[1]

$$h_0 := \int_\mathbb{R} f_1^s(z)\rho(\mathrm{d}z) \quad \text{and} \quad h_n(x_1,\ldots,x_n) := \int_\mathbb{R} f_{n+1}^s(x_1,\ldots,x_n,z)\rho(\mathrm{d}z)$$

for $n \in \mathbb{N}$, define the càdlàg L_2-martingale $\varphi(f,\rho) = (\varphi_t(f,\rho))_{t\in[0,T)}$ by the chaos expansion

$$\varphi_t(f,\rho) := \sum_{n=0}^\infty (n+1) I_n(h_n \mathbb{1}_{(0,t]}^{\otimes n}), \tag{8.3}$$

and the càdlàg martingale $X^D = (X_t^D)_{t\in[0,T]}$ by $X_0^D \equiv 0$ and $X_t^D := I_1(\mathbb{1}_{(0,t]}\otimes D)$ a.s. Denote by $\mathcal{P}_{X^D}\colon L_2 \to I(X^D) \subseteq L_2$ the orthogonal projection onto the closed subspace

$$I(X^D) := \left\{ \int_{(0,T)} \vartheta_t \mathrm{d}X_t^D : \vartheta \text{ is predictable with } \mathbb{E}\int_0^T |\vartheta_t|^2\mathrm{d}t < \infty \right\}.$$

Then

$$\mathcal{P}_{X^D}(f(X_T)) = \frac{\int_\mathbb{R} D\mathrm{d}\mu}{\int_\mathbb{R} D^2\mathrm{d}\mu}\int_{(0,T)} \varphi_{t-}(f,\rho)\mathrm{d}X_t^D \text{ a.s.}$$

For $D \equiv 1$ this was shown in [17, (8), (10), Example (c1) on. p. 209, Lemma 4]. We omit the proof of this extension. The following statement is one motivation of Section 8 and will be used in Section 8.5.

[1]There might be a symmetric $\mu^{\otimes n}$-null-set in $(x_1,\ldots,x_n)$ on which the integral does not exist. On this set we set h_n to be 0.

Proposition 8.2 (Gradient of GKW-projection)**.** *Assume that the Lévy process X is an L_2-martingale, that $f \in \mathcal{D}_X \cap L_2(\mathbb{R}, \mathbb{P}_{X_T})$ and F is given by (8.2), that $\mathrm{d}\rho = D\mathrm{d}\mu / \int_{\mathbb{R}} D\mathrm{d}\mu$ as above, and that $t \in (0,T)$. Then there is a null-set $N_t \in \mathcal{F}$ such that for $\omega \notin N_t$ one has*[2]

$$\varphi_t(f,\rho)(\omega) = \rho(\{0\}) \frac{\partial F}{\partial x}(t, X_t(\omega)) + \int_{\mathbb{R}\setminus\{0\}} \frac{F(t, X_t(\omega)+z) - F(t, X_t(\omega))}{z} \rho(\mathrm{d}z). \tag{8.4}$$

We prove Proposition 8.2 in Appendix D for the convenience of the reader. Results related to Proposition 8.2 are provided in [28, Theorem 2.4], [6, Theorems 2.1, 3.11, 4.1], and [12, Proposition 2]. Other techniques use the Fourier transform (see, e.g., [9]).

8.3. **Upper bounds for the gradient process.** Gradient estimates in the Lévy setting are studied in different ways in the literature. In [10, Theorem 1.1 and Remark 2.4] Hölder regularities are studied, where one looks for an improvement of the Hölder regularity caused by the transition group. In a way, this is opposite to our question. The result from the literature we contribute to is [39, Theorem 1.3] (see Remark 8.22 below). Finally, [33] investigates when $f(X_T)$ belongs to $\mathbb{D}_{1,2}$ or $(L_2, \mathbb{D}_{1,2})_{\theta,\infty}$ in dependence on $f \in \mathrm{H\ddot{o}l}^0_{\eta,\infty}(\mathbb{R})$ and properties of the underlying Lévy process X. In our article we look for L_∞ and BMO bounds for vector-valued gradient processes generated by an $f(X_T)$ when $f \in \mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})$, where we do not need and consider any Malliavin smoothness of $f(X_T)$ itself. Moreover, for a given $f(X_T)$ the fractional smoothness of the gradient process depends on the direction in which the gradient process is tested. So far, we do not see a way to exploit the results from [33] for our purpose, but it would be worthy to understand connections.

For this section we assume the following setting:

(1) $X = (X_t)_{t\in[0,T]}$ is a Lévy process with $\mu(\mathbb{R}) \in (0,\infty]$.
(2) ρ is a probability measure on $\mathcal{B}(\mathbb{R})$.

Let us start by formalizing the right-hand side of (8.4):

Definition 8.3. For an $F : [0,T) \times \mathbb{R} \to \mathbb{R}$, such that $x \mapsto F(t,x)$ is measurable for all $t \in [0,T)$, and $(t,x) \in [0,T) \times \mathbb{R}$ we define

$$D_\rho F(t,x) := \int_{\mathbb{R}\setminus\{0\}} \frac{F(t,x+z) - F(t,x)}{z} \rho(\mathrm{d}z) \quad \text{if} \quad \int_{\mathbb{R}\setminus\{0\}} \frac{|F(t,x+z) - F(t,x)|}{|z|} \rho(\mathrm{d}z) < \infty.$$

If additionally we have that $F(t,\cdot) \in C^1(\mathbb{R})$, then we let

$$\overline{D}_\rho F(t,x) := \rho(\{0\}) \frac{\partial F}{\partial x}(t,x) + \int_{\mathbb{R}\setminus\{0\}} \frac{F(t,x+z) - F(t,x)}{z} \rho(\mathrm{d}z).$$

One point of this definition is that the measure ρ is general. This allows us to capture different aspects: If ρ is as in Proposition 8.2, then we can study Galtchouk-Kunita-Watanabe projections, if ρ is a Dirac measure in $z \in \mathbb{R}\setminus\{0\}$, then we study the point-wise behaviour of $(F(t,x+z) - F(t,x))/z$. A general background is provided in Appendix D.3 in terms of a vector-valued gradient process associated to a functional $f(X_T)$.

We recall a class of functions that are of local bounded variation:

Definition 8.4. A Borel function $f : \mathbb{R} \to \mathbb{R}$ belongs to $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ provided that f is right-continuous and there are Borel measures μ^+ and μ^- on $\mathcal{B}(\mathbb{R})$, finite on each compact interval, and disjoint $S^+, S^- \in \mathcal{B}(\mathbb{R})$ with $S^+ \cup S^- = \mathbb{R}$ and $\mu^+(S^-) = \mu^-(S^+) = 0$, such that

$$f(b) - f(a) = \mu^+((a,b]) - \mu^-((a,b]) \quad \text{for all} \quad -\infty < a < b < \infty.$$

Furthermore, we let $|f'| := \mu^+ + \mu^-$ and, for a Borel function $g : \mathbb{R} \to \mathbb{R}$ with $\int_{\mathbb{R}} |g(x)||f'|(\mathrm{d}x) < \infty$,

$$\int_{\mathbb{R}} g(x) f'(\mathrm{d}x) := \int_{\mathbb{R}} g(x)\mu^+(\mathrm{d}x) - \int_{\mathbb{R}} g(x)\mu^-(\mathrm{d}x).$$

The pair of measures (μ^+, μ^-) is unique and we will identify f' with (μ^+, μ^-). The space $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ consists of functions that are of bounded variation of on each compact interval (cf. [37, Chapter 8]). The next definition is the key for what follows and defines two functionals to obtain $D_\rho F(t,x)$, the second term on the right-hand side of (8.4), from a given terminal condition f. The first functional simply rephrases $D_\rho F$, the second one uses some kind of partial integration.

[2]The integral with respect to $\rho(\mathrm{d}z)$ exists for $\omega \notin N_t$ and we omit $\rho(\{0\})(\partial F/\partial x)(t, X_t(\omega))$ if $\rho(\{0\}) = 0$.

Definition 8.5. (1) For $t \in [0,T)$ we define $\Gamma^0_{t,\rho} : \mathrm{Dom}(\Gamma^0_\rho) \to \mathbb{R}$ by

$$\mathrm{Dom}(\Gamma^0_\rho) := \Bigg\{ f \in \mathcal{D}_X \text{ and } \forall s \in [0,T)\ \forall\, 0 \leqslant \delta \leqslant s < T\ \forall x \in \mathbb{R} : \\ \mathbb{E}\int_{\mathbb{R}\setminus\{0\}} \left| \frac{F(s,x+X_\delta+z) - F(s,x+X_\delta)}{z} \right| \rho(\mathrm{d}z) < \infty \Bigg\},$$

$$\langle f, \Gamma^0_{t,\rho}\rangle := D_\rho F(t,0).$$

(2) For $t \in [0,T)$ we define the Borel function $\gamma_{t,\rho} : \mathbb{R} \to [0,\infty]$ and $\Gamma^1_{t,\rho} : \mathrm{Dom}(\Gamma^1_\rho) \to \mathbb{R}$ by

$$\gamma_{t,\rho}(v) := \int_{\mathbb{R}\setminus\{0\}} \frac{\mathbb{P}(X_{T-t} \in J(v;z))}{|z|} \rho(\mathrm{d}z) \quad \text{with} \quad J(v;z) := v + [-z^+, z^-),$$

$$\mathrm{Dom}(\Gamma^1_\rho) := \Bigg\{ f \in \mathcal{D}_X \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}) \text{ and } \forall\, 0 \leqslant \delta \leqslant s < T\ \forall x \in \mathbb{R} : \\ \mathbb{E}\int_{\mathbb{R}} \gamma_{s,\rho}(v - x - X_\delta)|f'|(\mathrm{d}v) < \infty \Bigg\},$$

$$\langle f, \Gamma^1_{t,\rho}\rangle = \langle f', \gamma_{t,\rho}\rangle := \int_{\mathbb{R}} \gamma_{t,\rho}(v)\, f'(\mathrm{d}v).$$

In Definition 8.5 we use L_1-conditions instead of L_2-conditions which is sufficient at this point. The L_1-conditions are chosen to guarantee a point-wise definition of $D_\rho F$ and the properties stated in Remark 8.6 below.

In Theorem 8.10 we prove $\int_{\mathbb{R}} \gamma_{t,\rho}(v)\mathrm{d}v = \rho(\mathbb{R}\setminus\{0\})$, $\mathrm{Dom}(\Gamma^1_\rho) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$, and that

$$\langle f', \gamma_{t,\rho}\rangle = \langle f, \Gamma^0_{t,\rho}\rangle \quad \text{for} \quad f \in \mathrm{Dom}(\Gamma^1_\rho).$$

If $D(\mathbb{R})$ is the test function space that consists of $f \in C^\infty(\mathbb{R})$ with compact support, then $D(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^1_\rho)$ (for $f \in D(\mathbb{R})$ we have $f'(\mathrm{d}v) = f'(v)\mathrm{d}v$ and $|f'|(\mathrm{d}v) = |f'(v)|\mathrm{d}v$, where f' on the right-hand sides is the classical derivative). If we consider $\gamma_{t,\rho} \in L_1(\mathbb{R})$ as distribution $\gamma_{t,\rho} \in D'(\mathbb{R})$ (see [36, Section 6.11]), then we have the interpretation

$$D\gamma_{t,\rho} = -\Gamma^0_{t,\rho}, \tag{8.5}$$

see [36, Section 6.12] and $\Gamma^0_{t,\rho}$ can be seen as distributional derivative of a distribution of L_1-type.

Before we continue, let us list some facts we exploit later:

Remark 8.6. For $f \in \mathrm{Dom}(\Gamma^0_\rho)$ the following holds:
(1) $D_\rho F(t,x) = \langle f(x+\cdot), \Gamma^0_{t,\rho}\rangle$.
(2) One has that $t \mapsto d(t) := \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \in [0,\infty]$ is non-decreasing.
(3) The process $(D_\rho F(t,X_t))_{t\in[0,T)}$ is a martingale.
(4) There exists a càdlàg modification $\varphi = (\varphi_t)_{t\in[0,T)}$ of $(D_\rho F(t,X_t))_{t\in[0,T)}$ such that

$$|\varphi_t| \leqslant d(t+) \quad \text{on} \quad [0,T)\times\Omega.$$

It will be useful to consider $\Gamma^0_{t,\rho}$ as linear functional on semi-normed spaces:

Definition 8.7. For $t \in [0,T)$ and a linear space $E \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ equipped with a semi-norm $|\cdot|_E$ we let $\|\Gamma^0_{t,\rho}\|_{E^*} := \inf c$, where the infimum is taken over all $c > 0$ such that

$$|\langle f, \Gamma^0_{t,\rho}\rangle| \leqslant c|f|_E \quad \text{for all} \quad f \in E.$$

In this article we aim for estimates of type

$$\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \le c_{(8.6)}(t)|f|_E \quad \text{for all} \quad f \in E. \tag{8.6}$$

If E contains only functions f such that $f(0) = 0$ (to have a norm $\|\cdot\|_E$ rather than a semi-norm $|\cdot|_E$ later) and are 'translation invariant' in the sense that $\|f\|_E = \|x \mapsto f(x_0+x) - f(x_0)\|_E$ for any $x_0 \in \mathbb{R}$, then the estimate (8.6) is equivalent to

$$|D_\rho F(t,0)| = |\langle f, \Gamma^0_{t,\rho}\rangle| \le c_{(8.6)}(t)\|f\|_E \quad \text{for all} \quad f \in E.$$

This is the reasoning for the definition of $\langle f, \Gamma^0_{t,\rho}\rangle$, i.e. for the estimates (8.6) one does not need to work with the Banach space $B_b(\mathbb{R})$. One application of the results of this section are the upper gradient estimates provided by Corollary 8.13 that can be seen as a counterpart to Theorem 6.6 proved on the Wiener space. To prove Corollary 8.13 we use the interpolation result Theorem 7.1 with end-point estimates derived by Theorem 8.9 and Theorem 8.12. As an application, inequality (8.15) of Corollary 8.13 allows for BMO-estimates of $(D_\rho F(t,X_t))_{t\in[0,T)}$ after applying our Riemann-Liouville operators to its càdlàg version by exploiting Theorem 8.11.

To start with, we introduce a variational quantity that is one key for us to obtain upper bounds for gradient processes:

Definition 8.8. For $\eta \in [0,1]$ and $s \in [0,T]$ we let

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)} := \inf_P \left\{ \int_{\mathbb{R}\setminus\{0\}} P(z)^{1-\eta} \rho(\mathrm{d}z) \right\} \in [0,\infty],$$

where the infimum is taken over all measurable $P : \mathbb{R}\setminus\{0\} \to [0,\infty)$ such that

$$\frac{\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} \leqslant P(z) \quad \text{for} \quad z \in \mathbb{R}\setminus\{0\}.$$

We use the potentials P to avoid a discussion about the measurability of the map $z \mapsto \|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}$ (which would not be necessary for us). We have the following special cases:

(1) $\|X_s\|_{\mathrm{TV}(\rho,1)} = \rho(\mathbb{R}\setminus\{0\}) < \infty$ for $s \in [0,T]$.
(2) $\|X_0\|_{\mathrm{TV}(\rho,\eta)} = 2^{1-\eta} \int_{\mathbb{R}\setminus\{0\}} |z|^{\eta-1} \rho(\mathrm{d}z)$ for $\eta \in [0,1]$.
(3) $\|X_s\|_{\mathrm{TV}(\delta_z,\eta)} = \left(\frac{\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} \right)^{1-\eta} < \infty$, $\eta \in [0,1]$, if δ_z is the Dirac measure in $z \in \mathbb{R}\setminus\{0\}$.

We will not use $\|X_0\|_{\mathrm{TV}(\rho,\eta)}$, whereas our idea is to use $\|X_s\|_{\mathrm{TV}(\rho,\eta)}$ for $s \in (0,T]$, where we exploit the behaviour of $\|\mathbb{P}_{z+X_s} - \mathbb{P}_{X_s}\|_{\mathrm{TV}}$. This enables us to obtain the correct blow-up of gradient processes when considering β-stable-like processes. Upper bounds for $\|X_s\|_{\mathrm{TV}(\delta_z,\eta)}$ can be found in the literature, see [39, Theorem 3.1], Theorem 8.9(2) is a variant for our setting.

In Theorem 8.9 and Theorem 8.10 below we provide basic properties of $\Gamma^0_{t,\rho}$ and $\Gamma^1_{t,\rho}$. We will use Theorem 8.9 to deduce upper and Theorem 8.10 to deduce lower bounds for our gradient processes. Moreover, Theorem 8.10 gives the interpretation (8.5) of $\Gamma^0_{t,\rho}$ and $\Gamma^1_{t,\rho}$ as distributions.

Theorem 8.9 (Properties of the functional $\Gamma^0_{t,\rho}$)**.** *Suppose that $\eta \in [0,1]$ and $(X_t)_{t\in[0,T]} \subseteq L_\eta$.*

(1) *If $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for $s \in (0,T]$, then $\mathrm{H\ddot{o}l}_\eta(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ and*

$$\left\|\Gamma^0_{t,\rho}\right\|_{(\mathrm{H\ddot{o}l}_\eta(\mathbb{R}))^*} \leqslant \|X_{T-t}\|_{\mathrm{TV}(\rho,\eta)}, \tag{8.7}$$

where $\mathrm{H\ddot{o}l}_\eta(\mathbb{R})$ is equipped with the semi-norm $|f|_{\eta,\infty} := \|f - f(0)\|_{\mathrm{H\ddot{o}l}^0_{\eta,\infty}(\mathbb{R})}$ if $\eta \in (0,1)$.

(2) *If $t \in [0,T)$ and X_{T-t} has a density $p_{T-t} \in C^1(\mathbb{R})$, then*

$$\|X_{T-t}\|_{\mathrm{TV}(\rho,\eta)} \leqslant \int_{\mathbb{R}\setminus\{0\}} \left(\min\left\{ \frac{2}{|z|}, \left\| \frac{\partial p_{T-t}}{\partial y} \right\|_{L_1(\mathbb{R})} \right\} \right)^{1-\eta} \rho(\mathrm{d}z).$$

In particular, if $\sigma > 0$, then $p_{T-t} \in C^1(\mathbb{R})$ with $\left\| \frac{\partial p_{T-t}}{\partial y} \right\|_{L_1(\mathbb{R})} \leqslant \sqrt{\frac{2}{\pi\sigma^2}}(T-t)^{-\frac{1}{2}}$.

Proof. (1) First we remark that $(X_t)_{t\in[0,T]} \subseteq L_\eta$ implies that $\mathrm{H\ddot{o}l}_\eta(\mathbb{R}) \subseteq \mathcal{D}_X$. Moreover, for fixed $z \in \mathbb{R}\setminus\{0\}$, $t \in [0,T)$, and $f \in \mathrm{H\ddot{o}l}_1(\mathbb{R})$ we obtain the estimate

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant |f|_1 \tag{8.8}$$

and, for $f \in B_b(\mathbb{R})$ and $x' \in \mathbb{R}$,

$$\begin{aligned} |F(t,x+z) - F(t,x)| &= \left| \int_{\mathbb{R}} (f(x+y) - f(x'))\mathbb{P}_{z+X_{T-t}}(\mathrm{d}y) - \int_{\mathbb{R}} (f(x+y) - f(x'))\mathbb{P}_{X_{T-t}}(\mathrm{d}y) \right| \\ &\leqslant \int_{\mathbb{R}} |f(x+y) - f(x')| \, |\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}|(\mathrm{d}y) \end{aligned}$$

$$\leqslant \|f - f(x')\|_{B_b(\mathbb{R})} \|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}.$$

Therefore,

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant c_{(8.9)} \frac{\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}}{|z|} \tag{8.9}$$

for $c_{(8.9)} := \|f\|_{C_b^0(\mathbb{R})}$ if $f \in C_b^0(\mathbb{R})$ (take $x' = 0$) and $c_{(8.9)} := |f|_0$ if $f \in \mathrm{Höl}_0(\mathbb{R})$ (take the supremum over $x' \in \mathbb{R}$ on the right-hand side). Moreover, real interpolation between (8.9) for $C_b^0(\mathbb{R})$ and (8.8) for $\mathrm{Höl}_1^0(\mathbb{R})$ (for fixed x and z) implies that

$$\left| \frac{F(t,x+z) - F(t,x)}{z} \right| \leqslant \|f\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})} \left[\frac{\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}}{|z|} \right]^{1-\eta} \tag{8.10}$$

for $\eta \in (0,1)$ by (2.2). From (8.9) and (8.8) we deduce $\mathrm{Höl}_\eta(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ and (8.7) for $\eta \in \{0,1\}$. If $\eta \in (0,1)$, then (8.10) implies $\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ and (8.7) with $\mathrm{Höl}_\eta(\mathbb{R})$ replaced by $\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})$. But if $f \in \mathrm{Höl}_\eta(\mathbb{R})$, then we replace f by $f_0 := f - f(0) \in \mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})$ and get (8.10) with constant $\|f - f(0)\|_{\mathrm{Höl}^0_{\eta,\infty}(\mathbb{R})}$. This concludes the proof of (1).

(2) We observe that

$$\begin{aligned}
\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}} = \|p_{T-t}(\cdot - z) - p_{T-t}\|_{L_1(\mathbb{R})} &= \int_{\mathbb{R}} \left| \int_{x-z}^{x} \frac{\partial p_{T-t}}{\partial y}(y)\mathrm{d}y \right| \mathrm{d}x \\
&\leqslant \mathrm{sign}(z) \int_{\mathbb{R}} \int_{x-z}^{x} \left| \frac{\partial p_{T-t}}{\partial y}(y) \right| \mathrm{d}y\mathrm{d}x \\
&= |z| \int_{\mathbb{R}} \left| \frac{\partial p_{T-t}}{\partial y}(y) \right| \mathrm{d}y. \qquad (8.11)
\end{aligned}$$

As we have $\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}} \leqslant 2$ as well, we obtain the first part of item (2). If $\sigma > 0$ and $s \in (0,T]$, then the density of X_s is given by $p_s(y) := \mathbb{E} p_{\sigma W_s}(y - J_s)$ where $p_{\sigma W_s}$ is the C^∞-density of σW_s and satisfies

$$\left\| \frac{\partial p_s}{\partial y} \right\|_{L_1(\mathbb{R})} = \left\| \mathbb{E} \frac{\partial p_{\sigma W_s}}{\partial y}(\cdot - J_s) \right\|_{L_1(\mathbb{R})} \leqslant \left\| \frac{\partial p_{\sigma W_s}}{\partial y} \right\|_{L_1(\mathbb{R})} = \sqrt{\frac{2}{\pi\sigma^2}} s^{-\frac{1}{2}}. \qquad \square$$

Theorem 8.10 (Properties of the functional $\Gamma^1_{t,\rho}$)**.** *Let* $t \in [0,T)$.

(1) *One has* $\int_{\mathbb{R}} \gamma_{t,\rho}(v)\mathrm{d}v = \rho(\mathbb{R}\backslash\{0\})$.

(2) *One has* $\mathrm{Dom}(\Gamma^1_\rho) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ *and for* $f \in \mathrm{Dom}(\Gamma^1_{t,\rho})$ *and* $x \in \mathbb{R}$ *that*

$$D_\rho F(t,x) = \langle f^x, \Gamma^1_{t,\rho} \rangle = \langle f^x, \Gamma^0_{t,\rho} \rangle \quad \text{if} \quad f^x(\cdot) := f(\cdot + x).$$

(3) *If* $q, r \in [1,\infty]$, X_{T-t} *has a density* $p_{T-t} \in L_r(\mathbb{R})$, *and* $s := \min\{r,q\}$, *then*

$$\|\gamma_{t,\rho}\|_{L_q(\mathbb{R})} \leqslant \|p_{T-t}\|_{L_s(\mathbb{R})} \int_{\mathbb{R}\backslash\{0\}} |z|^{\frac{1}{q} - \frac{1}{s}} \rho(\mathrm{d}z).$$

Proof. Recall the notation $J(v;z) = v + [-z^+, z^-)$. (1) follows from

$$\int_{\mathbb{R}} \gamma_{t,\rho}(v)\mathrm{d}v = \int_{\mathbb{R}\backslash\{0\}} \int_\Omega \left[\int_{\mathbb{R}} \mathbb{1}_{\{X_{T-t} \in J(v;z)\}} \frac{1}{|z|} \mathrm{d}v \right] \mathrm{d}\mathbb{P}\rho(\mathrm{d}z) = \int_{\mathbb{R}\backslash\{0\}} \int_\Omega \mathrm{d}\mathbb{P}\rho(\mathrm{d}z) = \rho(\mathbb{R}\backslash\{0\}).$$

(2) For $f \in \mathrm{Dom}(\Gamma^1_\rho)$ and $x \in \mathbb{R}$ we observe that

$$\begin{aligned}
\int_{\mathbb{R}\backslash\{0\}} \left| \frac{F(t,x+z) - F(t,x)}{z} \right| \rho(\mathrm{d}z) &\leqslant \int_{\mathbb{R}\backslash\{0\}} \left[\int_{\mathbb{R}} \left| \frac{f(x+z+y) - f(x+y)}{z} \right| \mathbb{P}_{X_{T-t}}(\mathrm{d}y) \right] \rho(\mathrm{d}z) \\
&\leqslant \int_{\mathbb{R}\backslash\{0\}} \int_{\mathbb{R}} \frac{1}{|z|} \int_{(x+y-z^-, x+y+z^+]} |f'|(\mathrm{d}v) \mathbb{P}_{X_{T-t}}(\mathrm{d}y) \rho(\mathrm{d}z) \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}\backslash\{0\}} \frac{1}{|z|} \int_{\mathbb{R}} \mathbb{1}_{(x+y-z^-, x+y+z^+]}(v) \mathbb{P}_{X_{T-t}}(\mathrm{d}y) \rho(\mathrm{d}z) \right] |f'|(\mathrm{d}v) \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}\backslash\{0\}} \frac{\mathbb{P}(X_{T-t} \in J(v-x;z))}{|z|} \rho(\mathrm{d}z) \right] |f'|(\mathrm{d}v)
\end{aligned}$$

$$= \int_{\mathbb{R}} \gamma_{t,\rho}(v-x)|f'|(\mathrm{d}v)$$

which implies $\mathrm{Dom}(\Gamma^1_\rho) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ and also enables us to compute, exactly along the previous computation,

$$\int_{\mathbb{R}\setminus\{0\}} \frac{F(t,x+z)-F(t,x)}{z}\rho(\mathrm{d}z) = \int_{\mathbb{R}} \left[\int_{\mathbb{R}\setminus\{0\}} \frac{\mathbb{P}(X_{T-t}\in J(v;z))}{|z|}\rho(\mathrm{d}z)\right](f^x)'(\mathrm{d}v).$$

(3) Let $z \neq 0$. Then the assertion follows from

$$\begin{aligned}
\|\mathbb{P}(X_{T-t}\in J(\cdot;z))\|_{L_q(\mathbb{R})} &= |z| \left\| \frac{1}{|z|}\int_{J(\cdot;z)} p_{T-t}(y)\mathrm{d}y \right\|_{L_q(\mathbb{R})} \\
&\leqslant |z|^{1-\frac{1}{s}} \left\| v \mapsto \left\| \mathbb{1}_{J(v;z)} p_{T-t} \right\|_{L_s(\mathbb{R})} \right\|_{L_q(\mathbb{R})} \\
&\leqslant |z|^{1-\frac{1}{s}} \left\| y \mapsto \left\| \mathbb{1}_{J(\cdot;z)}(y) p_{T-t}(y) \right\|_{L_q(\mathbb{R})} \right\|_{L_s(\mathbb{R})} \\
&= |z|^{1-\frac{1}{s}} \left\| y \mapsto \left\| \mathbb{1}_{J(\cdot;z)}(y) \right\|_{L_q(\mathbb{R})} p_{T-t}(y) \right\|_{L_s(\mathbb{R})} \\
&= |z|^{1-\frac{1}{s}+\frac{1}{q}} \left\| p_{T-t} \right\|_{L_s(\mathbb{R})},
\end{aligned}$$

where we use Hölder's inequality for the first inequality and (2.1) in the second one. □

We return to the Riemann-Liouville type operators and aim for correct upper bounds for (say) $\|\mathcal{I}^\alpha\varphi - \varphi_0\|_{\mathrm{BMO}_2([0,T))}$. *Point-wise* bounds for $\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}$, in the sense that $t\in[0,T)$ is *fixed*, will not yield to optimal results. Instead, we exploit integral bounds expressed by $|\!|\!| f |\!|\!|_{\rho,\alpha}$ below.

Theorem 8.11. *Assume that $\alpha > 0$, $f \in \mathrm{Dom}(\Gamma^0_\rho)$, and*

$$|\!|\!| f |\!|\!|^2_{\rho,\alpha} := \frac{2\alpha}{T^{2\alpha}}\int_0^T (T-t)^{2\alpha-1}\|D_\rho F(t,\cdot)\|^2_{B_b(\mathbb{R})}\mathrm{d}t < \infty,$$

and define

$$\varepsilon(a)^2 := \frac{2\alpha}{T^{2\alpha}}\int_a^T (T-t)^{2\alpha-1}\|D_\rho F(t,\cdot)\|^2_{B_b(\mathbb{R})}\mathrm{d}t \leqslant |\!|\!| f |\!|\!|^2_{\rho,\alpha}$$

so that $\varepsilon(a)\downarrow 0$ if $a\uparrow T$. For a càdlàg modification $\varphi = (\varphi_t)_{t\in[0,T)}$ of $(D_\rho F(t,X_t))_{t\in[0,T)}$ one has

(1) $\left(\frac{T-a}{T}\right)^\alpha \|D_\rho F(a,\cdot)\|_{B_b(\mathbb{R})} \leqslant \varepsilon(a)$ *for* $a\in[0,T)$,

(2) $\mathbb{E}^{\mathcal{F}_a}\left[|\mathcal{I}^\alpha_t\varphi - \mathcal{I}^\alpha_a\varphi|^2\right] \leqslant \varepsilon(a)^2$ *a.s. for* $0\leqslant a < t < T$,

(3) $\left\|(\mathcal{I}^\alpha_t\varphi - \mathcal{I}^\alpha_a\varphi)_{t\in[a,T)}\right\|_{\mathrm{BMO}_2([a,T))} \leqslant 3\varepsilon(a)$ *for* $a\in[0,T)$.

Proof. (1) follows from

$$\begin{aligned}
\frac{(T-a)^{2\alpha}}{2\alpha}\|D_\rho F(a,\cdot)\|^2_{B_b(\mathbb{R})} &= \int_a^T (T-t)^{2\alpha-1}\|D_\rho F(a,\cdot)\|^2_{B_b(\mathbb{R})}\mathrm{d}t \\
&\leqslant \int_a^T (T-t)^{2\alpha-1}\|D_\rho F(t,\cdot)\|^2_{B_b(\mathbb{R})}\mathrm{d}t = \frac{T^{2\alpha}}{2\alpha}\varepsilon(a)^2.
\end{aligned}$$

(2) We assume $B\in\mathcal{F}_a$ of positive measure and apply Proposition 3.8, formula (3.4), to get

$$\begin{aligned}
\int_B |\mathcal{I}^\alpha_t\varphi - \mathcal{I}^\alpha_a\varphi|^2\mathrm{d}\mathbb{P}_B &= 2\gamma T^{-2\alpha}\int_B\int_a^T (T-u)^{2\alpha-1}\left|\varphi_{u\wedge t}-\varphi_a\right|^2\mathrm{d}u\mathrm{d}\mathbb{P}_B \\
&\leqslant 2\alpha T^{-2\alpha}\int_B\int_a^T (T-u)^{2\alpha-1}\left|\varphi_{u\wedge t}\right|^2\mathrm{d}u\mathrm{d}\mathbb{P}_B \\
&\leqslant 2\alpha T^{-2\alpha}\int_a^T (T-u)^{2\alpha-1}\left\|D_\rho F(u,\cdot)\right\|^2_{B_b(\mathbb{R})}\mathrm{d}u \\
&= \varepsilon(a)^2.
\end{aligned}$$

(3) Because the $\mathrm{BMO}_2([a,T))$-norm is invariant when passing to càdlàg modifications, we may assume the bound from Remark 8.6(4) for φ and use Proposition 3.3(3) in order to get

$$|\Delta \mathcal{I}_t^\alpha \varphi| = \left(\frac{T-t}{T}\right)^\alpha |\Delta\varphi_t| \leqslant 2\varepsilon(t) \quad \text{on} \quad [0,T)\times\Omega.$$

The statement follows from item (2), Proposition A.4, and Proposition A.5(1) (applied to the time interval $[a,T)$). ☐

Theorem 8.12 (End point estimate)**.** *Let $X=(X_t)_{t\in[0,T]}$ be a Lévy process. If there are $\varepsilon\in(0,1)$ and $\beta\in(0,\infty]$ such that*

$$c_{(8.12)} := \sup_{n\in\mathbb{N}} 2^{\varepsilon n}\rho(\{2^{-n}\leqslant |z| < 2^{-n+1}\}) < \infty, \tag{8.12}$$

$$c_{(8.13)} := \sup_{s\in(0,T]}\ \sup_{z\in \operatorname{supp}(\rho)\setminus\{0\}} s^{\frac{1}{\beta}} \frac{\|\mathbb{P}_{z+X_s}-\mathbb{P}_{X_s}\|_{\mathrm{TV}}}{|z|} < \infty, \tag{8.13}$$

then, for $\eta\in[0,1-\varepsilon)$ there is a constant $c=c(\varepsilon,\beta,\eta,c_{(8.12)},c_{(8.13)})>0$ such that

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)} \leqslant c\, s^{\frac{\varepsilon+\eta-1}{\beta}} \quad \textit{for} \quad s\in(0,T].$$

Proof. With $A(z):=(2/|z|)^{1-\eta}$ and $\gamma:=\frac{1-\eta}{\beta}\in[0,\infty)$ we get

$$\|X_s\|_{\mathrm{TV}(\rho,\eta)} \leqslant \int_{\operatorname{supp}(\rho)\setminus\{0\}} \left(\min\left\{\frac{2}{|z|}, c_{(8.13)}s^{-\frac{1}{\beta}}\right\}\right)^{1-\eta}\rho(\mathrm{d}z) = \int_{\operatorname{supp}(\rho)\setminus\{0\}} \min\{A(z), c_{(8.13)}^{1-\eta}s^{-\gamma}\}\rho(\mathrm{d}z).$$

Moreover, for $\kappa:=\frac{\varepsilon}{1-\eta}\in(0,1)$, our assumption implies

$$\rho(\{2^n < A\leqslant 2^{n+1}\}) \leq c_{(8.14)}2^{-\kappa n} \quad \text{for } n\in\mathbb{N}_0 \tag{8.14}$$

with some $c_{(8.14)}=c(c_{(8.12)},\varepsilon)>0$. Then we use the first step of the proof of Theorem 7.1 and the relation $(\kappa-1)\gamma = \left(\frac{\varepsilon}{1-\eta}-1\right)\frac{1-\eta}{\beta} = \frac{\varepsilon+\eta-1}{\beta}$. ☐

Corollary 8.13. *Assume that $(X_t)_{t\in[0,T]}\subseteq L_1$ and either that*

(1) $\sigma>0$, $\beta=2$, or

(2) $\sigma=0$, $(\varepsilon,\beta)\in(0,1)\times(1,2)$, and that (8.12) and (8.13) hold.

Then one has for $\eta\in(0,1-\varepsilon)$, $\alpha:=\frac{1-(\varepsilon+\eta)}{\beta}\in\left(0,\frac{1}{\beta}\right)$, and $q\in[1,\infty]$ that $\text{Höl}_{\eta,q}^0(\mathbb{R})\subseteq \operatorname{Dom}(\Gamma_\rho^0)$ and

$$\left\| t\mapsto (T-t)^\alpha \|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})}\right\|_{L_q([0,T),\frac{\mathrm{d}t}{T-t})} \leqslant c_{(8.15)}^{(q)}\|f\|_{\text{Höl}_{\eta,q}^0(\mathbb{R})} \tag{8.15}$$

for $f\in \text{Höl}_{\eta,q}^0(\mathbb{R})$, where $c_{(8.15)}^{(q)}>0$ is a constant independent from f. In particular, for $q=2$ we obtain

$$|\!|\!| f |\!|\!|_{\rho,\alpha} \leqslant \frac{\sqrt{2\alpha}}{T^\alpha} c_{(8.15)}^{(2)}\|f\|_{\text{Höl}_{\eta,2}^0(\mathbb{R})}. \tag{8.16}$$

Proof. In case of (1) we have $\text{Höl}_{\eta,\infty}^0(\mathbb{R})\subseteq\operatorname{Dom}(\Gamma_\rho^0)$ for all $\eta\in(0,1)$ by Theorem 8.9. In case of (2) we have $\text{Höl}_{\eta,\infty}^0(\mathbb{R})\subseteq\operatorname{Dom}(\Gamma_\rho^0)$ for $\eta\in(0,1-\varepsilon)$ by Theorem 8.12 and Theorem 8.9. To interpolate we choose $0<\eta_0<\eta<\eta_1<1-\varepsilon$ and find a $\delta\in(0,1)$ with $\eta=(1-\delta)\eta_0+\delta\eta_1$. Then, by (8.10),

$$\begin{aligned}\sup_{x\in\mathbb{R}}\left|\frac{F(t,x+z)-F(t,x)}{z}\right| &\leqslant \|f\|_{\text{Höl}_{\eta_i,\infty}^0(\mathbb{R})}\left[\frac{\|\mathbb{P}_{z+X_{T-t}}-\mathbb{P}_{X_{T-t}}\|_{\mathrm{TV}}}{|z|}\right]^{1-\eta_i}\\ &\leqslant \|f\|_{\text{Höl}_{\eta_i,\infty}^0(\mathbb{R})}\min\{A(z), c_{(8.13)}^{1-\eta_i}(T-t)^{-\gamma_i}\}\end{aligned}$$

with $\gamma_i:=\frac{1-\eta_i}{\beta}$ and $A_i(z):=(2/|z|)^{1-\eta_i}$. Let $\kappa_i:=\frac{\varepsilon}{1-\eta_i}$. As in the proof of Theorem 8.12 we get

$$\sup_{n\in\mathbb{N}_0} 2^{\kappa_i n}\rho(\{2^n < A_i\leqslant 2^{n+1}\})<\infty.$$

Now the statement follows from Theorem 7.1, where we note that

$$(1-\delta)(1-\kappa_0)\gamma_0 + \delta(1-\kappa_1)\gamma_1 = \alpha,$$

and the reiteration theorem in the form of (2.5). □

Remark 8.14. The assumption on the existence of the density p_{T-t} of a Lévy process is a time dependent distributional property (see, e.g., [38, Ch.5]). In Theorem 8.9 and Theorem 8.10 we used $\|\partial p_{T-t}/\partial x\|_{L_1(\mathbb{R})}$ and $\|p_{T-t}\|_{L_s(\mathbb{R})}$. Results for $\partial p_t/\partial x$ and p_t for a Lévy process can be found, for example, in [29, 31, 39, 42].

8.4. **Lower bounds for the oscillation of gradient processes.** Theorem 8.19 and Theorem 8.20 are the main results of this section. Their background is Proposition 8.2 where we compute the gradient of the Galtchouk-Kunita-Watanabe projection. Theorem 8.19 proves the maximal oscillation of these gradients and Theorem 8.20 determines the quantitative behaviour of the maximal oscillation as a counterpart to Corollary 8.13.

To handle the oscillation we exploit the supports of the laws $\mathbb{P}_{X_t}$ and transform the Lévy process $(X_t)_{t\in[0,T]}$ into the process $(Y_t)_{t\in[0,T]}$ below which has independent and stationary increments as well. The statements Theorem 8.16, Example 8.17, and Example 8.18, are formulated for the Y-process, before we return to the X-process. Let us start with the basic setting of this section:

Assumption 8.15. (1) In the notation of Example 5.6 we use $\operatorname{supp}(X_t) = Q + \ell t$, $t \in (0,T]$, and

$$Y_t = (X_t - \ell t)\mathbb{1}_{\{X_t \in \operatorname{supp}(X_t)\}} \quad \text{for } t \in [0,T].$$

(2) The function $H\colon [0,T)\times Q \to \mathbb{R}$ is Y-*consistent*, which means
  (a) $H(t,\cdot)$ is continuous on Q for all $t\in[0,T)$,
  (b) $\mathbb{E}|H(t,y+Y_{t-s})| < \infty$ for all $0 \leqslant s \leqslant t < T$ and $y \in Q$,
  (c) $\mathbb{E}H(t,y+Y_{t-s}) = H(s,y)$ for all $0 \leqslant s \leqslant t < T$ and $y \in Q$.
(3) ρ is a probability measure on $\mathcal{B}(\mathbb{R})$.

The reason for this definition is the following statement:

Theorem 8.16. *Let H be Y-consistent and $\varphi_t := H(t,Y_t)$, $t\in[0,T)$. Then $\varphi = (\varphi_t)_{t\in[0,T)}$ is a martingale of maximal oscillation with constant 2 in the sense of Definition 5.1. Moreover, if for all $t\in[0,T)$ there is an $\bar{t}\in(t,T)$ such that $H(t,Y_{\bar{t}})\in L_2$, then the following assertions are equivalent:*

(1) $\inf_{t\in(0,T)} \underline{\operatorname{Osc}}_t(\varphi) = 0$.
(2) $\varphi_t = \varphi_0$ *a.s. for all $t\in(0,T)$.*

Item 2 of Theorem 8.16 implies a *forward uniqueness*: If there is an $s\in(0,T)$ such that $\varphi_0 = \varphi_s$ a.s., then the martingale is constant a.s.

Proof of Theorem 8.16. The martingale property follows by the definition and the maximal oscillation with constant 2 follows from Example 5.6. Regarding the equivalence we only need to show (1)⇒(2). For $0<s<t<T$, $y_1', y_2' \in Q$ and $\omega \in (Y_t - Y_s)^{-1}(Q)$ we obtain that

$$\begin{aligned}
\|\varphi_t - \varphi_s\|_{L_\infty} &= \sup_{y,y'\in Q} |H(t,y+y') - H(s,y')| \\
&\geqslant |H(t,y_1' + (Y_t - Y_s)(\omega) + y_2' + Y_s(\omega)) - H(s,y_2' + Y_s(\omega))| \\
&= |H(t,y_1' + y_2' + Y_t(\omega)) - H(s,y_2' + Y_s(\omega))|,
\end{aligned}$$

where the first inequality comes from $\varphi_t - \varphi_s = H(t, Y_t - Y_s + Y_s) - H(s,Y_s)$, $\operatorname{supp}(Y_t - Y_s, Y_s) = Q\times Q$, and from the continuity of $Q\times Q \ni (y,y') \mapsto H(t,y+y') - H(s,y')$. This implies

$$\begin{aligned}
\|\varphi_t - \varphi_s\|_{L_\infty} \geqslant \sup_{y,y'\in Q} |\mathbb{E}H(t,y+y'+Y_t) - \mathbb{E}H(s,y'+Y_s)| &= \sup_{y,y'\in Q} |H(0,y+y') - H(0,y')| \\
&\geqslant \sup_{y\in Q} |H(0,y) - H(0,0)|.
\end{aligned}$$

For $s = 0$ we use the same idea with $y' = y_2' = 0$ to get $\|\varphi_t - \varphi_0\|_{L_\infty} \geqslant \sup_{y\in Q} |H(0,y) - H(0,0)|$. So (1) yields to $C := H(0,0) = H(0,y)$ for all $y \in Q$. Fix $0 \leqslant t < \bar{t} < T$ as in our assumption. According to Lemma D.1, we have a chaos expansion

$$H(t, Y_{\bar{t}}) = \mathbb{E}H(t, Y_{\bar{t}}) + \sum_{n=1}^{\infty} I_n\left(\tilde{h}_n \mathbb{1}_{[0,\bar{t}]}^{\otimes n}\right)$$

with $\tilde{h}_n \in L_2(\mu^{\otimes n})$. Let $\widetilde{Y}$ be an independent copy of Y with the corresponding expectation $\widetilde{\mathbb{E}}$. For $\Delta t := \bar{t} - t > 0$ this implies $\mathbb{E}^{\mathcal{F}_{\Delta t}}[H(t, Y_{\bar{t}})] \overset{a.s.}{=} \widetilde{\mathbb{E}}H(t, Y_{\Delta t} + \widetilde{Y}_t) = H(0, Y_{\Delta t}) = C$ and

$$C = \mathbb{E}^{\mathcal{F}_{\Delta t}}[H(t, Y_{\bar{t}})] = \mathbb{E}H(t, Y_{\bar{t}}) + \sum_{n=1}^{\infty} I_n\left(\tilde{h}_n \mathbb{1}_{[0,\Delta t]}^{\otimes n}\right) \text{ a.s.}$$

Therefore, $\tilde{h}_n = 0$ in $L_2(\mu^{\otimes n})$ for all $n \geqslant 1$, which yields $H(t, Y_{\bar{t}}) = C$ a.s. Since $\operatorname{supp}(Y_{\bar{t}}) = Q = \operatorname{supp}(Y_t)$, together with the continuity of $H(t,\cdot)$ on Q, we derive that $H(t,y) = C$ for all $y \in Q$. Therefore $\varphi_t = H(t, Y_t) = C$ a.s. □

The next two results provide the fundamental examples of Y-consistent functions:

Example 8.17. We assume

(1) that $k\colon Q \to \mathbb{R}$ is a Borel function with $\mathbb{E}|k(y + Y_s)| < \infty$ for $(s,y) \in [0,T] \times Q$ and that $K\colon [0,T) \times Q \to \mathbb{R}$ with $K(t,y) := \mathbb{E}k(y + Y_{T-t})$ satisfies
$$\mathbb{E}\int_{Q\setminus\{0\}} \left| \frac{K(t, y + Y_\delta + z) - K(t, y + Y_\delta)}{z} \right| \rho(\mathrm{d}z) < \infty \quad \text{for} \quad 0 \leqslant \delta \leqslant t < T,$$
(2) that $y \mapsto K(t,y)$ is continuous on Q for $t \in [0,T)$,
(3) that for all $(t,y) \in [0,T) \times Q$ there is an $\varepsilon > 0$ such that the family of functions $z \mapsto \frac{K(t,y'+z) - K(t,y')}{z}$, indexed by $y' \in Q$ with $|y - y'| < \varepsilon$, is uniformly integrable on $(Q\setminus\{0\}, \rho)$.

Then we obtain a Y-consistent function by

$$H(t,y) := \int_{Q\setminus\{0\}} \frac{K(t, y+z) - K(t,y)}{z} \rho(\mathrm{d}z) \quad \text{for} \quad (t,y) \in [0,T) \times Q.$$

Proof. (2b) Taking $\delta = 0$ in assumption (1), we see that $H(t,y)$ is well-defined, and for $\delta := t - s$ we obtain that

$$\mathbb{E}|H(t, y + Y_{t-s})| \leqslant \mathbb{E}\int_{Q\setminus\{0\}} \left| \frac{K(t, y + Y_{t-s} + z) - K(t, y + Y_{t-s})}{z} \right| \rho(\mathrm{d}z) < \infty. \tag{8.17}$$

(2c) Because of (8.17) we can apply Fubini's theorem to get

$$\begin{aligned}
\mathbb{E}H(t, y + Y_{t-s}) &= \mathbb{E}\int_{Q\setminus\{0\}} \frac{K(t, y + Y_{t-s} + z) - K(t, y + Y_{t-s})}{z} \rho(\mathrm{d}z) \\
&= \int_{Q\setminus\{0\}} \frac{\mathbb{E}K(t, y + Y_{t-s} + z) - \mathbb{E}K(t, y + Y_{t-s})}{z} \rho(\mathrm{d}z) \\
&= \int_{Q\setminus\{0\}} \frac{K(s, y + z) - K(s,y)}{z} \rho(\mathrm{d}z) \\
&= H(s,y)
\end{aligned}$$

where we use $\mathbb{E}K(t, y + Y_{t-s}) = K(s,y)$. (2a) If we have $y_n, y \in Q$ with $y_n \to y$, then we take $\varepsilon = \varepsilon(t,y) > 0$ from assumption (3) and obtain $\lim_n H(t, y_n) = H(t,y)$ by the uniform integrability imposed in (3) and assumption (2). □

Example 8.18. Let $\sigma > 0$. Then $Q = \mathbb{R}$ and the following holds:

(1) If $k\colon \mathbb{R} \to \mathbb{R}$ is a Borel function with $\mathbb{E}|k(Y_T)|^q < \infty$ for some $q \in (1,\infty)$, then $\mathbb{E}|k(y + Y_{T-t})| < \infty$ for $(t,y) \in [0,T] \times \mathbb{R}$. If $K(t,y) := \mathbb{E}k(y + Y_{T-t})$ on $[0,T] \times \mathbb{R}$, then $K(t,\cdot) \in C^\infty(\mathbb{R})$ for $t \in [0,T)$ and we obtain a Y-consistent function $H : [0,T) \times \mathbb{R} \to \mathbb{R}$ by
$$H(t,y) := \frac{\partial K}{\partial y}(t,y) \quad \text{with} \quad H(t,y) = \frac{1}{\sigma}\mathbb{E}\left[k(y + Y_{T-t})\frac{W_{T-t}}{T-t}\right].$$

(2) If $k \in \text{Höl}_\eta(\mathbb{R})$ for some $\eta \in [0,1]$ (and $\mathbb{E}|k(Y_T)|^q < \infty$ as above if $\eta \in (0,1]$), then

$$\|H(t,\cdot)\|_{B_b(\mathbb{R})} \leqslant |k|_\eta \sigma^{\eta-1}(T-t)^{\frac{\eta-1}{2}} \int_{\mathbb{R}} |x|^{\eta+1} \mathrm{e}^{-\frac{x^2}{2}} \frac{\mathrm{d}x}{\sqrt{2\pi}}. \tag{8.18}$$

Proof. $Q = \mathbb{R}$ follows from [38, Theorem 24.10].

(1) Let $k \geqslant 0$, $J := Y - \sigma W$, and fix $t \in [0,T)$.

(a) Since $\mathbb{E}|k(\sigma W_T + J_t + (J_T - J_t))|^q = \mathbb{E}|k(Y_T)|^q < \infty$, independence and Fubini's theorem yield to $\mathbb{E}|k(\sigma W_T + a_t + (J_T - J_t))|^q < \infty$ for some $a_t \in \mathbb{R}$. If

$$N^{(t)} := \{x \in \mathbb{R} : \mathbb{E}|k(\sigma x + a_t + (J_T - J_t))|^q = \infty\},$$

then $N^{(t)}$ is a Borel set of Lebesgue measure zero. We define

$$f^{(t)}(x) := \mathbb{1}_{N_t^c}(x)\ \mathbb{E}k(\sigma x + a_t + (J_T - J_t))$$

so that $\mathbb{E}|f^{(t)}(W_T)|^q < \infty$. Now we can apply [19, Lemma A.2] to $f^{(t)}$ and get for $(s,x) \in [0,T) \times \mathbb{R}$ and $F^{(t)}(s,x) := \mathbb{E}f^{(t)}(x + W_{T-s})$ that

$$F^{(t)}(s,\cdot) \in C^\infty(\mathbb{R}) \quad \text{and} \quad \frac{\partial F^{(t)}}{\partial x}(s,x) = \mathbb{E}f^{(t)}(x + W_{T-s})\frac{W_{T-s}}{T-s}.$$

(b) Because $k \geqslant 0$, $N^{(t)}$ has Lebesgue measure zero, and $T - t > 0$, we verify by Fubini's theorem (regardless of the finiteness of the integrals) that

$$K(t,y) = F^{(t)}\left(t, \frac{y - a_t}{\sigma}\right) < \infty \quad \text{so that} \quad K(t,\cdot) \in C^\infty(\mathbb{R}).$$

(c) We choose $\tilde{q} \in (1,q)$ so that $\mathbb{E}k(y + Y_T - Y_s)|W_T - W_t| \leqslant c_{\tilde{q},T-t}\|k(y + Y_T - Y_s)\|_{L_{\tilde{q}}} < \infty$ where the finiteness of the last term is obtained as in (a-b) by starting with the function $y \mapsto k(y)^{\tilde{q}}$. This moment estimate enables us to apply Fubini's theorem in the sequel.

(d) Using [19, Lemma A.2] we deduce that

$$\begin{aligned}\sigma\frac{\partial K}{\partial y}(t,y) = \frac{\partial F^{(t)}}{\partial x}\left(t, \frac{y-a_t}{\sigma}\right) &= \mathbb{E}f^{(t)}\left(\frac{y-a_t}{\sigma} + W_{T-t}\right)\frac{W_{T-t}}{T-t} \\ &= \mathbb{E}(\widetilde{\mathbb{E}}k(y + \sigma W_{T-t} + \widetilde{J}_{T-t}))\frac{W_{T-t}}{T-t} = \mathbb{E}k(y + Y_{T-t})\frac{W_{T-t}}{T-t}\end{aligned}$$

(e) To check $\mathbb{E}\frac{\partial K}{\partial y}(t, y + Y_{t-s}) = \frac{\partial K}{\partial y}(s,y)$ for $s \in [0,t]$ we have to verify

$$\mathbb{E}\left[k(y + Y_T - Y_s)\frac{W_T - W_t}{T-t}\right] = \mathbb{E}\left[k(y + Y_T - Y_s)\frac{W_T - W_s}{T-s}\right].$$

As $\frac{W_T - W_t}{T-t} - \frac{W_T - W_s}{T-s}$ is of mean zero and independent of $Y_T - Y_s$, the last equality is true. To conclude the proof of (1) we remove the assumption $k \geqslant 0$ by considering the positive and negative parts separately.

(2) Now we additionally assume that $k \in \text{Höl}_\eta(\mathbb{R})$. Assume that $t \in [0,T)$ and $y = \sigma x + a_t$ with $x \notin N^{(t)}$ and $N^{(t)}$ defined as in step (a). Then $\mathbb{E}|k(y + J_{T-t})| = \mathbb{E}|k(\sigma x + a_t + J_{T-t})| < \infty$ and

$$\begin{aligned}\left|\frac{\partial K}{\partial y}(t,y)\right| &= \left|\frac{1}{\sigma}\mathbb{E}\left[k(y + Y_{T-t})\frac{W_{T-t}}{T-t}\right]\right| = \frac{1}{\sigma}\mathbb{E}\left|\left[(k(y + Y_{T-t}) - k(y + J_{T-t}))\frac{W_{T-t}}{T-t}\right]\right| \\ &\leqslant \frac{|k|_\eta}{\sigma}\mathbb{E}\left[|\sigma W_{T-t}|^\eta \frac{|W_{T-t}|}{T-t}\right] = |k|_\eta \sigma^{\eta-1}(T-t)^{\frac{\eta-1}{2}}\mathbb{E}|g|^{\eta+1}.\end{aligned}$$

Because $\lambda(N^{(t)}) = 0$ and $y \mapsto \frac{\partial K}{\partial y}(t,y)$ continuous, the estimate is true for all $y \in \mathbb{R}$. □

Now we are in a position to return to the setting of Proposition 8.2:

Theorem 8.19 (Maximal oscillation)**.** *Suppose that*

(a) *the Lévy process $(X_t)_{t\in[0,T]}$ is an L_2-martingale and $\rho := \mu/\mu(\mathbb{R})$,*
(b) *$\eta \in [0,1]$ and $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for all $s \in (0,T]$ if $\eta \in [0,1)$,*
(c) *$f \in \text{Höl}_\eta(\mathbb{R})$, where we additionally assume that $y \mapsto f(y + \ell T)$ is continuous on Q if $\eta = \sigma = 0$.*

Then $f \in \mathrm{Dom}(\Gamma_\rho^0)$ and, additionally, $F(t,\cdot) \in C^\infty(\mathbb{R})$ for $t \in [0,T)$ if $\sigma > 0$. Letting $\varphi_t := \overline{D}_\rho F(t, X_t)$ for $t \in [0,T)$, the following holds:

(1) $\|\varphi_t\|_{L_\infty} = \sup_{x\in\operatorname{supp}(X_t)} |\overline{D}_\rho F(t,x)|$ *for* $t \in [0,T)$.
(2) $(\varphi_t)_{t\in[0,T)}$ *is an* L_2*-martingale of maximal oscillation with constant 2.*
(3) *Unless* $\varphi_t = \varphi_0$ *a.s. for all* $t \in [0,T)$*, one has* $\inf_{t\in(0,T)} \underline{\mathrm{Osc}}_t(\varphi) > 0$.

Proof. Theorem 8.9 implies that $f \in \operatorname{Dom}(\Gamma^0_\rho)$. Now we let $k(y) := f(y+\ell T)$ and $K(t,y) := \mathbb{E}k(y+Y_{T-t})$ for $(t,y) \in [0,T)\times Q$, so that

$$K(t,y) = F(t, y+\ell t) \quad \text{for} \quad (t,y) \in [0,T)\times Q. \tag{8.19}$$

Let

$$H(t,y) := \rho(\{0\})\frac{\partial K}{\partial y}(t,y) + \int_{Q\setminus\{0\}} \frac{K(t,y+z)-K(t,y)}{z}\rho(\mathrm{d}z) \quad \text{for} \quad (t,y)\in[0,T)\times Q.$$

(a) H is Y-consistent: By Example 8.18 and $\mathbb{E}|k(Y_T)|^2 < \infty$ the first term $\rho(\{0\})\frac{\partial K}{\partial y}$ is well-defined and Y-consistent given $\rho(\{0\}) > 0$. For the second term we verify the assumptions of Example 8.17:

Assumption (1) follows by (8.19) and $f \in \operatorname{Dom}(\Gamma^0_\rho)$.

Assumption (2) follows by Example 8.18 if $\sigma > 0$. If $\sigma = 0$, then $k : Q \to \mathbb{R}$ is continuous by assumption. Then we use $k \in \text{Höl}_\eta(\mathbb{R})$ and $Y_{T-t} \in L_2$ to deduce the uniform integrability of $(k(y_n + Y_{T-t}))_{n\in\mathbb{N}}$ if $y_n \to y$ in Q which implies the continuity of $K(t,\cdot)$ on Q for $t \in [0,T)$.

Assumption (3) follows from the proof of Theorem 8.9(1) that gives

$$\left|\frac{K(t,y+z)-K(t,y)}{z}\right| \leqslant c_\eta(f)\|\mathbb{P}_{z+X_{T-t}} - \mathbb{P}_{X_{T-t}}\|^{1-\eta}_{\mathrm{TV}}|z|^{\eta-1}. \tag{8.20}$$

(b) We have $\sup_{y\in Q}|H(t,y)| < \infty$ for all $t\in[0,T)$ because of (8.18) and (8.20).

Now assertion (1) follows from the continuity of $H(t,\cdot)$ on Q, which implies the continuity of $\overline{D}_\rho F(t,\cdot)$ on $\operatorname{supp}(X_t)$. Assertions (2) and (3) follow from Theorem 8.16 and again by $\overline{D}_\rho F(t,x) = H(t, x-\ell t)$ for $x\in\operatorname{supp}(X_t)$ ($\operatorname{supp}(X_0) = \{0\} \subseteq Q$) which implies $\overline{D}_\rho F(t,X_t) = H(t,Y_t)$ a.s. □

Now we provide the corresponding lower bounds for Corollary 8.13. The conditions (8.21) and (8.22) are a counterpart to (8.12) and (8.13) assumed in Corollary 8.13.

Theorem 8.20 (Size of maximal oscillation)**.** *Suppose that*
(a) *the Lévy process* $(X_t)_{t\in[0,T]}$ *is an* L_2*-martingale and* $\rho := \mu/\mu(\mathbb{R})$*,*
(b) $\eta\in[0,1)$ *and* $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ *for all* $s\in(0,T]$*,*
(c) $f_\eta : \mathbb{R}\to\mathbb{R} \in \text{Höl}_\eta(\mathbb{R})$ *is given by* $f_\eta(x) := (x\vee 0)^\eta$ *if* $\eta\in(0,1)$ *and* $f_0(x) := \mathbb{1}_{[0,\infty)}(x)$.

If $F_\eta(t,x) := \mathbb{E}f_\eta(x+X_{T-t})$ *for* $(t,x)\in[0,T)\times\mathbb{R}$*, then one has*
(1) $\inf_{t\in[0,T)}(T-t)^{1-\frac{1+\eta}{2}}\frac{\partial F_\eta}{\partial x}(t,0) > 0$ *if* $\sigma > 0$*,*
(2) $\inf_{t\in[0,T)}(T-t)^{1-\frac{1+\eta}{\beta}} D_\rho F_\eta(t,0) > 0$ *if* $\sigma = 0$ *and* $\beta\in[1+\eta,2)$*, and if*

$$\rho([-\varepsilon,\varepsilon]) \geqslant c_{(8.21)}\varepsilon^{2-\beta} \quad \textit{for} \quad \varepsilon\in(0,\varepsilon_{(8.21)}], \tag{8.21}$$

$$\inf_{|v|\vee|z|\leqslant \tilde{c}_{(8.22)}s^{\frac{1}{\beta}},\, z\neq 0} \frac{\mathbb{P}(X_s\in J(v;z))}{|z|} \geqslant c_{(8.22)}s^{-\frac{1}{\beta}} \quad \textit{for} \quad s\in(0,T], \tag{8.22}$$

for some constants $c_{(8.21)}$, $\varepsilon_{(8.21)}$, $c_{(8.22)}$, $\tilde{c}_{(8.22)} > 0$ *and where* $J(v;z) = v + [-z^+, z^-)$.

Proof. (1) For $s\in(0,T]$ we let $p_s = p_{\sigma W_s} * \mathbb{P}_{J_s}$ be the continuous density of the law of X_s (see Theorem 8.9(2)). Then we have

$$p_s(x) = \frac{1}{\sigma\sqrt{2\pi s}}\int_{\mathbb{R}} \mathrm{e}^{-\frac{(x-z)^2}{2\sigma^2 s}}\,\mathbb{P}_{J_s}(\mathrm{d}z) \geqslant \frac{1}{\sigma\sqrt{2\pi s}}\,\mathrm{e}^{-\frac{x^2}{\sigma^2 s}}\int_{\mathbb{R}}\mathrm{e}^{-\frac{z^2}{\sigma^2 s}}\,\mathbb{P}_{J_s}(\mathrm{d}z) \geqslant c_{(8.23)}s^{-\frac{1}{2}}\,\mathrm{e}^{-\frac{x^2}{\sigma^2 s}} \tag{8.23}$$

with $c_{(8.23)} := (\sigma\sqrt{2\pi})^{-1}\,\mathrm{e}^{-\sigma^{-2}\int_{\mathbb{R}} z^2\nu(\mathrm{d}z)}$ because by Jensen's inequality,

$$\int_{\mathbb{R}}\mathrm{e}^{-\frac{z^2}{\sigma^2 s}}\,\mathbb{P}_{J_s}(\mathrm{d}z) \geqslant \mathrm{e}^{-\sigma^{-2}s^{-1}\mathbb{E}J_s^2} = \mathrm{e}^{-\sigma^{-2}\int_{\mathbb{R}} z^2\nu(\mathrm{d}z)} > 0.$$

Moreover, for $x = 0$ and $\varepsilon > 0$ we have

$$\frac{\partial F_\eta}{\partial x}(t,0) = \eta\int_0^\infty v^{\eta-1}p_{T-t}(v)\mathrm{d}v \geqslant \left[\inf_{v\in[0,\varepsilon]} p_{T-t}(v)\right]\varepsilon^\eta$$

for $\eta \in (0,1)$, where the first equality follows by a direct computation, and (1) follows with $\varepsilon := \sqrt{T-t}$. If $\eta = 0$, then $\frac{\partial F_\eta}{\partial x}(t,0) = p_{T-t}(0) \geqslant c_{(8.23)}(T-t)^{-\frac{1}{2}}$.

(2) Let $\varepsilon = d(T-t)^{\frac{1}{\beta}} := \min\{\tilde{c}_{(8.22)}, \varepsilon_{(8.21)}T^{-\frac{1}{\beta}}\}(T-t)^{\frac{1}{\beta}}$. We observe that $\frac{F_\eta(t,x+z)-F_\eta(t,x)}{z} \geqslant 0$ for all $z \in \mathbb{R}\backslash\{0\}$ and that f' in the sense of measure is a non-negative measure. For this reason we can use the proof of item (2) of Theorem 8.10 *without checking integrability assumptions to get*

$$\begin{aligned}\int_{\mathbb{R}\backslash\{0\}} \frac{F_\eta(t,z)-F_\eta(t,0)}{z}\rho(\mathrm{d}z) &= \int_{\mathbb{R}}\int_{\mathbb{R}\backslash\{0\}} \frac{\mathbb{P}(X_{T-t}\in J(v;z))}{|z|}\rho(\mathrm{d}z)f'_\eta(\mathrm{d}v)\\ &\geqslant \int_{|v|\leqslant\varepsilon}\int_{0<|z|\leqslant\varepsilon} \frac{\mathbb{P}(X_{T-t}\in J(v;z))}{|z|}\rho(\mathrm{d}z)f'_\eta(\mathrm{d}v)\\ &\geqslant \inf_{0<|z|\leqslant\varepsilon,|v|\leqslant\varepsilon} \frac{\mathbb{P}(X_{T-t}\in J(v;z))}{|z|}\varepsilon^\eta\rho([-\varepsilon,\varepsilon])\\ &\geqslant c_{(8.22)}(T-t)^{-\frac{1}{\beta}}(d(T-t)^{\frac{1}{\beta}})^\eta c_{(8.21)}(d(T-t)^{\frac{1}{\beta}})^{2-\beta}\\ &= d'(T-t)^{\frac{1+\eta}{\beta}-1}.\end{aligned}$$

□

8.5. **Sharpness of the results - β-stable-like processes.** In this section we assume a Lévy process $X = (X_t)_{t\in[0,T]}$ with $\sigma = 0$, which is an L_2-martingale, and $\beta \in (1,2)$ such that the Lévy measure satisfies $\nu(\mathrm{d}z) = p_\nu(z)\mathrm{d}z$, where p_ν is symmetric and

$$0 < \liminf_{|z|\to 0} |z|^{1+\beta}p_\nu(z) \leqslant \limsup_{|z|\to 0}|z|^{1+\beta}p_\nu(z) < \infty. \tag{8.24}$$

We consider a functional $D \in L_2(\mathbb{R},\mu)$ with $D \geqslant 0$ and $\int_{\mathbb{R}} D^2\mathrm{d}\mu > 0$, and set

$$\mathrm{d}\rho := \frac{1}{\int_{\mathbb{R}} D\mathrm{d}\mu}D\mathrm{d}\mu.$$

Given $\varepsilon \in (0,1)$, the small-ball assumption (8.12) on the functional $D : L_2(\mathbb{R}\backslash\{0\},\mu) \to \mathbb{R}$ reads as

$$\left(\int_{\mathbb{R}} D\mathrm{d}\mu\right)\sup_{n\in\mathbb{N}} 2^{\varepsilon n}\rho(\{2^{-n}\leqslant|z|<2^{-n+1}\}) = \sup_{n\in\mathbb{N}} 2^{\varepsilon n}\int_{[2^{-n},2^{-n+1})} D\mathrm{d}\mu < \infty. \tag{8.25}$$

Given $f \in \mathcal{D}_X \cap L_2(\mathbb{R},\mathbb{P}_{X_T})$, we also discuss the Riemann approximation of the stochastic integral

$$\int_{(0,T)} \varphi_{t-}(f,\rho)\mathrm{d}X_t^D$$

that represents by Proposition 8.2 the Galtchouk-Kunita-Watanabe projection of $f(X_T)$ on $I(X^D)$ up to a factor. The corresponding error process with respect to the time-net $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$ is

$$E_t(f;\tau,D) := \int_{(0,t]}\varphi_{s-}(f,\rho)\mathrm{d}X_s^D - \sum_{i=1}^n \varphi_{t_{i-1}-}(f,\rho)(X^D_{t_i\wedge t} - X^D_{t_{i-1}\wedge t}), \quad t\in[0,T).$$

Theorem 8.21. *Let $\eta \in (0,1-\varepsilon)$, $\alpha := \frac{1-(\varepsilon+\eta)}{\beta}$, $\theta := 1-2\alpha$, and assume that the functional D satisfies the ε-small ball property (8.25). Then $\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ and the following holds:*

<u>UPPER BOUNDS:</u> *For $f \in \mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})$ and the parameters $\Theta := (\beta,\nu,\varepsilon,D,\eta,T)$ the following holds:*

(1) *There is a $c_{(8.26)} = c(\Theta) > 0$ such that one has*

$$\sup_{t\in[0,T)}(T-t)^\alpha\|\varphi_t(f,\rho)\|_{L_\infty} + \|\mathcal{I}^\alpha\varphi(f,\rho)-\varphi_0(f,\rho)\|_{\mathrm{BMO}_2([0,T))} \leqslant c_{(8.26)}\|f\|_{\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})}, \tag{8.26}$$

$$\lim_{a\uparrow T}\left\|(\mathcal{I}^\alpha_t\varphi(f,\rho)-\mathcal{I}^\alpha_a\varphi(f,\rho))_{t\in[a,T)}\right\|_{\mathrm{BMO}_2([a,T))} = 0. \tag{8.27}$$

(2) *There is a $c_{(8.28)} = c(\Theta) > 0$ such that one has*

$$\|E(f;\tau,D)\|_{\mathrm{bmo}_2([0,T))} \leqslant c_{(8.28)}\sqrt{\|\tau\|_\theta}\;\|f\|_{\mathrm{H\ddot{o}l}^0_{\eta,2}(\mathbb{R})}. \tag{8.28}$$

(3) *$\varphi(f,\rho)$ has maximal oscillation with constant 2.*

(4) *Unless $\varphi(f,\rho)$ is almost surely constant, one has $\inf_{t\in[0,T)}\underline{\mathrm{Osc}}_t(\varphi(f,\rho)) > 0$.*

(5) *If $p \in [2,\infty)$, then there is a $c_{(8.29)} > 0$ such that for $0 \leqslant a < t < T$, $\Phi \in \mathrm{CL}^+([0,t])$ with $1 \vee |\Delta X_s| \leqslant \Phi_s$ on $[0,t]$, $\sup_{u\in[0,t]} \Phi_u \in L_p$, and $\lambda > 0$ one has*

$$\mathbb{P}_{\mathcal{F}_a}\left(|E_t(f;\tau_n^\theta, D) - E_a(f;\tau_n^\theta, D)| > \lambda\right) \leqslant c_{(8.29)} \min\left\{\frac{1}{n\lambda^2}, \frac{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,t]}\Phi_u^p\right]}{\lambda^p (T-t)^{p\alpha}}\right\} \text{ a.s.} \tag{8.29}$$

<u>Lower bounds:</u> *For $D \equiv 1$ we can take $\varepsilon = 2-\beta$ and there is an $f_\eta \in \mathrm{Höl}_\eta(\mathbb{R})$ such that for $\varphi_t := \varphi_t(f_\eta, \rho)$ one has:*

(6) $\inf_{t\in(0,T)} (T-t)^\alpha \underline{\mathrm{Osc}}_t(\varphi) > 0$.

(7) *There is a $c_{(8.30)} > 0$ such that for all $\tau = \{t_i\}_{i=0}^n \in \mathcal{T}$ with $\|\tau\|_\theta = \frac{t_k - t_{k-1}}{(T-t_{k-1})^{1-\theta}}$ one has*

$$\inf_{\vartheta_{i-1}\in L_0(\mathcal{F}_{t_{i-1}})} \sup_{a\in[t_{k-1},t_k)} \left\| \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} \left|\varphi_u - \sum_{i=1}^n \vartheta_{i-1}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}u\right]\right\|_{L_\infty} \geqslant c_{(8.30)}^2 \|\tau\|_\theta. \tag{8.30}$$

(8) $\|E(f;\tau,1)\|_{\mathrm{bmo}_2([0,T))} \geqslant \sqrt{\mu(\mathbb{R}) c_{(8.30)}}\sqrt{\|\tau\|_\theta}$ *for all $\tau \in \mathcal{T}$.*

Remark 8.22. From the above theorem we get that

$$\|\varphi_t(f,\rho)\|_{L_\infty} \leqslant c_{(8.26)} (T-t)^{-\frac{1-(\varepsilon+\eta)}{\beta}} \|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}.$$

Let us take a sequence of real numbers $|z_l| = 2^{-l}$, $l \in \mathbb{N}$, and consider the corresponding Dirac-measures $\rho_l = \delta_{z_l}$. Suppose that the small ball condition

$$\rho_l(\{2^{-n} \leqslant |z| < 2^{-n+1}\}) \leqslant c_{(8.12)} 2^{-\varepsilon n}$$

holds uniformly in l and n. Because $\rho_l(\{2^{-l} \leqslant |z| < 2^{-l+1}\}) = 1$ this implies that $1 \leqslant c_{(8.12)} 2^{-\varepsilon n}$ for all $n \in \mathbb{N}$ and finally $\varepsilon = 0$. If we interpret $f \in B_b(\mathbb{R})$ as $\eta = 0$, then we would get an exponent

$$(T-t)^{-\frac{1-(\varepsilon+\eta)}{\beta}} = (T-t)^{-\frac{1}{\beta}}$$

which is the upper bound of [39, Theorem 1.3].

For the proof of the theorem we first need the following Lemma:

Lemma 8.23. *For $0 \leqslant a \leqslant t \in \mathbb{I}$, $0 < r < p < \infty$, $Y \in \mathrm{CL}_0(\mathbb{I})$, $\Phi, \overline{\Phi} \in \mathrm{CL}^+([0,t])$ with $\Phi_s \leqslant \overline{\Phi}_s$ for $s \in [0,t]$ and $\sup_{u\in[0,t]} \overline{\Phi}_u \in L_p$, and for $\lambda > 0$ one has, a.s.,*

$$\begin{aligned}&\mathbb{P}_{\mathcal{F}_a}(|Y_t - Y_a| > \lambda)\\ &\quad\leqslant c_{(8.31)} \min\left\{\frac{\Phi_a^r}{\lambda^r}\|Y\|^r_{\mathrm{bmo}_r^\Phi([0,t])}, \frac{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,t]}\overline{\Phi}_u^p\right]}{\lambda^p}\left[\|Y\|^p_{\mathrm{bmo}_r^{\overline{\Phi}}([0,t])} + |\Delta Y|^p_{\overline{\Phi},[0,t]}\right]\right\}\end{aligned} \tag{8.31}$$

where $c_{(8.31)} > 0$ depends at most on (r,p).

Proof. First we observe that

$$\mathbb{P}_{\mathcal{F}_a}(|Y_t - Y_a| > \lambda) \leqslant \frac{\Phi_a^r}{\lambda^r}\|Y\|^r_{\mathrm{bmo}_r^\Phi} \text{ a.s.} \tag{8.32}$$

Moreover, from Proposition A.5(1) we know that

$$\|Y\|_{\mathrm{BMO}_r^{\overline{\Phi}}([0,t])} \leqslant 2^{\left(\frac{1}{r}-1\right)^+}\left[\|Y\|_{\mathrm{bmo}_r^{\overline{\Phi}}([0,t])} + |\Delta Y|_{\overline{\Phi},[0,t]}\right]. \tag{8.33}$$

Using Proposition A.6(2) this yields to

$$(\mathbb{E}^{\mathcal{F}_a}[|Y_t - Y_a|^p])^{\frac{1}{p}} \leqslant c_{(A.2)}\|Y\|_{\mathrm{BMO}_r^{\overline{\Phi}}([0,t])}\left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,t]}\overline{\Phi}_u^p\right]\right)^{\frac{1}{p}} \text{ a.s.}$$

which implies

$$\mathbb{P}_{\mathcal{F}_a}(|Y_t - Y_a| > \lambda) \leqslant \frac{c^p_{(A.2)}}{\lambda^p}\|Y\|^p_{\mathrm{BMO}_r^{\overline{\Phi}}([0,t])}\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u\in[a,t]}\overline{\Phi}_u^p\right]. \tag{8.34}$$

Combining (8.32), (8.34), and (8.33) implies our statement. □

Proof of Theorem 8.21. (a) <u>Estimates on the density of X_s</u>. Let ψ be the characteristic exponent of X, i.e. $\mathbb{E}\,\mathrm{e}^{\mathrm{i}uX_s} = \mathrm{e}^{-s\psi(u)}$ (see [38, Theorem 8.1]) for $s \in [0,T]$. By (8.24) we obtain

$$0 < \liminf_{|u|\to\infty} \frac{\mathrm{Re}\psi(u)}{|u|^\beta} \leqslant \limsup_{|u|\to\infty} \frac{\mathrm{Re}\psi(u)}{|u|^\beta} < \infty. \tag{8.35}$$

If $s \in (0,T]$, then X_s has a density $p_s \in C^\infty(\mathbb{R})$ with $\lim_{|x|\to\infty}(\partial^m p_s/\partial x^m)(x) = 0$ for $m \in \mathbb{N}_0$ by [34] (see [38, Proposition 28.3]). We combine (8.35) with [39, Theorem 1.3] and [31, Lemma 4.1] and obtain $s_0 \in (0,T]$ and $c_0 > 0$ such that $\|\partial p_s/\partial x\|_{L_1(\mathbb{R})} \leqslant c_0 s^{-\frac{1}{\beta}}$ for $s \in (0,s_0]$. If $s \in (s_0,T]$, then $\|\partial p_s/\partial x\|_{L_1(\mathbb{R})} = \|(\partial p_{s_0}/\partial x) * p_{s-s_0}\|_{L_1(\mathbb{R})} \leqslant c_0 s_0^{-\frac{1}{\beta}}$, so that

$$\left\| \frac{\partial p_s}{\partial x} \right\|_{L_1(\mathbb{R})} \leqslant c_{(8.36)} s^{-\frac{1}{\beta}} \quad \text{for} \quad s \in (0,T] \tag{8.36}$$

with $c_{(8.36)} := c_0 \vee c_0(\frac{T}{s_0})^{\frac{1}{\beta}}$. On the other hand, by (8.35) there is a $c_{(8.37)} = c(\beta, p_\nu) > 0$ such that, for $s \in (0,T]$,

$$\frac{1}{c_{(8.37)}} s^{-\frac{1}{\beta}} \leqslant \int_{\mathbb{R}} \mathrm{e}^{-s\mathrm{Re}\psi(u)}\,\mathrm{d}u \quad \text{and} \quad \int_{\mathbb{R}} \mathrm{e}^{-s\mathrm{Re}\psi(u)}\,|u|\mathrm{d}u \leqslant c_{(8.37)} s^{-\frac{2}{\beta}}. \tag{8.37}$$

When p_ν is symmetric, then p_s is symmetric and $0 \in \mathrm{supp}(X_s)$ for $s \in (0,T]$. Combining (8.37) with the proof of [30, Lemma 7] yields $c_{(8.38)}, \tilde{c}_{(8.38)} > 0$, not depending on (t,x), such that

$$p_s(x) \geqslant c_{(8.38)} s^{-\frac{1}{\beta}} \quad \text{for} \quad |x| < \tilde{c}_{(8.38)} s^{\frac{1}{\beta}} \quad \text{and} \quad s \in (0,T]. \tag{8.38}$$

(b) <u>Upper bounds</u>: (8.13) follows from (8.36) and (8.11). Theorem 8.12 gives $\|X_s\|_{\mathrm{TV}(\rho,\eta)} < \infty$ for $s \in (0,T]$ and $\eta \in (0,1-\varepsilon)$ and $\mathrm{Höl}^0_{\eta,2}(\mathbb{R}) \subseteq \mathrm{Dom}(\Gamma^0_\rho)$ by Theorem 8.9. Now let us check our assertions:

(1) follows from Corollary 8.13 and Theorem 8.11.

(2) For $0 \leqslant a \leqslant t < T$ we use Itô's isometry and and choose $\mathrm{d}\langle X^D\rangle_u = \left(\int_{\mathbb{R}} D^2 \mathrm{d}\mu\right)\mathrm{d}u$ to get, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_a}\big[|E_t(f;\tau,D) - E_a(f;\tau,D)|^2\big] &= \mathbb{E}^{\mathcal{F}_a}\left[\int_a^t \left|\varphi_{u-}(f,\rho) - \sum_{i=1}^n \varphi_{t_{i-1}-}(f,\rho)\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \mathrm{d}\langle X^D\rangle_u\right] \\
&= \left(\int_{\mathbb{R}} D^2\mathrm{d}\mu\right)\mathbb{E}^{\mathcal{F}_a}[[\varphi(f,\rho),\tau]_t - [\varphi(f,\rho),\tau]_a]
\end{aligned}$$

where we use $\varphi_u(f,\rho) = \varphi_{u-}(f,\rho)$ a.s., $u \in [0,T)$, which follows from the chaos expansion. Hence

$$\frac{1}{\left(\int_{\mathbb{R}} D^2\mathrm{d}\mu\right)} \|E(f;\tau,D)\|^2_{\mathrm{bmo}_2([0,T))} = \|[\varphi(f,\rho),\tau]\|_{\mathrm{bmo}_1([0,T))} = \|[\varphi(f,\rho),\tau]\|_{\mathrm{BMO}_1([0,T))}. \tag{8.39}$$

Next we use inequality (8.16) and observe that

$$\begin{aligned}
\frac{(T-t)^{2\alpha}}{2\alpha}\|\varphi_t(f,\rho)\|^2_{L_\infty} &= \|\varphi_t(f,\rho)\|^2_{L_\infty}\int_{[t,T)}(T-u)^{2\alpha-1}\mathrm{d}u \leqslant \int_{[0,T)}(T-u)^{2\alpha-1}\|D_\rho F(u,\cdot)\|^2_{B_b(\mathbb{R})}\mathrm{d}u \\
&\leqslant c^2_{(8.16)}\|f\|^2_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})} < \infty
\end{aligned}$$

which implies

$$\|\varphi_t(f,\rho) - \varphi_0(f,\rho)\|_{L_\infty} \leqslant 2c_{(8.16)}\sqrt{2\alpha}(T-t)^{-\alpha}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}$$

and $\|D_\rho F(t,\cdot)\|_{B_b(\mathbb{R})} \leqslant \sqrt{2\alpha}c_{(8.16)}(T-t)^{-\alpha}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}$. From this, for $0 \leqslant s \leqslant a < T$ the proof of Lemma C.1 gives

$$\|\varphi_a(f,\rho) - \varphi_s(f,\rho)\|_{L_\infty} \leqslant 4c_{(8.16)}\sqrt{2\alpha}\frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}.$$

Now (2) follows from last inequality, (8.26), and Theorem 4.7 (equation (4.6)). Assertions (3) and (4) are a consequence of Theorem 8.19. Regarding (5) we first observe that

$$\left\|E(f;\tau_n^\theta,D)\right\|_{\mathrm{bmo}_2([0,T))} \leqslant c_{(8.28)}\sqrt{\|\tau_n^\theta\|_\theta}\;\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})} \leqslant c_{(8.28)}\sqrt{\frac{T^\theta}{\theta n}}\;\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}. \tag{8.40}$$

Moreover, one has a.s. that

$$|\Delta E_s(f;\tau_n^\theta,D)| \leqslant 2\sqrt{2\alpha}c_{(8.16)}(T-s)^{-\alpha}|\Delta X_s^D|\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})} \quad \text{for} \quad s\in[0,t]$$

where we use $\|D_\rho F(s,\cdot)\|_{B_b(\mathbb{R})} \leqslant \sqrt{2\alpha}c_{(8.16)}(T-s)^{-\alpha}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}$ and Remark 8.6(4). Hence

$$|\Delta E(f;\tau_n^\theta,D)|_{\Phi,[0,t]} \leqslant 2\sqrt{2\alpha}c_{(8.16)}(T-t)^{-\alpha}\|f\|_{\mathrm{Höl}^0_{\eta,2}(\mathbb{R})}. \tag{8.41}$$

Now the statement follows from (8.40), (8.41), and Lemma 8.23.

(c) Lower bounds: We take the function f_η from Theorem 8.20. By (8.24) we derive (8.25) for $\varepsilon = 2-\beta$. Regarding Theorem 8.20 assumption (8.21) follows from (8.24) and assumption (8.22) from (8.38). (6) Theorem 8.19(1) and Theorem 8.20(2) imply

$$\|\varphi_t\|_{L_\infty} = \sup_{x\in\operatorname{supp}(X_t)} |D_\rho(t,x)| \geqslant \frac{1}{c}(T-t)^{\frac{1+\eta}{\beta}-1} = \frac{1}{c}(T-t)^{-\alpha}$$

for $t\in[0,T)$. By Theorem 8.19(3) this gives $\inf_{t\in(0,T)}\underline{\mathrm{Osc}}_t(\varphi)>0$. To verify (6) it is sufficient to prove for some $\varepsilon\in(0,T)$ that $\inf_{t\in[\varepsilon,T)}(T-t)^\alpha\underline{\mathrm{Osc}}_t(\varphi)>0$. As by Theorem 8.19(2) we know that $(\varphi_t)_{t\in[0,T)}$ is of maximal oscillation with constant 2 we get, for $t\in[\varepsilon,T)$,

$$\underline{\mathrm{Osc}}_t(\varphi) \geqslant \frac{1}{2}\|\varphi_t-\varphi_0\|_{L_\infty} \geqslant \frac{1}{2}\|\varphi_t\|_{L_\infty} - \frac{1}{2}\|\varphi_0\|_{L_\infty} \geqslant \frac{1}{2c}(T-t)^{-\alpha} - \frac{1}{2}\|\varphi_0\|_{L_\infty}.$$

Choosing ε appropriate, (6) follows. Items (7)-(8) follow from Theorem 5.7 and (8.39). □

Appendix A. The class $\mathcal{SM}_p(\mathbb{I})$ and BMO-spaces

We summarize some basic facts about the class $\mathcal{SM}_p(\mathbb{I})$ and BMO-spaces that are used in the article. For this we assume a stochastic basis $(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\in[0,T]})$ with $T\in(0,\infty)$ such that $(\Omega,\mathcal{F},\mathbb{P})$ is complete, $\mathcal{F}_0$ contains all null-sets, and such that $\mathcal{F}_t=\bigcap_{s\in(t,T]}\mathcal{F}_s$ for all $t\in[0,T)$. We do not assume that $\mathcal{F}_0$ is generated by the null-sets only. In the computations below we exploit the following fact: given stopping times $\sigma,\tau:\Omega\to\mathbb{I}$ and an integrable random variable $Z:\Omega\to\mathbb{R}$, we have $\{\sigma=\tau\}\in\mathcal{F}_{\sigma\wedge\tau}$ and

$$\mathbb{E}^{\mathcal{F}_\sigma}\big[\mathbb{1}_{\{\sigma=\tau\}}Z\big] = \mathbb{E}^{\mathcal{F}_{\sigma\wedge\tau}}\big[\mathbb{1}_{\{\sigma=\tau\}}Z\big] = \mathbb{E}^{\mathcal{F}_\tau}\big[\mathbb{1}_{\{\sigma=\tau\}}Z\big] \text{ a.s.}$$

Moreover, we again use $\inf\emptyset:=\infty$.

A.1. **Properties of the class $\mathcal{SM}_p$.** We start by a convenient reduction. Since $\mathcal{F}_0$ does not need to be trivial we add the assumption $\Phi_0\in L_p$ to the definition of $\mathcal{SM}_p(\mathbb{I})$ in Definition 2.2.

Proposition A.1. *For $p\in(0,\infty)$ and $\Phi\in\mathrm{CL}^+(\mathbb{I})$ with $\Phi_0\in L_p$ one has $|\Phi|_{\mathcal{SM}_p(\mathbb{I})}=\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})}$, where $|\Phi|_{\mathcal{SM}_p(\mathbb{I})}:=\inf c$ is the infimum over $c\in[1,\infty)$ such that for all $a\in\mathbb{I}$ one has*

$$\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}\Phi_t^p\right] \leqslant c^p\Phi_a^p \quad a.s.$$

Proof. It is clear that $|\Phi|_{\mathcal{SM}_p(\mathbb{I})}\leqslant\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})}$, so that we assume that $c:=|\Phi|_{\mathcal{SM}_p(\mathbb{I})}<\infty$. Let $\rho:\Omega\to\mathbb{I}$ be a stopping time, $h:[0,T)\to[0,\infty)$ be given by $h(t):=\frac{1}{T-t}-\frac{1}{T}$. For $k,N\in\mathbb{N}_0$ set

$$[a_k^N,b_k^N) := h^{-1}\left(\left[\frac{k}{2^N},\frac{k+1}{2^N}\right)\right)\subseteq[0,T) \quad \text{and let} \quad H^N(t):=\mathbb{1}_{\{T\}}(t)T+\sum_{k=0}^\infty \mathbb{1}_{[a_k^N,b_k^N)}(t)b_k^N.$$

Then $H^N(t)\downarrow t$ for all $t\in[0,T]$ and $\rho^N:=H^N(\rho):\Omega\to\mathbb{I}$ is a stopping time as well. Then, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_{\rho^N}}\left[\sup_{\rho^N\leqslant t\in\mathbb{I}}\Phi_t^p\right] &= \mathbb{E}^{\mathcal{F}_{\rho^N}}\big[\mathbb{1}_{\{\rho^N=T\}}\Phi_T^p\big] + \sum_{k=0}^\infty \mathbb{E}^{\mathcal{F}_{\rho^N}}\left[\mathbb{1}_{\{\rho^N=b_k^N\}}\sup_{b_k^N\leqslant t\in\mathbb{I}}\Phi_t^p\right] \\
&= \mathbb{1}_{\{\rho^N=T\}}\Phi_T^p + \sum_{k=0}^\infty \mathbb{1}_{\{\rho^N=b_k^N\}}\mathbb{E}^{\mathcal{F}_{b_k^N}}\left[\mathbb{1}_{\{\rho^N=b_k^N\}}\sup_{b_k^N\leqslant t\in\mathbb{I}}\Phi_t^p\right] \\
&\leqslant \mathbb{1}_{\{\rho^N=T\}}\Phi_T^p + \sum_{k=0}^\infty \mathbb{1}_{\{\rho^N=b_k^N\}}c^p\Phi_{b_k^N}^p \\
&\leqslant c^p\Phi_{\rho^N}^p
\end{aligned}$$

where we omit $\mathbb{1}_{\{\rho^N=T\}}\Phi_T^p$ if $\mathbb{I}=[0,T)$. This implies that $\mathbb{E}^{\mathcal{F}_\rho}\big[\sup_{\rho^N\leqslant t\in\mathbb{I}}\Phi_t^p\big]\leqslant c^p\mathbb{E}^{\mathcal{F}_\rho}\Big[\Phi_{\rho^N}^p\Big]$ a.s. By $N\to\infty$, monotone convergence on the left-hand side and because Φ is càdlàg, and dominated convergence on the right-hand side (Φ is càdlàg and $\mathbb{E}\sup_{t\in\mathbb{I}}\Phi_t^p<\infty$) we obtain the assertion. □

We continue with structural properties of the class $\mathcal{SM}_p$:

Proposition A.2. *For $0<p,p_0,p_1<\infty$ with $\frac{1}{p}=\frac{1}{p_0}+\frac{1}{p_1}$ the following holds:*

(1) *$\mathcal{SM}_q(\mathbb{I})\subseteq\mathcal{SM}_p(\mathbb{I})$ and $\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})}\leqslant\|\Phi\|_{\mathcal{SM}_q(\mathbb{I})}$ whenever $0<p<q<\infty$.*

(2) *If $\Phi\in\mathcal{SM}_p(\mathbb{I})$, then $\Phi^*\in\mathcal{SM}_p(\mathbb{I})$ and $\|\Phi^*\|_{\mathcal{SM}_p(\mathbb{I})}\leqslant\sqrt[p]{1+\|\Phi\|^p_{\mathcal{SM}_p(\mathbb{I})}}$.*

(3) *For $\Phi^i\in\mathcal{SM}_{p_i}(\mathbb{I})$, $i=0,1$, and $\Phi=(\Phi_a)_{a\in[0,T)}$ with $\Phi_a:=\Phi_a^0\Phi_a^1$, one has*
$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{I})}\leqslant\|\Phi^0\|_{\mathcal{SM}_{p_0}(\mathbb{I})}\|\Phi^1\|_{\mathcal{SM}_{p_1}(\mathbb{I})}.$$

Proof. (1) follows from the definition. Now let $a\in\mathbb{I}$. To check (2) we observe $\Phi_0^*=\Phi_0\in L_p$ and
$$\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}|\Phi_t^*|^p\right]=\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in\mathbb{I}}\Phi_t^p\right]\leqslant|\Phi_a^*|^p+\|\Phi\|^p_{\mathcal{SM}_p(\mathbb{I})}\Phi_a^p\leqslant(1+\|\Phi\|^p_{\mathcal{SM}_p(\mathbb{I})})|\Phi_a^*|^p \text{ a.s.}$$
(3) We get $\Phi_0^0\Phi_0^1\in L_p$ and by the conditional Hölder inequality that, a.s.,
$$\begin{aligned}\sqrt[p]{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}\Phi_t^p\right]}&=\sqrt[p]{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}(\Phi_t^0\Phi_t^1)^p\right]}\leqslant\sqrt[p]{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}(\Phi_t^0)^p\sup_{a\leqslant t\in\mathbb{I}}(\Phi_t^1)^p\right]}\\&\leqslant\sqrt[p_0]{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}(\Phi_t^0)^{p_0}\right]}\sqrt[p_1]{\mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\in\mathbb{I}}(\Phi_t^1)^{p_1}\right]}\\&\leqslant\|\Phi^0\|_{\mathcal{SM}_{p_0}(\mathbb{I})}\|\Phi^1\|_{\mathcal{SM}_{p_1}(\mathbb{I})}\Phi_a^0\Phi_a^1\\&=\|\Phi^0\|_{\mathcal{SM}_{p_0}(\mathbb{I})}\|\Phi^1\|_{\mathcal{SM}_{p_1}(\mathbb{I})}\Phi_a.\end{aligned}$$
□

A.2. **Simplifications in the definitions of BMO-spaces.** The first simplification concerns the case $\mathbb{I}=[0,T]$:

Proposition A.3. *For $p\in(0,\infty)$, $Y\in\mathrm{CL}_0([0,T])$, and $\Phi\in\mathrm{CL}^+([0,T])$ define $|Y|_{\mathrm{BMO}_p^\Phi([0,T])}:=\inf c$ and $|Y|_{\mathrm{bmo}_p^\Phi([0,T])}:=\inf c$, respectively, to be the infimum over all $c\in[0,\infty)$ such that, for all $\rho\in\mathcal{S}_T$,*
$$\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_{\rho-}|^p]\leqslant c^p\Phi_\rho^p \text{ a.s.} \quad\text{and}\quad \mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_\rho|^p]\leqslant c^p\Phi_\rho^p \quad\text{a.s.,}$$
respectively. Then one has
$$|Y|_{\mathrm{BMO}_p^\Phi([0,T])}\leqslant\|Y\|_{\mathrm{BMO}_p^\Phi([0,T])}\leqslant 2^{(\frac{1}{p}-1)^+}[1+\|\Phi\|_{\mathcal{SM}_p([0,T])}]|Y|_{\mathrm{BMO}_p^\Phi([0,T])},$$
$$|Y|_{\mathrm{bmo}_p^\Phi([0,T])}\leqslant\|Y\|_{\mathrm{bmo}_p^\Phi([0,T])}\leqslant 2^{(\frac{1}{p}-1)^+}[1+\|\Phi\|_{\mathcal{SM}_p([0,T])}]|Y|_{\mathrm{bmo}_p^\Phi([0,T])},$$
where we additionally assume for the right-hand side inequalities that $\Phi\in\mathcal{SM}_p([0,T])$.

Proof. The inequalities on the left are obvious. To check the inequalities on the right we may assume that $c:=|Y|_{\mathrm{BMO}_p^\Phi([0,T])}$ or $c:=|Y|_{\mathrm{bmo}_p^\Phi([0,T])}$ are finite. To treat both cases simultaneously, we let $t\in[0,T]$, $\rho\in\mathcal{S}_t$, and $A=Y_{\rho-}$ or $A=Y_\rho$, respectively. Then, a.s.,
$$\begin{aligned}\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_t-A|^p]\right)^{\frac{1}{p}}&\leqslant 2^{(\frac{1}{p}-1)^+}\left[\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-A|^p]\right)^{\frac{1}{p}}+\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_t|^p]\right)^{\frac{1}{p}}\right]\\&\leqslant 2^{(\frac{1}{p}-1)^+}\left[c\Phi_\rho+\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_t|^p]\right)^{\frac{1}{p}}\right].\end{aligned}$$
To estimate the second term we may assume $t\in[0,T)$. In case of bmo-spaces this term can be estimated by
$$\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_t|^p]\right)^{\frac{1}{p}}=\left(\mathbb{E}^{\mathcal{F}_\rho}[\mathbb{E}^{\mathcal{F}_t}[|Y_T-Y_t|^p]]\right)^{\frac{1}{p}}\leqslant c\left(\mathbb{E}^{\mathcal{F}_\rho}[\Phi_t^p]\right)^{\frac{1}{p}}\leqslant c\|\Phi\|_{\mathcal{SM}_p([0,T])}\Phi_\rho \quad\text{a.s.}$$
In case of BMO-spaces we find a sequence $t_n\in(t,T]$ with $t_n\downarrow t$. Using Fatou's Lemma for conditional expectations we get, a.s.,
$$\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_t|^p]\right)^{\frac{1}{p}}\leqslant\liminf_n\left(\mathbb{E}^{\mathcal{F}_\rho}[|Y_T-Y_{t_n-}|^p]\right)^{\frac{1}{p}}\leqslant\liminf_n c\left(\mathbb{E}^{\mathcal{F}_\rho}[\Phi_{t_n}^p]\right)^{\frac{1}{p}}\leqslant c\|\Phi\|_{\mathcal{SM}_p([0,T])}\Phi_\rho.$$

□

The second simplification concerns the bmo-spaces. For $p \in (0,\infty)$, $Y \in \mathrm{CL}_0(\mathbb{I})$, and $\Phi \in \mathrm{CL}^+(\mathbb{I})$ we let $|Y|^{\det}_{\mathrm{bmo}_p^\Phi(\mathbb{I})} := \inf c$ be the infimum over all $c \in [0,\infty)$ such that

$$\mathbb{E}^{\mathcal{F}_a}[|Y_t - Y_a|^p] \leqslant c^p \Phi_a^p \text{ a.s.} \quad \text{for all} \quad t \in \mathbb{I} \text{ and } a \in [0,t].$$

With this definition we obtain:

Proposition A.4. *One has* $|\cdot|^{\det}_{\mathrm{bmo}_p^\Phi(\mathbb{I})} = \|\cdot\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})}$ *for all* $p \in (0,\infty)$.

Proof. It is obvious that $|Y|^{\det}_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant \|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})}$. To show $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant |Y|^{\det}_{\mathrm{bmo}_p^\Phi(\mathbb{I})}$ we assume that $c := |Y|^{\det}_{\mathrm{bmo}_p^\Phi(\mathbb{I})} < \infty$, otherwise there is nothing to prove. For $t \in \mathbb{I}$, $\rho \in \mathcal{S}_t$, and $L \in \mathbb{N}_0$ we define the new stopping times $\rho_L(\omega) := \psi_L(\rho(\omega))$ where $\psi_L(0) := 0$ and $\psi_L(s) = s_\ell^L := \ell 2^{-L} t$ when $s \in \left(s_{\ell-1}^L, s_\ell^L\right]$ for $\ell \in \{1,\ldots,2^L\}$. By definition, $\rho_L(\omega) \downarrow \rho(\omega)$ for all $\omega \in \Omega$ as $L \to \infty$. Then

$$\mathbb{E}^{\mathcal{F}_{s_\ell^L}}\left[|Y_t - Y_{s_\ell^L}|^p\right] \leqslant c^p \Phi^p_{s_\ell^L} \text{ a.s.}$$

for $\ell = 0,\ldots,2^L$. Multiplying both sides with $\mathbb{1}_{\{\rho_L = s_\ell^L\}}$ and summing over $\ell = 0,\ldots,2^L$, we get that

$$\mathbb{E}^{\mathcal{F}_{\rho_L}}[|Y_t - Y_{\rho_L}|^p] \leqslant c^p \Phi^p_{\rho_L} \text{ a.s.}$$

For any $M > 0$ this implies

$$\mathbb{E}^{\mathcal{F}_{\rho_L}}[|Y_t - Y_{\rho_L}|^p \wedge M] \leqslant (c^p \Phi^p_{\rho_L}) \wedge M \text{ a.s.}$$

and

$$\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_{\rho_L}|^p \wedge M] \leqslant \mathbb{E}^{\mathcal{F}_\rho}\left[(c^p \Phi^p_{\rho_L}) \wedge M\right] \text{ a.s.}$$

The càdlàg properties of Y and Φ imply

$$\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_\rho|^p \wedge M] \leqslant \mathbb{E}^{\mathcal{F}_\rho}\left[(c^p \Phi^p_\rho) \wedge M\right] \text{ a.s.}$$

By $M \uparrow \infty$ it follows that $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant c$ as desired. □

A.3. **The relation between** BMO_p^Φ **and** bmo_p^Φ**.** The BMO- and bmo-spaces are related to each other as follows:

Proposition A.5. *For* $\Phi \in \mathrm{CL}^+(\mathbb{I})$, $Y \in \mathrm{CL}_0(\mathbb{I})$,

$$|\Delta Y|_{\Phi,\mathbb{I}} := \inf\{c > 0 : |\Delta Y_t| \leqslant c\Phi_t \textit{ for all } t \in \mathbb{I} \textit{ a.s.}\},$$

and $p \in (0,\infty)$ *the following assertions are true:*

(1) $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \leqslant 2^{(\frac{1}{p}-1)^+}\left[\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} + |\Delta Y|_{\Phi,\mathbb{I}}\right]$.

(2) *If* $\mathbb{E}|\Phi_t^*|^p < \infty$ *for all* $t \in \mathbb{I}$, *then* $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$ *and* $|\Delta Y|_{\Phi,\mathbb{I}} \leqslant 2^{\frac{1}{p}\vee 1}\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$.

Proof. For the proof we set $c_p := 2^{(\frac{1}{p}-1)^+}$. (1) For $t \in \mathbb{I}$ and $\rho \in \mathcal{S}_t$ we have, a.s.,

$$\left|\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_{\rho-}|^p]\right|^{\frac{1}{p}} \leqslant c_p\left[\left|\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_\rho|^p]\right|^{\frac{1}{p}} + |\Delta Y_\rho|\right] \leqslant c_p\Phi_\rho\left[\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} + |\Delta Y|_{\Phi,\mathbb{I}}\right]$$

so that $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \leqslant c_p\left[\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} + |\Delta Y|_{\Phi,\mathbb{I}}\right]$. (2) For $t \in \mathbb{I}$ and $\rho \in \mathcal{S}_t$ we have, a.s.,

$$\begin{aligned}
\left|\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_\rho|^p]\right|^{\frac{1}{p}} &= \left|\mathbb{E}^{\mathcal{F}_\rho}\left[\mathbb{1}_{\{\rho<t\}}\lim_n |Y_t - Y_{((\rho+\frac{1}{n})\wedge t)-}|^p\right]\right|^{\frac{1}{p}} \\
&\leqslant \liminf_n \left|\mathbb{E}^{\mathcal{F}_\rho}\left[\mathbb{1}_{\{\rho<t\}}|Y_t - Y_{((\rho+\frac{1}{n})\wedge t)-}|^p\right]\right|^{\frac{1}{p}} \\
&= \liminf_n \left|\mathbb{E}^{\mathcal{F}_\rho}\left[\mathbb{E}^{\mathcal{F}_{(\rho+\frac{1}{n})\wedge t}}\left[\mathbb{1}_{\{\rho<t\}}|Y_t - Y_{((\rho+\frac{1}{n})\wedge t)-}|^p\right]\right]\right|^{\frac{1}{p}} \\
&\leqslant \liminf_n \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}\left|\mathbb{E}^{\mathcal{F}_\rho}\left[\Phi^p_{(\rho+\frac{1}{n})\wedge t}\right]\right|^{\frac{1}{p}} \\
&\leqslant \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}\Phi_\rho
\end{aligned}$$

where we used $\mathbb{E}|\Phi_t^*|^p < \infty$. Hence $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{I})} \leqslant \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$. Moreover, for $\rho \in \mathcal{S}_t$ with $t \in \mathbb{I}$ we get that, a.s.,

$$|\Delta Y_\rho| \leqslant c_p \left[\left|\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_{\rho-}|^p]\right|^{\frac{1}{p}} + \left|\mathbb{E}^{\mathcal{F}_\rho}[|Y_t - Y_\rho|^p]\right|^{\frac{1}{p}} \right] \leqslant 2c_p \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \Phi_\rho.$$

Now we show that this implies

$$|\Delta Y_s| \leqslant [2c_p \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}] \Phi_s \text{ for all } s \in \mathbb{I} \text{ a.s.} \tag{A.1}$$

which yields to $|\Delta Y|_{\Phi,\mathbb{I}} \leqslant 2c_p \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$. It is sufficient to check (A.1) for $s \in [0,t]$ for $0 < t \in \mathbb{I}$. So we define for $k \in \mathbb{N}$ that

$$\rho_1^k := \inf\left\{ s \in (0,t] : |\Delta Y_s| > \tfrac{1}{k} \right\} \wedge t,$$
$$\rho_n^k := \inf\left\{ s \in (\rho_{n-1}^k, t] : |\Delta Y_s| > \tfrac{1}{k} \right\} \wedge t, \quad n \geqslant 2.$$

Since the stochastic basis satisfies the usual conditions and Y is adapted and càdlàg, each $\rho_n^k : \Omega \to [0,t]$ is a stopping time (this is known and can be checked with [8, Lemma 1, Chapter 3]). Hence

$$|\Delta Y_{\rho_n^k}| \leqslant 2c_p \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \Phi_{\rho_n^k} \quad \text{a.s.},$$

and we denote by Ω_n^k the set in which the above inequality holds. Set $\Omega^* = \cap_{k=1}^\infty \cap_{n=1}^\infty \Omega_n^k$, then $\mathbb{P}(\Omega^*) = 1$ and

$$|\Delta Y_s(\omega)| \leqslant 2c_p \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \Phi_t(\omega) \quad \text{for all } (\omega, s) \in \Omega^* \times [0,t],$$

which gives the desired statement. □

A.4. **Distributional estimates.** The BMO-spaces allow for John-Nirenberg theorems. One consequence of the following equivalence of moments:

Proposition A.6. *Let $0 < p \leqslant q < \infty$, $r \in (0,\infty)$, and $\Phi \in \mathrm{CL}^+(\mathbb{I})$.*

(1) *If $\Phi \in \mathcal{SM}_q(\mathbb{I})$ with $\|\Phi\|_{\mathcal{SM}_q(\mathbb{I})} \leqslant d < \infty$, then there is a $c_{(1)} = c(p,q,d) \geqslant 1$ such that*

$$\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})} \sim_{c_{(1)}} \|\cdot\|_{\mathrm{BMO}_q^\Phi(\mathbb{I})}.$$

(2) *There is a constant $c_{(2)} = c(p,r) > 0$ such that, for $0 \leqslant a < t \in \mathbb{I}$ and $Y \in \mathrm{CL}_0(\mathbb{I})$,*

$$\mathbb{E}^{\mathcal{F}_a}\left[\sup_{u \in [a,t]} |Y_u - Y_a|^p \right] \leqslant c_{(\mathrm{A.2})}^p \mathbb{E}^{\mathcal{F}_a}\left[\sup_{u \in [a,t]} \Phi_u^p \right] \|Y\|_{\mathrm{BMO}_r^\Phi([0,t])}^p. \tag{A.2}$$

Proof. (1a) For $\mathbb{I} = [0,T]$ and $\Phi > 0$ on $[0,T] \times \Omega$ the result follows from [20, Corollary 1(i)], where we use Proposition A.3 to relate the formally different BMO-definitions to each other and Proposition A.2(1).

(1b) For $\mathbb{I} = [0,T)$ and $\Phi > 0$ on $[0,T) \times \Omega$ this follows from (1a) by considering the restrictions of the processes to $[0,t]$ for $t \in [0,T)$.

(1c) For $\mathbb{I} = [0,T]$ or $\mathbb{I} = [0,T)$, and $\Phi \geqslant 0$ on $\mathbb{I} \times \Omega$ we proceed as follows: For $\varepsilon > 0$ we consider $\Phi_t^\varepsilon := \Phi_t + \varepsilon$ and observe that $\|\Phi^\varepsilon\|_{\mathcal{SM}_p(\mathbb{I})} \leqslant c_p \|\Phi\|_{\mathcal{SM}_p(\mathbb{I})}$ and $\sup_{\varepsilon>0} \|\cdot\|_{\mathrm{BMO}_p^{\Phi^\varepsilon}(\mathbb{I})} = \|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{I})}$.

(2) Again we replace Φ by Φ^ε. Then we use the proof of [20, (6)] and [20, step (a) of the proof of Corollary 1] to derive the statement with Φ^ε, where the corresponding constant does not depend on $\varepsilon > 0$. By $\varepsilon \downarrow 0$ we arrive at our statement. □

Appendix B. Transition density

Theorem B.1 ([15, p. 263, p. 44]). *For $\hat{b}, \hat{\sigma} \in C_b^\infty$ with $\widehat{\sigma} \geqslant \varepsilon_0 > 0$ there is a jointly continuous transition density $\Gamma_X : (0,T] \times \mathbb{R} \times \mathbb{R} \to (0,\infty)$ such that $\mathbb{P}(X_t^x \in B) = \int_B \Gamma_X(t,x,\xi) \mathrm{d}\xi$ for $t \in (0,T]$ and $B \in \mathcal{B}(\mathbb{R})$, where $(X_t^x)_{t\in[0,T]}$ is the solution to the SDE (6.1) starting in $x \in \mathbb{R}$, such that the following is satisfied:*

(1) *One has $\Gamma_X(s,\cdot,\xi) \in C^\infty(\mathbb{R})$ for $(s,\xi) \in (0,T] \times \mathbb{R}$.*

(2) *For $k \in \mathbb{N}_0$ there is a $c = c(k) > 0$ such that for $(s,x,\xi) \in (0,T] \times \mathbb{R} \times \mathbb{R}$ one has that*

$$\left| \frac{\partial^k \Gamma_X}{\partial x^k}(s,x,\xi) \right| \leqslant c_{(\mathrm{B.1})} s^{-\frac{k}{2}} \gamma_{c_{(\mathrm{B.1})} s}(x - \xi) \quad \textit{where} \quad \gamma_t(\eta) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{\eta^2}{2t}}. \tag{B.1}$$

(3) *For $k \in \mathbb{N}$ and $f \in C_X$ (the set C_Y from Section 6 in the case (C1)) one has*

$$\frac{\partial^k}{\partial x^k}\int_{\mathbb{R}} \Gamma_X(s,x,\xi)f(\xi)\mathrm{d}\xi = \int_{\mathbb{R}} \frac{\partial^k \Gamma_X}{\partial x^k}(s,x,\xi)f(\xi)\mathrm{d}\xi \quad \textit{for} \quad (s,x) \in (0,T]\times\mathbb{R}.$$

Appendix C. A technical lemma

Lemma C.1. *For $\theta \in [0,1]$, a function $\varphi : [0,T) \to \mathbb{R}$, and a non-decreasing function $\Psi : [0,T) \to [0,\infty)$ the following assertions are equivalent:*

(1) *There is a $c_{(C.1)} > 0$ such that for any $0 \leqslant s \leqslant a < T$ one has*

$$|\varphi_a - \varphi_s| \leqslant c_{(C.1)} \frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}} \Psi_a. \tag{C.1}$$

(2) (a) *$\theta \in [0,1)$: There is a $c_{(C.2)} > 0$ such that for $a \in [0,T)$ one has*

$$|\varphi_a - \varphi_0| \leqslant c_{(C.2)} (T-a)^{\frac{\theta-1}{2}} \Psi_a. \tag{C.2}$$

(b) *$\theta = 1$: There is a $c_{(C.3)} > 0$ such that for $0 \leqslant s \leqslant a < T$ one has*

$$|\varphi_a - \varphi_s| \leqslant c_{(C.3)} \left(1 + \log \frac{T-s}{T-a}\right) \Psi_a. \tag{C.3}$$

Proof. (1) $\Rightarrow$ (2) We let $t_n := T - \frac{T}{2^n}$ for $n \geqslant 0$. If $s, a \in [t_{n-1}, t_n]$, $n \geqslant 1$, then (C.1) implies

$$|\varphi_a - \varphi_s| \leqslant c_{(C.1)} \Psi_a T^{\frac{\theta}{2}-\frac{1}{2}} \frac{[1-(1-\frac{1}{2^{n-1}})]^{\frac{\theta}{2}}}{[1-(1-\frac{1}{2^n})]^{\frac{1}{2}}} \leqslant c_{(C.1)} \Psi_a T^{\frac{\theta-1}{2}} (\sqrt{2})^{1+(1-\theta)n}.$$

We now let $s \in [t_{n-1}, t_n)$ and $a \in [t_{n+m-1}, t_{n+m})$ for $n \geqslant 1$, $m \geqslant 0$ arbitrarily. If $\theta \in [0,1)$, then the triangle inequality and the monotonicity of Ψ give

$$|\varphi_a - \varphi_0| \leqslant c_{(C.1)} \Psi_a T^{\frac{\theta-1}{2}} \sum_{k=1}^{n+m} (\sqrt{2})^{1+(1-\theta)k} \leqslant c_{(C.1)} c_\theta \Psi_a T^{\frac{\theta-1}{2}} (\sqrt{2})^{(1-\theta)(n+m-1)} \leqslant \frac{c_{(C.1)} c_\theta \Psi_a}{(T-a)^{\frac{1-\theta}{2}}}$$

for some $c_\theta > 0$ depending on θ only. When $\theta = 1$, similarly as above we get

$$|\varphi_a - \varphi_s| \leqslant c_{(C.1)} \Psi_a \sqrt{2}(1+m) \leqslant 2\sqrt{2} c_{(C.1)} \Psi_a \left(1 + \log \frac{T-s}{T-a}\right).$$

(2) $\Rightarrow$ (1) If $\theta \in [0,1)$, then (C.2) implies for any $0 \leqslant s \leqslant a < T$ that

$$\begin{aligned} |\varphi_a - \varphi_s| &\leqslant |\varphi_a - \varphi_0| + |\varphi_s - \varphi_0| \leqslant c_{(C.2)} \left[\Psi_a (T-a)^{\frac{\theta-1}{2}} + \Psi_s (T-s)^{\frac{\theta-1}{2}}\right] \\ &\leqslant c_{(C.2)} \Psi_a \left[\left(\frac{T-a}{T-s}\right)^{\frac{\theta}{2}} + \left(\frac{T-a}{T-s}\right)^{\frac{1}{2}}\right] \frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}} \leqslant 2c_{(C.2)} \Psi_a \frac{(T-s)^{\frac{\theta}{2}}}{(T-a)^{\frac{1}{2}}}. \end{aligned}$$

The case $\theta = 1$ is derived from the inequality $1 + \log x \leqslant 2\sqrt{x}$, $x \geqslant 1$. □

Appendix D. Malliavin Calculus

D.1. **Itô's chaos decomposition.** We assume the setting from Section 8.1. The random measure M is defined for sets $E \in \mathcal{B}((0,T]\times\mathbb{R})$ with $(\lambda\otimes\mu)(E) < \infty$ by

$$M(E) := \sigma \int_{\{t:(t,0)\in E\}} \mathrm{d}W_t + \lim_{n\to\infty} \int_{E\cap((0,T]\times\{\frac{1}{n}<|x|<n\})} x\widetilde{N}(\mathrm{d}t,\mathrm{d}x),$$

where the limit is taken in L_2. For $n \geqslant 1$, set

$$L_2^n := L_2\left(((0,T]\times\mathbb{R})^n, \mathcal{B}(((0,T]\times\mathbb{R})^n), (\lambda\otimes\mu)^{\otimes n}\right).$$

Let $I_n(f_n)$ denote the multiple integral of an $f_n \in L_2^n$ with respect to the random measure M in the sense of [27] and let $I_n(L_2^n) := \{I_n(f_n) : f_n \in L_2^n\}$. If $f_n^s(z_1,\dots,z_n) = \frac{1}{n!}\sum_\pi f_n(z_{\pi(1)},\dots,z_{\pi(n)})$ for $z_i = (t_i,x_i) \in [0,T]\times\mathbb{R}$ is the symmetrization of f_n, where the sum is taken over all permutations of $\{1,\dots,n\}$, then $I_n(f_n) = I_n(f_n^s)$ a.s. For $n = 0$ we agree about $L_2^0 = \mathbb{R}$ and that $I_0 : \mathbb{R}\to\mathbb{R}$ is the identity, so that $I_0(L_2^0) = \mathbb{R}$. We also use $f_i^s = f_i$ for $f_i \in L_2^i$, $i = 0, 1$. The orthogonal chaos expansion $L_2 = \bigoplus_{n=0}^\infty I_n(L_2^n)$ is due to Itô [27]: given $\xi \in L_2$ there are $f_n \in L_2^n$ such that

$$\xi = \sum_{n=0}^\infty I_n(f_n) \text{ a.s.},$$

so that $I_0(f_0) = \mathbb{E}\xi$. By orthogonality one has $\|\xi\|_{L_2}^2 = \sum_{n=0}^\infty \|I_n(f_n)\|_{L_2}^2 = \sum_{n=0}^\infty n!\|f_n^s\|_{L_2^n}^2$. The Malliavin-Sobolev space $\mathbb{D}_{1,2}$ consists of all $\xi = \sum_{n=0}^\infty I_n(f_n) \in L_2$ such that

$$\|\xi\|_{\mathbb{D}_{1,2}}^2 := \sum_{n=0}^\infty (n+1)\|I_n(f_n)\|_{L_2}^2 < \infty.$$

Given $\xi \in \mathbb{D}_{1,2}$, the Malliavin derivative $D.\xi : (0,T] \times \mathbb{R} \times \Omega \to \mathbb{R} \in L_2(\lambda \otimes \mu \otimes \mathbb{P})$ satisfies

$$\begin{aligned} &\int_{\mathbb{R}} \int_0^T \mathbb{E}\Big((D_{s,z}\xi) I_m(g_m) h(s,z)\Big) \mathrm{d}s\mu(\mathrm{d}z) \\ &= (m+1)! \int_{\mathbb{R}} \int_0^T \cdots \int_{\mathbb{R}} \int_0^T \Big(f_{m+1}^s((t_1,x_1),\ldots,(t_m,x_m),(s,z)) g_m((t_1,x_1),\ldots,(t_m,x_m)) h(s,z) \Big) \\ &\qquad \mathrm{d}t_1\mu(\mathrm{d}x_1)\cdots \mathrm{d}t_m\mu(\mathrm{d}x_m)\mathrm{d}s\mu(\mathrm{d}z) \end{aligned} \tag{D.1}$$

for $h \in L_2^1$, $m \in \mathbb{N}_0$, and symmetric $g_m \in L_2^m$.

Lemma D.1. *If a Borel function $f\colon \mathbb{R} \to \mathbb{R}$ satisfies $f(X_T) \in L_2$, then there exist symmetric $f_n^s \in L_2(\mu^{\otimes n}) := L_2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu^{\otimes n})$ such that the following holds:*

(1) *One has $f(X_T) = \mathbb{E}f(X_T) + \sum_{n=1}^\infty I_n(f_n^s \mathbb{1}_{(0,T]}^{\otimes n})$ a.s.*

(2) *For any $t \in [0,T)$ one has $\mathbb{E}^{\mathcal{F}_t}[f(X_T)] = \mathbb{E}f(X_T) + \sum_{n=1}^\infty I_n(f_n^s \mathbb{1}_{(0,t]}^{\otimes n})$ a.s. Consequently, $\mathbb{E}^{\mathcal{F}_t}[f(X_T)] \in \mathbb{D}_{1,2}$ for any $t \in [0,T)$.*

Proof. (1) follows from [4, Theorem 4]. (2) The first claim is known. For the latter consequence we use the isometry to obtain

$$\sum_{n=1}^\infty (n+1)\|I_n(f_n^s \mathbb{1}_{(0,t]}^{\otimes n})\|_{L_2}^2 = \sum_{n=1}^\infty (n+1)! t^n \|f_n^s\|_{L_2(\mu^{\otimes n})}^2 = \sum_{n=1}^\infty (n+1)\frac{t^n}{T^n}\|I_n(f_n^s \mathbb{1}_{(0,T]}^{\otimes n})\|_{L_2}^2 < \infty,$$

which verifies $\mathbb{E}^{\mathcal{F}_t}[f(X_T)] \in \mathbb{D}_{1,2}$ for $t \in [0,T)$. □

D.2. **Proof of Proposition 8.2.** We fix $t \in (0,T)$. Lemma D.1(2) implies $F(t,X_t) \in \mathbb{D}_{1,2}$ so that

$$D_{s,z}F(t,X_t) = \frac{\partial F}{\partial x}(t,X_t)\mathbb{1}_{(0,t]\times\{0\}}(s,z) + \frac{F(t,X_t+z) - F(t,X_t)}{z}\mathbb{1}_{(0,t]\times(\mathbb{R}\setminus\{0\})}(s,z) \tag{D.2}$$

for $\lambda \otimes \mu \otimes \mathbb{P}$-a.e. $(s,z,\omega) \in (0,T] \times \mathbb{R} \times \Omega$ by [32, Corollary 3.1 of the second article] (see also [40, 1, 41, 18]); if $\sigma > 0$, then $F(t,\cdot) := \mathbb{E}f(\cdot + X_{T-t}) \in C^\infty(\mathbb{R})$ by Example 8.18 for $q = 2$, and if $\sigma = 0$, then the first term on the right-hand side is omitted. As both sides are square-integrable in (s,z,ω) with respect to $\lambda \otimes \mu \otimes \mathbb{P}$ we apply Fubini's theorem to get

$$\begin{aligned} &\frac{1}{t}\int_0^t \int_{\mathbb{R}} (D_{s,z}F(t,X_t))(\omega)\rho(\mathrm{d}z)\mathrm{d}s \\ &= \frac{\partial F}{\partial x}(t,X_t(\omega))\rho(\{0\}) + \int_{\mathbb{R}\setminus\{0\}} \frac{F(t,X_t(\omega)+z) - F(t,X_t(\omega))}{z}\rho(\mathrm{d}z) = \overline{D}_\rho F(t,X_t(\omega)) \end{aligned} \tag{D.3}$$

for $\omega \in \Omega \setminus N_t$ for some null-set N_t, where the integrals on the left-hand side and on the right-hand side (with respect to $\rho(\mathrm{d}z)\mathrm{d}s$ and $\rho(\mathrm{d}z)$, respectively) exist for $\omega \notin N_t$. Then, for $m \in \mathbb{N}_0$ and a symmetric $g_m \in L_2^m$ we obtain from (D.1) with $h(s,z) := \mathbb{1}_{(0,t]}(s)(\mathrm{d}\rho/\mathrm{d}\mu)(z)$ that

$$\begin{aligned} &\mathbb{E}\left(\frac{1}{t}\int_0^t \int_{\mathbb{R}} D_{s,z}F(t,X_t)\rho(\mathrm{d}z)\mathrm{d}s\right) I_m(g_m) \\ &= \frac{(m+1)!}{t}\int_0^t \int_{\mathbb{R}} \int_0^T \int_{\mathbb{R}} \cdots \int_0^T \int_{\mathbb{R}} f_{m+1}^s((t_1,x_1),\ldots,(t_n,x_n),(s,z)) \\ &\qquad g_m((t_1,x_1),\ldots,(t_n,x_n))\mu(\mathrm{d}x_1)\mathrm{d}t_1 \cdots \mu(\mathrm{d}x_m)\mathrm{d}t_m\rho(\mathrm{d}z)\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&= (m+1)! \int_{\mathbb{R}} \int_0^t \int_{\mathbb{R}} \cdots \int_0^t \int_{\mathbb{R}} f^s_{m+1}(x_1,\ldots,x_n,z) \\
&\qquad\qquad\qquad g_m((t_1,x_1),\ldots,(t_n,x_n))\mu(\mathrm{d}x_1)\mathrm{d}t_1\cdots\mu(\mathrm{d}x_m)\mathrm{d}t_m\rho(\mathrm{d}z) \\
&= m! \int_0^T \int_{\mathbb{R}} \cdots \int_0^T \int_{\mathbb{R}} \Big[(m+1)h_m(x_1,\ldots,x_n)\mathbb{1}^{\otimes n}_{(0,t]}(t_1,\ldots,t_n)\Big] g_m((t_1,x_1),\ldots,(t_n,x_n)) \\
&\qquad\qquad\qquad \mu(\mathrm{d}x_1)\mathrm{d}t_1\cdots\mu(\mathrm{d}x_m)\mathrm{d}t_m \\
&= \mathbb{E}\varphi_t(f,\rho)I_m(g_m).
\end{aligned}
$$

This implies that $\overline{D}_\rho F(t,X_t) = \varphi_t(f,\rho)$ a.s. □

D.3. **Interpretation as vector-valued gradient.** Assume that $f \in \mathcal{D}_X \cap L_2(\mathbb{P}_{X_T})$, $\mathrm{d}\rho := D\mathrm{d}\mu / \int_{\mathbb{R}} D\mathrm{d}\mu$, and fix an orthonormal basis $(D_l)_{l\in J} \subseteq L_2(\mathbb{R},\mu)$ with $J = \{1,\ldots,L\}$ or $J = \mathbb{N}$ (note that $L_2(\mathbb{R},\mu)$ is separable). For $(t,\omega,z) \in (0,T)\times\Omega\times\mathbb{R}$ we let $M(t,\omega,z)$ be the right-hand side of (D.2) define the null-sets $P_t := \{\omega\in\Omega : \int_{\mathbb{R}} |M(t,\omega,z)|^2\mu(\mathrm{d}z) = \infty\}$ and

$$M_t^{(l)}(\omega) := \mathbb{1}_{\{\omega\notin P_t\}} \int_{\mathbb{R}} M(t,\omega,z)D_l(z)\mu(\mathrm{d}z) \quad \text{for} \quad l \in J.$$

We obtain random variables $M_t^{(l)} : \Omega \to \mathbb{R}$ such that $\sum_{l\in J} |M_t^{(l)}(\omega)|^2 < \infty$ for all $\omega \in \Omega$. This yields to the map $M_t := (M_t^{(l)})_{l\in J} : \Omega \to \ell_2^J \cong L_2(\mathbb{R},\mu)$. For $\omega \notin P_t$ this gives

$$\langle M_t(\omega), \overline{D}\rangle_{\ell_2^J} = \int_R M(t,\omega,z)D(z)\mu(\mathrm{d}z),$$

where $\overline{D}$ is D considered in ℓ_2^J, so that

$$\frac{\langle M_t, \overline{D}\rangle_{\ell_2^J}}{\int_{\mathbb{R}} D\mathrm{d}\mu} = \overline{D}_\rho F(t,X_t) = \frac{1}{t}\int_0^t \int_{\mathbb{R}} D_{s,z}F(t,X_t)\rho(\mathrm{d}z)\mathrm{d}s \text{ a.s.} \quad \text{for} \quad t\in(0,T). \tag{D.4}$$

References

1. E. Alòs, J.A. León and J. Vives, An anticipating Itô formula for Lévy processes. *Lat. Am. J. Probab. Math. Stat.* 4(2008)285–305.
2. D. Applebaum, R. Bañuelos, Probabilistic approach to fractional integrals and the Hardy-Littlewood-Sobolev inequality, *Analytic methods in interdisciplinary applications*, Springer Proc. Math. Stat., 116, Cham, 2015, 17–40.
3. R. Arai, E. Nakai, G. Sadasue, Fractional integrals and their commutators on martingale Orlicz spaces, *Journal of Mathematical Analysis and Applications* 487(2020).
4. F. Baumgartner, S. Geiss, Permutation invariant functionals of Lévy processes, *Trans. Amer. Math. Soc.* 369(2017)8607–8641.
5. C. Bennet and R. Sharpley, *Interpolation of operators*, Academic Press, 1988.
6. F. Benth, G. Di Nunno, A. Løkka, B. Øksendal, F. Proske, Explicit representation of the minimal variance portfolio in markets driven by Lévy processes, *Math. Finance* 13(2003)55–72.
7. J. Bergh, J. Löfström, *Interpolation Spaces*, Springer, 1976.
8. P. Billingsley, *Convergence of Probability Measures*, 2nd edition, John Wiley & Sons, 1999.
9. M. Brodén, P. Tankov, Tracking errors from discrete hedging in exponential Lévy model, *Int. J. Theor. Appl. Finance* 14(2011)803–837.
10. Z.-Q. Chen, Z. Hao, X. Zhang, Hölder regularity and gradient estimates Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes, arXiv:2001.03873, 2020.
11. T. Choulli, L. Krawczyk, C. Stricker, $\mathcal{E}$-martingales and their applications in mathematical finance, *Ann. Probab.* 26(1998)853–876.
12. R. Cont, P. Tankov, E. Voltchkova, *Hedging with options in models with jumps*, In: *Stochastic Analysis and Applications – The Abel Symposium 2005*, Springer, Berlin, 2007.
13. F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, C. Stricker, Weighted norm inequalities and hedging in incomplete markets, *Finance Stoch.* 1(1997)181–227.
14. C. Dellacherie, P.-A. Meyer, *Probabilities and potential B: Theory of martingales*, North-Holland, 1982.
15. A. Friedman. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, 1964.
16. C. Geiss, S. Geiss, On approximation of a class of stochastic integrals and interpolation, *Stochastics Stochastics Rep.* 76(2004)339–362.
17. C. Geiss, S. Geiss, E. Laukkarinen, A note on Malliavin fractional smoothness for Lévy processes and approximation, *Potential Anal.* 39(2013)203–230.
18. C. Geiss and A. Steinicke. Malliavin derivatives of random functions and applications to Lévy driven BSDEs. Elect. J. Prob. 21(2016), no. 10, 1–28.
19. S. Geiss, Quantitative approximation of certain stochastic integrals, *Stochastics Stochastics Rep.* 73(2002)241–270.

20. S. Geiss, Weighted BMO and discrete time hedging within the Black–Scholes model, *Probab. Theory Related Fields* 132(2005)13–38.
21. S. Geiss, On the sharpness of embeddings of Hölder spaces into Gaussian Besov spaces, in preparation, 2020.
22. S. Geiss, M. Hujo, Interpolation and approximation in $L_2(\gamma)$, *Journal of Approximation Theory* 144(2007)213–232.
23. S. Geiss, A. Toivola, Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces, *Bernoulli* 15(2009)925–954.
24. S. Geiss, A. Toivola, On fractional smoothness and L_p-approximation on the Gaussian space, *Ann. Probab.* 43(2015)605–638.
25. S. Geiss, J. Ylinen, Weighted bounded mean oscillation applied to backward stochastic differential equations, *Stoch. Proc. Appl.* 130(2020)3711–3752.
26. Y. Hu, D. Nualart, J. Song, Fractional martingales and characterization of the fractional Brownian motion, *Ann. Probab.* 37(2009)2404–2430.
27. K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, *Trans. Amer. Math. Soc.* 81(1956)253–263.
28. J. Jacob, S. Méléard, P.E. Protter, Explicit form and robustness of martingale representations, *Ann. Probab.* 28(2000)1747–1780.
29. V. Knopova, R.L. Schilling, A note on the existence of transition probability densities of Lévy process, *Forum Math.* 25(2013)125–149.
30. K. Kaleta, P. Sztonyk, Estimates of transition densities and their derivatives for jump Lévy processes, *J. Math. Anal. Appl.* 431(2015)260–282.
31. F. Kühn, R.L. Schilling, Strong convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs, *Stochastic Process. Appl.* 129(2019)2654–2680.
32. E. Laukkarinen, *On Malliavin calculus and approximation of stochastic integrals for Lévy process*, PhD thesis, University of Jyväskylä, 2013.
33. E. Laukkarinen, Malliavin smoothness on the Lévy space with Hölder continuous or BV functionals, *Stochastic Process. Appl.* 130 (2020)4766–4792.
34. S. Orey, On continuity properties of infinitely divisible distribution functions, *Ann. Math. Statist.* 39(1968)936-937.
35. P.E. Protter, *Stochastic integration and differential equations (2nd ed., ver. 2.1)*, Springer-Verlag, Berlin, 2005.
36. W. Rudin, *Functional Analysis*, McGraw-Hill, 1973.
37. W. Rudin, *Real and complex analysis*, McGraw-Hill, 1970.
38. K. Sato, *Lévy processes and infinitely divisible distributions (2nd ed.)*, Cambridge University Press, Cambridge, 2013.
39. R.L. Schilling, P. Sztonyk, J. Wang, Coupling property and gradient estimates of Lévy processes via the symbol, *Bernoulli* 18(2012)1128–1149.
40. J. Solé, F. Utzet, J. Vives, Canonical Lévy processes and Malliavin calculus, Stochastic Process. Appl. 117(2007)165–187.
41. A. Steinicke, Functionals of a Lévy process on canonical and generic probability spaces. J. Theoret. Probab. 29(2016)443?458.
42. P. Sztonyk, Transition density estimates for jump Lévy processes, *Stochastic Process. Appl.* 121(2011)1245–1265.
43. H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, North-Holland, 1978.

Department of Mathematics and Statistics, P.O.Box 35, FI-40014 University of Jyväskylä, Finland
Email address: stefan.geiss@jyu.fi

Department of Mathematics and Statistics, P.O.Box 35, FI-40014 University of Jyväskylä, Finland
Email address: thuan.t.nguyen@jyu.fi

APPROXIMATION OF STOCHASTIC INTEGRALS WITH JUMPS IN WEIGHTED BOUNDED MEAN OSCILLATION SPACES

NGUYEN TRAN THUAN

Abstract. This article investigates discrete-time approximation methods of stochastic integrals driven by semimartingales with jumps. The error process is measured with two weighted bounded mean oscillation norms (which coincide in the case of no jumps) and lead especially to L_p-estimates. Besides, this approach also allows a change of the underlying measure which leaves the error estimates unchanged if the change of measure satisfies a reverse Hölder inequality. We propose a new approximation scheme and discuss a way to optimise the approximation rate by adapting the discretization times to the setting, especially to the jump behavior of the considered semimartingale. The research was inspired by Mathematical Finance: We apply the methods in the special case where the semimartingale is an exponential Lévy process to mean variance hedging of European type options. To do this, an explicit representation for the hedging strategy is shown under a general condition using Malliavin calculus. The results reveal the interplay between properties of the Lévy measure, the regularity of the pay-off function and the approximation rate.

1. Introduction

1.1. **The problem.** This article is concerned 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 $T \in (0,\infty)$ 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 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,

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

for the deterministic time-net $\tau = \{0 = t_0 < t_1 < \cdots < t_n = T\}$. We will study the corresponding error E^{Rm} in L_2, but *locally in time*, which means 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 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

---

Date: November 7, 2020.

2010 *Mathematics Subject Classification.* Primary: 60H05, 41A25; Secondary: 60G51, 91G99.

Key words and phrases. Approximation of stochastic integral, Convergence rate, Lévy process, Semimartingale, Weighted bounded mean oscillation.

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

two different spaces). More precisely, we will work with *weighted* bmo-norms introduced in [17, 18], because we consider

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

Here, $\mathbb{E}^{\mathcal{F}_\rho}$ stands for the conditional expectation with respect to $\mathcal{F}_\rho$, and the *weight function* $\Phi = (\Phi_t)_{t\in[0,T]}$ will be specified later. We will denote the infimum of the $c_{(1.2)} > 0$ by $\|E^{\mathrm{Rm}}\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})}$. In Theorem 3.5 we state that under certain conditions it holds that

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

where $\theta \in (0, 1]$ is related to the growth property of the integrand ϑ. Here, $\|\tau\|_\theta$ denotes a nonlinear mesh size, and in Subsection 3.3 we discuss that τ can be chosen such that $\|\tau\|_\theta \leqslant \frac{c}{n}$, implying the optimal approximation rate

$$\|E^{\mathrm{Rm}}\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})} \leqslant \frac{c}{\sqrt{n}}.$$

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 Subsection 1.3 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 $\mathbb{Q}$, 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 the existence of possibly large jumps of S, which is in contrast to the geometric Brownian setting in [17]. To overcome this difficulty, we adapt and develop further the idea using a *small-large jump decomposition* of S presented in Dereich and Heidenreich [9] 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 adjusted method**. The time-net used in this approximation method is a combination of the given deterministic time-net in the Riemann sum and random times of carefully chosen large jumps of S. One also stresses that this method is different from that in Rosenbaum and Tankov [32], where the authors track jumps of the integrand of the approximated stochastic integral, while here we only observe the jumps of the integrator which is less expensive (in computation).

Let E^{adj} denote the error caused from the approximation with the jump adjustment scheme. To formulate the result, we assume that S is given as the 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 Subsection 2.3). We also use the weight $\overline{\Phi}$, which is a variant of Φ, given in (3.9). Then, Theorem 3.14 implies that for suitably chosen time-nets and corrections it holds that

$$\|E^{\mathrm{adj}}\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant \frac{c}{\sqrt{n}}$$

under the condition that the random measure ν of the characteristics of Z satisfies

$$\sup_{r>0}\left|\int_{|z|>r} z\nu_t(\omega, \mathrm{d}z)\right| \leqslant c$$

almost everywhere with respect to $\mathbb{P} \otimes \lambda$, where λ is the Lebesgue measure, and

$$\|E^{\mathrm{adj}}\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant \frac{c}{\sqrt[2\alpha]{n}}$$

provided that

$$(\omega, t) \mapsto \int_{|z| \leqslant 1} |z|^\alpha \nu_t(\omega, \mathrm{d}z)$$

has a finite essential supremum with respect to $\mathbb{P} \otimes \lambda$.

As an application, we choose S to be an exponential Lévy process and measure the discretization error in mean-variance hedging of a European payoff. To measure the hedging error we provide in Theorem 4.2 using Malliavin calculus an explicit representation of the mean-variance hedging strategy for a European payoff which is to the best of our knowledge new in this generality.

1.2. **Background.** 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 stochastic differential equation (SDE) $\mathrm{d}S_t = \sigma(S_t)\mathrm{d}W_t$, where W is the standard Brownian motion and the function σ satisfies some suitable conditions. 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, and this substitution causes the discretization error.

The error represented by the difference between a stochastic integral and its discretization has been extensively analysed 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 to [14, 21, 39], where the error along with its convergence rates was examined. The weak convergence of the error was treated in [20, 21]. When the driving process is a continuous semimartingale, the convergence in the L_2-sense was studied in [13], and in the almost sure sense it was considered in [22].

In this article, we allow the semimartingale to jump since many important processes used in financial modelling are not continuous (see [7]), 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 [34] and the references therein). The present article focuses on the first type of hedging error only. The discretization error was studied within Lévy models in the weak convergence sense in [37], in the L_2-sense in [5, 15], and for a more general jump model under the L_2-setting in [32].

1.3. **Why a weighted BMO-approach?** 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 in the Wiener space in [19]. A remarkably different route given in [17] 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 ([17, 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 $\alpha, \beta > 0$,*

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

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. Moreover, as a benefit to further applications in mathematical finance, the weighted BMO-approach also allows 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 notations are contained in Section 2. The main results are provided in Section 3 and theirs proofs are given in Section 5. In Section 4, we give some applications of those main results in exponential Lévy models. Section 6 presents briefly Malliavin calculus for Lévy processes which is the main tool to obtain an explicit mean-variance hedging strategy for a European type option in Theorem 4.2. The regularity of weight processes used in this article is shown in Section 7. In Section 8, we establish some gradient type estimates for a Lévy semigroup on Hölder spaces, which are used to verify the main results in the Lévy setting (Theorem 4.6).

2. Preliminaries

2.1. Notations and conventions.

General notations. Denote $\mathbb{R}_+ := (0, \infty)$ 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 $A, B \geqslant 0$ and $c \geqslant 1$, the notation $A \sim_c B$ stands for $\frac{1}{c}A \leqslant B \leqslant cA$. Subindexing a symbol by a label means the place where that symbol appears (e.g., $c_{(2.2)}$ refers to the relation (2.2)).

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 (non-negative), 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 $L_p(\mathbb{P}) := L_p(\Omega, \mathcal{F}, \mathbb{P})$.

Notations 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 notations:

- 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 "$\geqslant$", "$\leqslant$", 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. $X \geqslant 0$);
- 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 $\|X\|_{\mathcal{L}_p(\mathbb{P})} := \|\sup_{t\in[0,T]} |X_t|\|_{L_p(\mathbb{P})}$.

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 that the property resembles a supermartingale). Let $\mathcal{S}([0,T])$ denote the family of all stopping times $\rho\colon \Omega \to [0,T]$ and set $\inf\emptyset := \infty$.

Definition 2.1 ([17, 18])**.** For $p \in (0,\infty)$, $Y \in \mathrm{CL}_0([0,T])$ and $\Phi \in \mathrm{CL}^+([0,T])$, define

$$\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} := \inf\left\{c \geqslant 0 : \mathbb{E}^{\mathcal{F}_\rho}[|Y_T - Y_{\rho-}|^p] \leqslant c^p\Phi_\rho^p \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T])\right\},$$
$$\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} := \inf\left\{c \geqslant 0 : \mathbb{E}^{\mathcal{F}_\rho}[|Y_T - Y_{\rho}|^p] \leqslant c^p\Phi_\rho^p \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T])\right\},$$
$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} := \inf\left\{c \geqslant 0 : \mathbb{E}^{\mathcal{F}_\rho}\left[\sup_{\rho\leqslant t\leqslant T}\Phi_t^p\right] \leqslant c^p\Phi_\rho^p \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T])\right\}.$$

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 non-weighted case, i.e. $\Phi \equiv 1$, we drop Φ and simply use the notation $\mathrm{BMO}_p(\mathbb{P})$ or $\mathrm{bmo}_p(\mathbb{P})$.

Remark 2.2. Thanks to [18, 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, i.e.

$$\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} = \inf\left\{c \geqslant 0 : \mathbb{E}^{\mathcal{F}_a}[|Y_T - Y_a|^p] \leqslant c^p\Phi_a^p \quad \text{a.s.}, \forall a \in [0,T]\right\},$$
$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} = \inf\left\{c \geqslant 0 : \mathbb{E}^{\mathcal{F}_a}\left[\sup_{a\leqslant t\leqslant T}\Phi_t^p\right] \leqslant c^p\Phi_a^p \quad \text{a.s.}, \forall a \in [0,T]\right\}.$$

The theory of classical non-weighted BMO/bmo-martingales can be found in [11, Ch.VII] or [31, Ch.IV], and they were used later in different contexts (see, e.g., [6, 10]). The notion of weighted BMO space above was introduced and discussed in [17] where it was developed for general càdlàg processes which are not necessarily martingales.

It is clear from the definition 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 ([18], 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

$$|\Delta Y|_\Phi := \inf\{c \geqslant 0 : |\Delta Y_t| \leqslant c\Phi_t \ \text{ for all } t \in [0,T] \text{ a.s.}\}. \tag{2.1}$$

Definition 2.4 ([17]). 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 $c_{(2.2)} > 0$ such that U satisfies the following *reverse Hölder inequality*

$$\sqrt[s]{\mathbb{E}^{\mathcal{F}_\rho}[U^s]} \leqslant c_{(2.2)}\mathbb{E}^{\mathcal{F}_\rho}[U] \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T]), \tag{2.2}$$

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

We recall in Proposition 2.5 some features of weighted BMO which play a key role in our applications. Notice that Proposition 2.5 is *not* valid for weighted bmo in general.

Proposition 2.5 ([17]). *Let $p \in (0,\infty)$.*

(1) *There exists a constant $c_1 = c_1(p) > 0$ such that $\|\cdot\|_{L_p(\mathbb{P})} \leqslant c_1\|\Phi\|_{L_p(\mathbb{P})}\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$.*

(2) *If $\Phi \in \mathcal{SM}_p(\mathbb{P})$, then for any $r \in (0,p]$ there is a constant $c_2 = c_2(r,p,\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}) > 0$ 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 $c_3 = c(s,p) > 0$ such that $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{Q})} \leqslant c_3\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$.*

Proof. Items (1) and (2) are due to [18, Proposition A.6]. For item (3), we apply [17, 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.

• $S \in \mathrm{CL}([0,T])$ satisfies the SDE[1]

$$\mathrm{d}S_t = \sigma(S_{t-})\mathrm{d}Z_t, \quad S_0 \in \mathcal{R}_S, \tag{2.3}$$

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

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

• $Z \in \mathrm{CL}([0,T])$ is a square integrable semimartingale defined on $(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\in[0,T]})$ with the representation

$$Z_t = Z_0 + Z_t^{\mathrm{c}} + \int_0^t\int_{\mathbb{R}_0} z(N_Z - \pi_Z)(\mathrm{d}u,\mathrm{d}z) + \int_0^t V_u\mathrm{d}u, \quad t \in [0,T], \tag{2.4}$$

where $Z_0 \in \mathbb{R}$, V is a progressively measurable process, Z^{c} is a pathwise continuous square integrable 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 for Z are the following:

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

$$\pi_Z(\omega,\mathrm{d}t,\mathrm{d}z) = \nu_t(\omega,\mathrm{d}z)\mathrm{d}t, \tag{2.5}$$

where the transition kernel $\nu_t(\omega,\cdot)$ is a Lévy measure, i.e. 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 C such that $\langle Z^{\mathrm{c}}\rangle = \int_0^\cdot C_u^2\mathrm{d}u$.

---

[1]See, for example, [31, Ch.V, Sec.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 $0 \leqslant s < t \leqslant T$, $B \in \mathcal{B}(\mathbb{R}_0)$.

[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 non-negative 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 [26, Ch.II, Sec.1] for more details).

($Z3$) The processes V and K, where $K_t := (C_t^2 + \int_{\mathbb{R}} z^2 \nu_t(\mathrm{d}z))^{1/2}$, satisfy that

$$V_{(2.6)} := \| \|V\|_{L_2([0,T],\lambda)} \|_{L_\infty(\mathbb{P})} < \infty, \quad K_{(2.6)} := \|K\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty. \tag{2.6}$$

- ϑ belongs to the family Σ_S^{adm} of *admissible integrands*, where

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

Remark 2.6. (a) By a standard stopping argument and Gronwall's lemma, (2.3) implies that S is an $L_2(\mathbb{P})$-semimartingale and

$$\mathbb{E}\int_0^T \sigma(S_u)^2 \mathrm{d}u = \mathbb{E}\int_0^T \sigma(S_{u-})^2 \mathrm{d}u < \infty. \tag{2.7}$$

(b) For each $t \in [0,T]$, it follows from (2.5) 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. Thus, it is natural to assume $\Delta\vartheta_t = 0$ a.s. for admissible integrands.

3. Main results

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 [18, Assumption 4.1].

Assumption 3.1. For $\vartheta \in \Sigma_S^{\mathrm{adm}}$, we assume that there exists a random measure

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

such that

$$\Upsilon(\omega,(0,t]) < \infty, \quad \forall (\omega,t) \in \Omega\times(0,T),$$

and such that there exists a constant $c_{(3.1)} > 0$ such that for any $0 \leqslant a < b < T$,

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

Examples for Assumption 3.1 when S is a diffusion on the Wiener space are discussed in [18, Section 6], and in that context the random measure Υ describes some kind of curvature of the stochastic integral. In the Lévy setting when S is a Lévy process and $\sigma \equiv 1$, an example for Υ is also given in [18, Section 8]. We now provide in Example 3.2 another formula for Υ which is used in the exponential Lévy setting in Section 4.

Example 3.2. Assume that $M := \vartheta\sigma(S) \in \mathrm{CL}([0,T))$ is an $L_2(\mathbb{P})$-martingale. Then, the random measure Υ defined by

$$\Upsilon(\omega,\mathrm{d}t) := \mathrm{d}\langle M\rangle_t(\omega) + |\sigma|_{\mathrm{Lip}}^2 |M_t(\omega)|^2 \mathrm{d}t$$

satisfies (3.1) with $c_{(3.1)}^2 = 2 + 8(K_{(2.6)}^2 + V_{(2.6)}^2)\mathrm{e}^{4T|\sigma|_{\mathrm{Lip}}^2 (K_{(2.6)}^2 + V_{(2.6)}^2)}$. Indeed, for any $0 \leqslant a < b < T$, using the triangle inequality and Lemma 5.1 we have

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |\vartheta_t - \vartheta_a|^2 \sigma(S_t)^2 \mathrm{d}t\right] \\
&\leqslant \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |M_t - M_a|^2 \mathrm{d}t\right] + \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} \vartheta_a^2 |\sigma(S_t) - \sigma(S_a)|^2 \mathrm{d}t\right]
\end{aligned}$$

$$\leqslant \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]}\int_{(a,t]} \mathrm{d}\langle M\rangle_u \mathrm{d}t\right] + c_{(5.2)}^2 (b-a)^2 M_a^2$$
$$\leqslant \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} (b-u)\mathrm{d}\langle M\rangle_u\right] + 2c_{(5.2)}^2 \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} (b-u) M_u^2 \mathrm{d}u\right],$$

where in the last inequality, we apply Fubini's theorem for first term and use the martingale property of M for the second term. Hence, the assertion follows from (5.5) and (5.4).

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

Assumption 3.3. Let $\theta \in (0,1]$. Assume that Assumption 3.1 is satisfied and there is an a.s. non-decreasing process $\Theta \in \mathrm{CL}^+([0,T])$ such that the following two conditions hold:

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

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

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

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

where

$$\Phi := \Theta\sigma(S).$$

The parameter θ in Assumption 3.3 describes the growth (pathwise and relatively to Θ) of ϑ when the time variable a approaches the terminal time T. For the Black–Scholes model with the delta-hedging strategy ϑ, the parameter θ can be interpreted as the fractional smoothness of the payoff in the sense of [14, 19].

Various specifications of Assumption 3.3 in the Brownian setting or in the Lévy setting are provided in [18]. In Section 4, we use Assumption 3.3 in the exponential Lévy setting which extends [18].

3.1. The basic method: Riemann approximation.

Definition 3.4. (1) Let $\mathcal{T}_{\mathrm{det}}$ be the family of all *deterministic* time-nets $\tau = (t_i)_{i=0}^n$ on $[0,T]$ with $0 = t_0 < t_1 < \cdots < t_n = T$, $n \geqslant 1$. The mesh size of $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ 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 $\vartheta \in \Sigma_S^{\mathrm{adm}}$, $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ and $t \in [0,T]$, we let

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

Below is the main result in this subsection.

Theorem 3.5. *Let Assumption 3.3 hold for some $\theta \in (0,1]$. Then, there exists a constant $c_{(3.4)} > 0$ such that for any $\tau \in \mathcal{T}_{\mathrm{det}}$,*

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

Since the weighted bmo and weighted BMO-norms of $E^{\mathrm{Rm}}(\vartheta, \tau)$ coincide when the driving process S is continuous, we derive directly from Theorem 3.5 and Proposition 2.5 the following result.

Corollary 3.6. *Let Assumption 3.3 hold for some $\theta \in (0,1]$. If S is continuous, then the following assertions hold, where the constants $c_1, c_2, c_3 > 0$ do not depend on τ.*

(1) *One has $\|E^{\mathrm{Rm}}(\vartheta,\tau)\|_{\mathrm{BMO}_2^{\Phi}(\mathbb{P})} \leqslant c_1\sqrt{\|\tau\|_\theta}$ for any $\tau \in \mathcal{T}_{\det}$.*
Furthermore, if $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in [2,\infty)$, then for any $\tau \in \mathcal{T}_{\det}$,
$$\|E^{\mathrm{Rm}}(\vartheta,\tau)\|_{L_p(\mathbb{P})} \leqslant c_2\sqrt{\|\tau\|_\theta}.$$
(2) *If $\mathbb{Q} \in \mathcal{RH}_s(\mathbb{P})$ for some $s \in (1,\infty)$ and $\Phi \in \mathcal{SM}_2(\mathbb{Q})$, then for any $\tau \in \mathcal{T}_{\det}$,*
$$\|E^{\mathrm{Rm}}(\vartheta,\tau)\|_{\mathrm{BMO}_2^{\Phi}(\mathbb{Q})} \leqslant c_3\sqrt{\|\tau\|_\theta}.$$

In particular, when S is a geometric Brownian motion and ϑ is the delta-hedging strategy of a Lipschitz functional of S_T, then Corollary 3.6 gives the upper bound part in [17, Theorem 7].

3.2. **The jump adjusted method.** In Corollary 3.6, the continuity of S is crucial to derive the conclusions. If S has jumps, then those results may fail as shown in the following example.

Example 3.7. In the notations of Subsection 2.3, we let $Z = \widetilde{J}$, where $\widetilde{J}_t := J_t - rt$ is a compensated Poisson process with intensity $r > 0$. Let $\sigma \equiv 1$ (i.e. $S = Z$). Let $f\colon (0,T] \times \mathbb{N} \to \mathbb{R}$ 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 $\vartheta_t = \vartheta_0 + \int_{(0,t\wedge\rho_2]} f(s, J_{s-})\mathrm{d}\widetilde{J}_s$, $t \in (0,T]$. It is not difficult to check that $\vartheta \in \Sigma_S^{\mathrm{adm}}$ 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 3.2. In addition, it is straightforward to check that Assumption 3.3 holds true for $\Theta \equiv \Phi \equiv 1$ and for any $\theta \in (0,1]$.

Take $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ arbitrarily. On the set $\{0 < \rho_1 < \rho_2 < t_1\}$ we have
$$\begin{aligned}
|\Delta E^{\mathrm{Rm}}_{\rho_2}(\vartheta,\tau)| &= \sum_{i=1}^n |\vartheta_{\rho_2-} - \vartheta_{t_{i-1}-}|\mathbb{1}_{(t_{i-1},t_i]}(\rho_2)|\Delta J_{\rho_2}| \\
&= |\vartheta_{\rho_2-} - \vartheta_0| = \left| f(\rho_1, J_{\rho_1-}) - r\int_{(0,\rho_2)} f(s, J_{s-})\mathrm{d}s \right| \\
&\geqslant |f(\rho_1, 0)| - rT\|f\|_\infty \geqslant \delta.
\end{aligned}$$
Since $\mathbb{P}(0 < \rho_1 < \rho_2 < t_1) > 0$, it implies that $\inf_{\tau\in\mathcal{T}_{\det}} \|\Delta E^{\mathrm{Rm}}_{\rho_2}(\vartheta,\tau)\|_{L_\infty(\mathbb{P})} \geqslant \delta$. 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 weighted BMO to derive results as in Corollary 3.6 for jump models, we propose another 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 Subsection 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 $\rho(\varepsilon,\kappa) = (\rho_i(\varepsilon,\kappa))_{i\geqslant 0}$ based on tracking the jumps of S as follows (recall that $\inf\emptyset := \infty$).

Definition 3.8. For $\varepsilon > 0$ and $\kappa \geqslant 0$, let $\rho_0(\varepsilon,\kappa) := 0$ and

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

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

The quantity $\varepsilon(T-t)^\kappa$ above is the level at time t where we decide which jumps of S are (relatively) large, and moreover, this level shrinks when t approaches the terminal time T and $\kappa > 0$. 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 $\kappa = 0$, the control parameter ε can be interpreted as the *jump size threshold*.

The scheme of Riemann approximation with correction is as follows.

Definition 3.9. Let $\varepsilon > 0$, $\kappa \in [0,\frac{1}{2})$ and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\text{det}}$.

(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 $t \in [0,T]$, we define

$$\vartheta_t^\tau := \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(t),$$

$$A_t^{\text{adj}}(\vartheta,\tau|\varepsilon,\kappa) := A_t^{\text{Rm}}(\vartheta,\tau) + \sum_{\rho_i(\varepsilon,\kappa)\in[0,t]\cap[0,T)} \left(\vartheta_{\rho_i(\varepsilon,\kappa)-} - \vartheta^\tau_{\rho_i(\varepsilon,\kappa)}\right)\Delta S_{\rho_i(\varepsilon,\kappa)}, \tag{3.8}$$

$$E_t^{\text{adj}}(\vartheta,\tau|\varepsilon,\kappa) := \int_0^t \vartheta_{u-}\mathrm{d}S_u - A_t^{\text{adj}}(\vartheta,\tau|\varepsilon,\kappa),$$

where $A^{\text{Rm}}(\vartheta,\tau)$ is given in Definition 3.4.

As verified in Subsection 5.2, each $\rho_i(\varepsilon,\kappa)$ is a stopping time. Moreover, in our setting the sum on the right-hand side of (3.8) is a finite sum a.s. as a consequence of Proposition 5.3 below. Hence, by adjusting this sum on a set of probability zero, we may assume that $A^{\text{adj}}(\vartheta,\tau|\varepsilon,\kappa) \in \text{CL}_0([0,T])$. 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, the strategy ϑ does not necessarily have the left-limit at T, and secondly, 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 in our context.

To formulate main results in this section, we need to modify the weight processes. For $\Phi \in \text{CL}^+([0,T])$ and $t\in[0,T]$, we define

$$\overline{\Phi}_t := \Phi_t + \sup\nolimits_{s\in[0,t]}|\Delta\Phi_s|. \tag{3.9}$$

The reason to consider $\overline{\Phi}$ is that in the calculation below we will end up with Φ_- which is not càdlàg and therefore is not a candidate for a weight process. For $\overline{\Phi}$, it is clear that

$\overline{\Phi} \in \mathrm{CL}^+([0,T])$ with $\Phi \vee \Phi_- \leqslant \overline{\Phi}$, and $\Phi \equiv \overline{\Phi}$ if and only if Φ is continuous. Moreover, Proposition 7.1(2) shows that $\Phi \in \mathcal{SM}_p(\mathbb{P})$ implies $\overline{\Phi} \in \mathcal{SM}_p(\mathbb{P})$.

Theorem 3.10. *Let Assumption 3.3 hold for some $\theta \in (0,1]$ and let $\Phi \in \mathcal{SM}_2(\mathbb{P})$.*

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

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

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

$$\left\|E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(3.11)} \max\left\{\varepsilon^{1-\alpha}\sqrt{\|\tau\|_\theta},\ \sqrt{\|\tau\|_\theta},\ \varepsilon\right\}. \tag{3.11}$$

(2) *If there is a constant $c_{(3.12)} > 0$ such that for $\mathbb{P}\otimes\lambda$-a.e. $(\omega,t) \in \Omega\times[0,T]$,*

$$\sup_{r>0}\left|\int_{|z|>r} z\nu_t(\omega,\mathrm{d}z)\right| \leqslant c_{(3.12)}, \tag{3.12}$$

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

$$\left\|E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(3.13)} \max\left\{\sqrt{\|\tau\|_\theta},\ \varepsilon\right\}. \tag{3.13}$$

Minimizing the right-hand side of (3.11) (resp. (3.13)) over $\varepsilon > 0$ leads us to the selection $\varepsilon = \sqrt[2\alpha]{\|\tau\|_\theta}$ (resp. $\varepsilon = \sqrt{\|\tau\|_\theta}$). Then, we have the following:

Corollary 3.11. *Let Assumption 3.3 hold for some $\theta \in (0,1]$ and let $\Phi \in \mathcal{SM}_2(\mathbb{P})$.*

(1) *If (3.10) is satisfied for some $\alpha \in [1,2]$, then*

$$\left\|E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \sqrt[2\alpha]{\|\tau\|_\theta}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant T^{\frac{\theta}{2}(1-\frac{1}{\alpha})} c_{(3.11)} \sqrt[2\alpha]{\|\tau\|_\theta}.$$

(2) *If (3.12) is satisfied, then*

$$\left\|E^{\mathrm{adj}}\left(\vartheta, \tau \middle| \sqrt{\|\tau\|_\theta}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(3.13)} \sqrt{\|\tau\|_\theta}.$$

Remark 3.12. (a) The assumption $K_{(2.6)} < \infty$ implies that

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

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

(b) Some obvious sufficient conditions for (3.12) are as follows: Since (3.14) holds in our setting, condition (3.12) is satisfied if (3.10) holds for $\alpha = 1$, or there is an $r_0 > 0$ such that the measure $\nu_t(\omega,\cdot)$ is symmetric on $(-r_0, r_0)$ for $\mathbb{P}\otimes\lambda$-a.e. $(\omega,t) \in \Omega\times[0,T]$.

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

Adapted time-net. The conclusions in Theorem 3.5, Corollaries 3.6 and 3.11 assert that the errors measured in $\mathrm{bmo}_2^{\Phi}(\mathbb{P})$ or $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$ are up to multiplicative constants upper bounded by $\|\tau\|_\theta^r$ with $r \in [\frac{1}{4}, \frac{1}{2}]$. Assume $\tau_n \in \mathcal{T}_{\det}$ with $\#\tau_n = n+1$, where $n \geqslant 1$ can be regarded as a parameter that controls the complexity of the approximation schemes. If one uses the equidistant nets $\tau_n = (T\frac{i}{n})_{i=0}^n$, then $\|\tau_n\|_\theta = \frac{T^\theta}{n^\theta}$, 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 $\|\tau_n\|_\theta \geqslant \frac{T^\theta}{n}$ for any $\tau_n \in \mathcal{T}_{\det}$ with $\#\tau_n = n+1$. Next, minimizing $\|\tau_n\|_\theta$ over $\tau_n \in \mathcal{T}_{\det}$ with $\#\tau_n = n+1$ leads us to the following adapted time-nets, which were used in [14, 15, 17, 19, 20]: For $\theta \in (0,1]$ and $n \geqslant 1$, the *adapted time-net* $\tau_n^\theta = (t_{i,n}^\theta)_{i=0}^n$ is defined by

$$t_{i,n}^\theta := T\left(1 - \sqrt[\theta]{1 - i/n}\right), \quad i = 1, \dots, n.$$

It is clear that the equidistant time-net corresponds to the case $\theta = 1$. By a computation, one can show that

$$\frac{T^\theta}{n} \leqslant \|\tau_n^\theta\|_\theta \leqslant \frac{T^\theta}{\theta n}. \tag{3.15}$$

Cardinality of the combined time-net. The time-net used in Theorem 3.10 is $\tau \sqcup \rho(\varepsilon, \frac{1-\theta}{2})$. Due to the randomness, a simple way to quantify the cardinality of this combined time-net is to compute its expected cardinality, i.e. $\mathbb{E}\left[\#\tau \sqcup \rho(\varepsilon, \frac{1-\theta}{2})\right]$ (see, e.g., [12]). 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 3.13. *Let $q \in [1,2]$, $r \in [2,\infty]$ with $\frac{q}{2} + \frac{1}{r} = 1$. Assume that $\mathbb{Q}$ is a probability measure absolutely continuous with respect to $\mathbb{P}$ and $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L_r(\mathbb{P})$. For $\theta \in (0,1]$ and $(\varepsilon_n)_{n\geqslant 1} \subset (0,\infty)$ with $\inf_{n\geqslant 1} \sqrt{n}\varepsilon_n > 0$, there is a constant $c_{(3.16)} > 0$ such that for any $n \geqslant 1$, $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$,*

$$\left\|\#\tau_n \sqcup \rho\left(\varepsilon_n, \tfrac{1-\theta}{2}\right)\right\|_{L_q(\mathbb{Q})} \sim_{c_{(3.16)}} n. \tag{3.16}$$

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

Theorem 3.14. *Assume that Assumption 3.3 holds for some $\theta \in (0,1]$.*

(1) *One has $\sup_{n\geqslant 1} n^{\frac{1}{2}} \|E^{\mathrm{Rm}}(\vartheta, \tau_n^\theta)\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})} < \infty$.*

(2) *If $\Phi \in \mathcal{SM}_2(\mathbb{P})$ and if (3.10) is satisfied for some $\alpha \in [1,2]$, then*

$$\sup_{n\geqslant 1} n^{\frac{1}{2\alpha}} \left\|E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \,\middle|\, n^{-\frac{1}{2\alpha}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty.$$

(3) *If $\Phi \in \mathcal{SM}_2(\mathbb{P})$ and if (3.12) is satisfied, then*

$$\sup_{n\geqslant 1} n^{\frac{1}{2}} \left\|E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \,\middle|\, n^{-\frac{1}{2}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty.$$

(4) *If in addition $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in (2,\infty)$, then the conclusions of items (2)–(3) hold for the $L_p(\mathbb{P})$-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 $\Phi \in \mathcal{SM}_2(\mathbb{Q})$, then the conclusions of items (2)–(3) hold 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 directly from combining Theorem 3.5 (resp. Theorem 3.10) with (3.15). Items (4)–(5) are due to Proposition 2.5 and Proposition 7.1(2). □

In the estimates of Theorem 3.14(1)–(4), applying Proposition 3.13 with $q = 2$, $r = \infty$ and $\mathbb{Q} = \mathbb{P}$ we find that the parameter n in front of the $\mathrm{bmo}_2^\Phi(\mathbb{P})$, $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$ or $L_p(\mathbb{P})$-norms can be regarded as the $L_2(\mathbb{P})$-norm of the cardinality of the time-net used in the corresponding approximation schemes. Regarding Theorem 3.14(5), thanks to Proposition 3.13 (choose $q = 1$, $r = 2$), if $s \in [2,\infty)$, then this observation still holds true after a change of measure: The parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{Q})$-norm can be considered as the expected cardinality of the time-net *under* $\mathbb{Q}$.

We get from Theorem 3.14(1) and (3) the convergence rate of order $n^{-1/2}$ which is *asymptotically optimal* in general (e.g., see [15, Theorem 5] in the Lévy case), while this rate is achieved in (2) for $\alpha = 1$. Furthermore, the convergence rate in (2) depends on the small jumps intensity of the underlying process Z, which is characterised by α. If we define

$$\beta^Z := \inf\left\{\alpha \in [0,2] : \left\|(\omega,t) \mapsto \textstyle\int_{|z|\leqslant 1} |z|^\alpha \nu_t(\omega, \mathrm{d}z)\right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty\right\},$$

then it follows from Theorem 3.14(2) that

$$\inf\left\{\alpha \in [1,2] : \sup_{n\geqslant 1} n^{\frac{1}{2\alpha}} \left\|E^{\mathrm{adj}}\left(\vartheta, \tau_n^\theta \,\Big|\, n^{-\frac{1}{2\alpha}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty\right\} \leqslant 1 \vee \beta^Z.$$

Notice that when Z is a Lévy process, then β^Z is the *Blumenthal–Getoor index* of Z (see [4]).

4. Applications to exponential Lévy models

We provide several examples for Assumption 3.3 in the Lévy setting so that the main results can be applied. As an important step to obtain them, we establish in Theorem 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 defined on $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $X_0 = 0$, X has independent and stationary increments and X has càdlàg paths. Let $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\in[0,T]}$ denote the augmented natural filtration of X, and we assume that $\mathcal{F} = \mathcal{F}_T^X$. According to the Lévy–Khintchine formula (see, e.g., [33, Theorem 8.1]), there is a *characteristic triplet* (γ, σ, ν), where $\gamma \in \mathbb{R}$, coefficient of Brownian component $\sigma \geqslant 0$, 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

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

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 [34]).

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 [38].

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

The SDE for S is $\mathrm{d}S_t = S_{t-}\mathrm{d}Z_t$ (eq. (6.1)), 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 [7, Proposition 8.20]), and hence all conditions in Subsection 2.3 are fulfilled.

Although results in Section 3 are stated in terms of the characteristic of Z (Theorems 3.10 and 3.14), main results in this section are formulated involving the characteristic of the log price process X which is slightly more convenient to verify in practice. Thanks to Remark 6.1, we can easily translate conditions imposed on X to Z (and vice versa). Especially, the equivalence between small jump behavior of X and Z is given in (6.3).

Galtchouk–Kunita–Watanabe (GKW) decomposition. We now turn to the quadratic hedging problem. Under Assumption 4.1, any $\xi \in L_2(\mathbb{P})$ admits the *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, $L^\xi = (L_t^\xi)_{t\in[0,T]}$ is an $L_2(\mathbb{P})$-martingale with zero mean and is strongly orthogonal to S, i.e. $\langle S, L^\xi\rangle = 0$. The integrand θ^ξ is called the *MVH strategy* corresponding to ξ, which is unique in $L_2(\mathbb{P}\otimes\lambda, \Omega\times[0,T])$. The reader is referred to [34] 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 the main tasks for us is to find a representation of θ^ξ which is convenient for verifying the conditions in Assumption 3.3. This issue is handled in Subsection 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., [25]) 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., [5, 23, 36]). 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., [3, 28]), 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, the result 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. The proof is given in Section 6 by exploiting Malliavin calculus. Recall that σ and ν are the coefficient of the Brownian component and the Lévy measure of X respectively.

Theorem 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 $t \in (0,T)$, 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 $c_{(4.2)} := (\sigma^2 + \int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x))^{1/2}$ 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.*

Formula (4.2) was also established in [8, Section 4] and in [36, Proposition 7] under some extra conditions for g and S. A similar formula of (4.2) in a general setting can be found in [25, Theorem 2.4].

Assumption 4.1 ensures that $c_{(4.2)} \in (0,\infty)$. For the case (3a), due to the presence of the Gaussian component of X, the function $G(t,\cdot)$ has derivatives of all orders on $\mathbb{R}_+$ (see [18, Example 8.18]).

4.4. **Growth of the MVH strategy and weight process regularity.**

Hölder spaces and α-stable-like processes. Let $\emptyset \neq U \subseteq \mathbb{R}$ be an open interval.

Definition 4.3. (1) Let $\eta \in [0,1]$. For a Borel function $f\colon U \to \mathbb{R}$, we define

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

where $\inf \emptyset := \infty$. For $\eta \in (0,1]$, the space $C^{0,\eta}(U)$ of all η-Hölder continuous functions on U is the set of all f with $|f|_{C^{0,\eta}(U)} < \infty$. For $\eta = 0$, the space $C^{0,0}(U)$ consists of all bounded Borel functions on U.

(2) For $q \in [1,\infty]$, we define

$$\mathring{W}^{1,q}(U) := \left\{ f\colon U \to \mathbb{R} : \exists k \in L_q(U),\ f(y) - f(x) = \int_x^y k(u)\mathrm{d}u, \forall x, y \in U, x < y \right\},$$

and let $|f|_{\mathring{W}^{1,q}(U)} := \|k\|_{L_q(U)}$.

For $q \in [1,\infty]$, Hölder's inequality yields the embedding $\mathring{W}^{1,q}(U) \subseteq C^{0,\eta}(U)$, where $\eta = 1 - \frac{1}{q}$, with $|f|_{C^{0,\eta}(U)} \leqslant |f|_{\mathring{W}^{1,q}(U)}$ for $f \in \mathring{W}^{1,q}(U)$. In particular, $\mathring{W}^{1,\infty}(U) = C^{0,1}(U)$, which is the collection of Lipschitz functions on U.

We next introduce some classes of α-stable-like Lévy measures.

Definition 4.4. Let ν be a Lévy measure and $\alpha \in (0,2)$.

(1) We let $\nu \in \mathcal{S}_1(\alpha)$ if one can decompose $\nu = \nu_1 + \nu_2$, where ν_1, ν_2 are Lévy measures that satisfy

$$\limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1 - \cos(ux))\nu_2(\mathrm{d}x) < \infty,$$
$$\nu_1(\mathrm{d}x) = \frac{k(x)}{|x|^{\alpha+1}} \mathbb{1}_{\{x\neq 0\}}\mathrm{d}x,$$

where $0 < \liminf_{x\to 0} k(x) \leqslant \limsup_{x\to 0} k(x) < \infty$, and the function $x \mapsto \frac{k(x)}{|x|^\alpha}$ is non-decreasing on $(-\infty, 0)$ and non-increasing on $(0,\infty)$.

(2) We let $\nu \in \mathcal{S}_2(\alpha)$ if

$$0 < \liminf_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1 - \cos(ux))\nu(\mathrm{d}x) \leqslant \limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1 - \cos(ux))\nu(\mathrm{d}x) < \infty.$$

Remark 4.5. Let $\alpha \in (0,2)$.

(a) One has $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$. Indeed, for $\nu \in \mathcal{S}_1(\alpha)$ with the decomposition $\nu = \nu_1 + \nu_2$, a computation shows that $\nu_1 \in \mathcal{S}_2(\alpha)$. Hence, $\nu \in \mathcal{S}_2(\alpha)$. Moreover, since $\nu(\mathrm{d}x) := x^{-1-\alpha}\mathbb{1}_{(0,1)}(x)\mathrm{d}x$ belongs to $\mathcal{S}_2(\alpha)\backslash\mathcal{S}_1(\alpha)$, the inclusion $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$ is strict.

(b) According to [4, Theorem 3.2], if $\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in (0,2)$, then α is equal to the *Blumenthal–Getoor index* of ν, i.e. $\alpha = \inf\{r \in [0,2] : \int_{|x|\leqslant 1} |x|^r \nu(\mathrm{d}x) < \infty\}$.

For $\eta \in [0,1]$, define processes $\Theta(\eta), \Phi(\eta) \in \mathrm{CL}^+([0,T])$ by setting

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

As mentioned earlier, Assumption 3.3 is crucial to obtain the main results in Section 3, and now we provide examples for Assumption 3.3 in the exponential Lévy setting.

Theorem 4.6. *Assume Assumption 4.1. Let $\eta \in [0,1]$. Then, the following assertions hold:*

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

(2) *(MVH strategy growth) If $g \in C^{0,\eta}(\mathbb{R}_+)$, then there exist a $\hat{\theta} \in [0,1]$ and a constant $c_{(4.4)} > 0$, which might depend on $\hat{\theta}$, such that for ϑ^g given in (4.2) one has*

$$|\vartheta_t^g| \leqslant c_{(4.4)}(T-t)^{\frac{\hat{\theta}-1}{2}}\Theta(\eta)_t \quad a.s., \ \forall t \in [0,T), \tag{4.4}$$

where $\hat{\theta}$ is provided in Table 1:

Table 1: Values of $\hat{\theta}$

	σ *and* η	*Small jump condition for* X	*Regularity of* g	*Conclusion for* $\hat{\theta}$
C1	$\sigma > 0$ $\eta \in [0,1]$		$g \in C^{0,\eta}(\mathbb{R}_+)$	$\hat{\theta} = \eta$
C2	$\sigma = 0$ $\eta \in [0,1]$	$\int_{\lvert x\rvert \leqslant 1} \lvert x\rvert^{1+\eta}\nu(\mathrm{d}x) < \infty$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\hat{\theta} = 1$
C3	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_1(\alpha)$ *for some* $\alpha \in [1+\eta, 2)$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\forall\hat{\theta} \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$
C4	$\sigma = 0$ $\eta \in [0,1)$	$\nu \in \mathcal{S}_2(\alpha)$ *for some* $\alpha \in [1+\eta, 2)$	$g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$	$\forall\hat{\theta} \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$

(3) *Denote $M := \vartheta^g S$. Then, Assumption 3.3 is satisfied for*

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

and for $\theta = 1$ if $\hat{\theta} = 1$, and for any $\theta \in (0, \hat{\theta})$ if $\hat{\theta} \in (0,1)$.

Proof. We recall from Assumption 4.1 that $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$.

(1) follows from Proposition 7.2.

(2) We let $\ell := \nu$ in (8.12) and obtain from (4.2) that

$$\vartheta_t^g = c_{(4.2)}^{-2}\mathit{\Gamma}_\nu(T-t, S_t) \quad \text{a.s., } \forall t \in [0,T).$$

We consider each case in Table 1 as follows. We apply Proposition 8.6(1) to get C1. The case C2 follows from Proposition 8.6(2). For C3, since $\nu \in \mathcal{S}_1(\alpha)$, Remark 4.5(b) implies that $0 < \int_{|x|\leqslant 1}|x|^{\alpha+\varepsilon}\nu(\mathrm{d}x) < \infty$ for any $\varepsilon \in (0, 2-\alpha]$. Moreover, applying Proposition 8.6(3) and Remark 8.7 with $\beta = \alpha + \varepsilon$ yields

$$|\vartheta_t^g| \leqslant c(\varepsilon)(T-t)^{\frac{\eta+1}{\alpha}-1-\frac{\varepsilon}{\alpha}}S_t^{\eta-1} \leqslant c(\varepsilon)(T-t)^{\frac{1}{2}\left(\left(\frac{2(\eta+1)}{\alpha}-1-\frac{2\varepsilon}{\alpha}\right)-1\right)}\Theta(\eta)_t \text{ a.s., } \forall t \in [0,T),$$

where $c(\varepsilon) > 0$ is a constant depending on ε. Since $\varepsilon > 0$ can be arbitrarily small, C3 follows. The case C4 is similar to C3 where we use Proposition 8.6(4) and Remark 8.7.

(3) Due to Theorem 4.2(2), M is an $L_2(\mathbb{P})$-martingale. Then, Assumption 3.1 holds because of Example 3.2. We now only need to check (3.3). If $\hat{\theta} = 1$, then the martingale M is closed 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.,

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

If $\hat\theta \in (0,1)$, then for any $\theta \in (0,\hat\theta)$ and any $a \in [0,T)$ one has, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta} M_t^2 \mathrm{d}t\right] \leqslant c^2_{(4.4)} T^{\hat\theta-\theta+1} \|\Phi(\eta)\|^2_{\mathcal{SM}_2(\mathbb{P})} \Phi(\eta)^2_a. \tag{4.5}$$

We apply conditional Itô's isometry and [18, Proposition 3.8] to obtain that, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta} \mathrm{d}\langle M\rangle_t\right] &= \lim_{b\uparrow T} \mathbb{E}^{\mathcal{F}_a}\left[\left|\int_{(a,b]} (T-t)^{\frac{1-\theta}{2}} \mathrm{d}M_t\right|^2\right]\\ &\leqslant (1-\theta)\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{-\theta} M_t^2 \mathrm{d}t\right]\\ &\leqslant (1-\theta) c^2_{(4.4)} \|\Phi(\eta)\|^2_{\mathcal{SM}_2(\mathbb{P})} \Phi(\eta)^2_a \int_{(a,T)} (T-t)^{\hat\theta-\theta-1}\mathrm{d}t\\ &\leqslant \frac{T^{\hat\theta-\theta}}{\hat\theta-\theta}(1-\theta) c^2_{(4.4)} \|\Phi(\eta)\|^2_{\mathcal{SM}_2(\mathbb{P})} \Phi(\eta)^2_a. \end{aligned} \tag{4.6}$$

Combining (4.5) with (4.6) yields the desired conclusion. □

Remark 4.7. The larger $\hat\theta$ is, the better estimate one can get for ϑ^g in (4.4). Furthermore, the parameter $\hat\theta$ comes from the interplay between the small jump intensity of the underlying Lévy process and the regularity of the payoff function which affects the convergence rate of the approximation error.

5. Proofs of the main results

5.1. Proofs of results in Subsection 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 $0 \leqslant a < b \leqslant T$, a.s.,*

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

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

Proof. Fix $a \in [0,T)$. For any $b \in (a,T]$, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |\sigma(S_t)-\sigma(S_a)|^2 \mathrm{d}t\right] \leqslant |\sigma|^2_{\mathrm{Lip}}\, \mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |S_t-S_a|^2\mathrm{d}t\right]\\ &= |\sigma|^2_{\mathrm{Lip}}\, \mathbb{E}^{\mathcal{F}_a}\left[\int_a^b \left|\int_a^t \sigma(S_{u-})\left(\mathrm{d}Z^{\mathrm{c}}_u + \int_{\mathbb{R}_0} z(N_Z-\pi_Z)(\mathrm{d}u,\mathrm{d}z)\right) + \int_a^t \sigma(S_{u-})V_u \mathrm{d}u\right|^2 \mathrm{d}t\right]\\ &\leqslant 2|\sigma|^2_{\mathrm{Lip}}\, \mathbb{E}^{\mathcal{F}_a}\left[\int_a^b \left(\int_a^t \sigma(S_{u-})^2 K_u^2 \mathrm{d}u + \int_a^t V_u^2\mathrm{d}u \int_a^t \sigma(S_{u-})^2\mathrm{d}u\right)\mathrm{d}t\right]\\ &\leqslant c^2_{(5.3)}\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b\int_a^t \sigma(S_u)^2\mathrm{d}u\mathrm{d}t\right], \end{aligned} \tag{5.3}$$

where in order to obtain the second inequality we use the conditional Itô isometry for the martingale term and apply Hölder's inequality for the finite variation term. The last inequality comes from the fact that $t \mapsto \sigma(S_t)$ has at most countable discontinuities, and

$$c^2_{(5.3)} := 2|\sigma|^2_{\mathrm{Lip}}(K^2_{(2.6)} + V^2_{(2.6)}). \tag{5.4}$$

Then, the triangle inequality implies that, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b \sigma(S_t)^2\mathrm{d}t\right] &\leqslant 2(b-a)\sigma(S_a)^2 + 2\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |\sigma(S_t)-\sigma(S_a)|^2\mathrm{d}t\right]\\ &\leqslant 2(b-a)\sigma(S_a)^2 + 2c^2_{(5.3)}\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b\int_a^t \sigma(S_u)^2\mathrm{d}u\mathrm{d}t\right].\end{aligned}$$

Now, for any $A \in \mathcal{F}_a$, it holds that

$$\int_a^b \mathbb{E}\mathbb{1}_A\sigma(S_t)^2\mathrm{d}t \leqslant 2(b-a)\mathbb{E}\mathbb{1}_A\sigma(S_a)^2 + 2c^2_{(5.3)}\int_a^b\int_a^t \mathbb{E}\mathbb{1}_A\sigma(S_u)^2\mathrm{d}u\mathrm{d}t.$$

Since $\mathbb{E}\int_0^T \sigma(S_u)^2\mathrm{d}u < \infty$ due to (2.7), using Gronwall's inequality yields

$$\int_a^b \mathbb{E}\mathbb{1}_A\sigma(S_t)^2\mathrm{d}t \leqslant 2(b-a)\mathbb{E}\mathbb{1}_A\sigma(S_a)^2\mathrm{e}^{2c^2_{(5.3)}(b-a)},$$

which verifies (5.1) with $c^2_{(5.1)} := 2\mathrm{e}^{2c^2_{(5.3)}T}$. In order to obtain (5.2), we apply (5.1) to the right-hand side of (5.3), and then we can let

$$c^2_{(5.2)} = \tfrac{1}{2}c^2_{(5.1)}c^2_{(5.3)} = c^2_{(5.3)}\mathrm{e}^{2c^2_{(5.3)}T}. \tag{5.5}$$

□

Proof of Theorem 3.5. For $\vartheta \in \Sigma_S^{\mathrm{adm}}$ 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 non-decreasing paths on $[0,T]$, by

$$\langle\vartheta,\tau\rangle_t := \sum_{i=1}^n \int_{t_{i-1}\wedge t}^{t_i\wedge t} \left|\vartheta_u - \vartheta_{t_{i-1}}\right|^2 \sigma(S_u)^2\mathrm{d}u. \tag{5.6}$$

From (5.7) below we see that $\langle\vartheta,\tau\rangle$ is "nearly" the predictable quadratic variation of $E^{\mathrm{Rm}}(\vartheta;\tau)$ (this is the reason for (slightly abusively) using angle brackets in the notation $\langle\vartheta,\tau\rangle$) which is known as a useful tool for studying $E^{\mathrm{Rm}}(\vartheta;\tau)$ in the mean square sense.

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

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\left[|E_T^{\mathrm{Rm}}(\vartheta,\tau) - E_a^{\mathrm{Rm}}(\vartheta,\tau)|^2\right]\\ &\leqslant 2\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\vartheta_{u-} - \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma(S_{u-})^2\left(K_u^2 + \int_a^T V_r^2\mathrm{d}r\right)\mathrm{d}u\right]\\ &\leqslant 2(K^2_{(2.6)} + V^2_{(2.6)})\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\vartheta_{u-} - \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma(S_{u-})^2\mathrm{d}u\right]\\ &= 2(K^2_{(2.6)} + V^2_{(2.6)})\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\vartheta_{u} - \sum_{i=1}^n \vartheta_{t_{i-1}}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma(S_{u})^2\mathrm{d}u\right]\\ &= 2(K^2_{(2.6)} + V^2_{(2.6)})\mathbb{E}^{\mathcal{F}_a}\left[\langle\vartheta,\tau\rangle_T - \langle\vartheta,\tau\rangle_a\right],\end{aligned} \tag{5.7}$$

where the first equality comes from the fact that the number of discontinuities of a càdlàg function is at most countable and $\vartheta \in \Sigma_S^{\mathrm{adm}}$ has no fixed-time discontinuity. We 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, Theorem 3.5 is a direct consequence of (5.7) and the following result.

Proposition 5.2. *Let Assumption 3.3 hold for some $\theta \in (0,1]$. Then, there exists a constant $c_{(5.8)} > 0$ such that for any $\tau \in \mathcal{T}_{\mathrm{det}}$ and any $a \in [0,T)$, a.s.,*

$$\mathbb{E}^{\mathcal{F}_a}[\langle \vartheta, \tau\rangle_T - \langle \vartheta, \tau\rangle_a] \leqslant c^2_{(5.8)} \|\tau\|_\theta \Phi_a^2. \tag{5.8}$$

Consequently, $\|\langle \vartheta, \tau\rangle\|_{\mathrm{BMO}_1^{\Phi^2}(\mathbb{P})} \leqslant c^2_{(5.8)}\|\tau\|_\theta$.

Proof. By the monotonicity of Θ and (3.2), we have that for $c_{(5.9)} := \sqrt{2}c_{(3.2)}$ and for any $0 \leqslant s < t < T$, a.s.,

$$|\vartheta_t - \vartheta_s|^2 \sigma(S_t)^2 \leqslant c^2_{(5.9)} \left((T-t)^{\theta-1} + (T-s)^{\theta-1}\right)\Phi_t^2. \tag{5.9}$$

We aim to apply [18, Theorem 4.3] to obtain (5.8). To do this, let us define the random measure

$$\Pi(\omega, \mathrm{d}t) := \sigma(S_t(\omega))^2 \mathrm{d}t, \quad \omega \in \Omega.$$

Then, it is clear that $\Pi(\omega, (0,t]) < \infty$ for any $(\omega, t) \in \Omega \times (0,T)$. For $0 \leqslant s \leqslant a < b < T$, the triangle inequality yields, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_a}&\left[\int_{(a,b]} |\vartheta_t - \vartheta_s|^2 \Pi(\cdot, \mathrm{d}t)\right] = \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |\vartheta_t - \vartheta_s|^2 \sigma(S_t)^2 \mathrm{d}t\right]\\
&\leqslant 2\mathbb{E}^{\mathcal{F}_a}\left[|\vartheta_a - \vartheta_s|^2 \int_{(a,b]} \sigma(S_t)^2\mathrm{d}t + \int_{(a,b]} |\vartheta_t - \vartheta_a|^2 \sigma(S_t)^2 \mathrm{d}t\right]\\
&\leqslant 2\mathbb{E}^{\mathcal{F}_a}\left[|\vartheta_a - \vartheta_s|^2 \Pi(\cdot, (a,b]) + c^2_{(3.1)} \int_{(a,b]} (b-t)\Upsilon(\cdot, \mathrm{d}t)\right].
\end{aligned}$$

Let $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ and $a \in [t_{k-1}, t_k)$ for $k \in [1,n]$. Applying [18, Theorem 4.3] yields a constant $c > 0$ independent of τ and a such that, a.s.,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{F}_a}[\langle \vartheta, \tau\rangle_T - \langle \vartheta, \tau\rangle_a]\\
&\leqslant c\|\tau\|_\theta \left(\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta}\Upsilon(\cdot, \mathrm{d}t) + \frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}} |\vartheta_a - \vartheta_{t_{k-1}}|^2 \int_{(a,t_k]} \sigma(S_t)^2 \mathrm{d}t\right]\right)\\
&\leqslant c\|\tau\|_\theta \left(c^2_{(3.3)}\Phi_a^2 + c^2_{(5.1)} \frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}} (t_k - a)|\vartheta_a - \vartheta_{t_{k-1}}|^2 \sigma(S_a)^2\right)\\
&\leqslant c\|\tau\|_\theta \left(c^2_{(3.3)} + c^2_{(5.1)} c^2_{(5.9)} \frac{(T-t_{k-1})^{1-\theta}}{t_k - t_{k-1}} (t_k - a)\left((T-a)^{\theta-1} + (T-t_{k-1})^{\theta-1}\right)\right)\Phi_a^2\\
&\leqslant c\|\tau\|_\theta (c^2_{(3.3)} + 2c^2_{(5.1)}c^2_{(5.9)})\Phi_a^2,
\end{aligned}$$

which implies (5.8) with $c^2_{(5.8)} = c(c^2_{(3.3)} + 2c^2_{(5.1)}c^2_{(5.9)})$. For the "Consequently" part, since $\langle \vartheta, \tau\rangle$ is continuous, it holds that $\|\langle \vartheta, \tau\rangle\|_{\mathrm{BMO}_1^{\Phi^2}(\mathbb{P})} = \|\langle \vartheta, \tau\rangle\|_{\mathrm{bmo}_1^{\Phi^2}(\mathbb{P})} \leqslant c^2_{(5.8)}\|\tau\|_\theta$. ☐

5.2. **Proofs of results in Subsections 3.2 and 3.3.** We let $\varepsilon > 0$, $\kappa \geqslant 0$ and recall $\rho(\varepsilon, \kappa) = (\rho_i(\varepsilon, \kappa))_{i \geqslant 0}$ in Definition 3.8. Due to (3.5) and the assumption $\sigma(S_-) > 0$, it holds that

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

Hence, we derive from (3.6) the relations

$$\rho_i(\varepsilon, \kappa) = \inf\{T \geqslant t > \rho_{i-1}(\varepsilon, \kappa) : |\Delta Z_t| > \varepsilon(T-t)^\kappa\} \wedge T, \ i \geqslant 1. \tag{5.10}$$

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 $1 \leqslant i \leqslant \mathcal{N}_{(3.7)}(\varepsilon,\kappa)$.

For a non-negative Borel function h defined on $\mathbb{R}$, denote

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

Then, condition (3.14) is re-written as

$$\|z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} < \infty. \tag{5.11}$$

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

$$\|\mathcal{N}_{(3.7)}(\varepsilon,\kappa)\|_{L_2(\mathbb{P})} \leqslant 1 + \sqrt{C_{(5.13)}} + C_{(5.13)}, \tag{5.12}$$

where

$$C_{(5.13)} := T\|\mathbb{1}_{\{|z|>1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} + \varepsilon^{-\alpha}\frac{T^{1-\alpha\kappa}}{2^{1-\alpha\kappa}-1}\|\mathbb{1}_{\{|z|\leqslant 1\}}|z|^\alpha \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}. \tag{5.13}$$

Proof. We may assume that $C_{(5.13)} < \infty$, otherwise the desired inequality is trivial.

Step 1. We show that, a.s.,

$$\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}\pi_Z(\mathrm{d}t,\mathrm{d}z) \leqslant C_{(5.13)}.$$

One decomposes

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

where the first term in the right-hand side is upper bounded by $T\|\mathbb{1}_{\{|z|>1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}$ a.s. The second term can be estimated as follows, a.s.,

$$\begin{aligned}
\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{1\geqslant|z|>\varepsilon(T-t)^\kappa\}}\pi_Z(\mathrm{d}t,\mathrm{d}z) &= \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{1\geqslant|z|>\varepsilon(T-t)^\kappa\}}\nu_t(\mathrm{d}z)\mathrm{d}t\\
&\leqslant \sum_{n=0}^\infty \int_{T(1-\frac{1}{2^n})}^{T(1-\frac{1}{2^{n+1}})}\int_{\mathbb{R}} \mathbb{1}_{\{1\geqslant|z|>\varepsilon(T/2^{n+1})^\kappa\}}\nu_t(\mathrm{d}z)\mathrm{d}t\\
&\leqslant \sum_{n=0}^\infty \int_{T(1-\frac{1}{2^n})}^{T(1-\frac{1}{2^{n+1}})}\int_{|z|\leqslant 1} \left(\frac{|z|}{\varepsilon(T/2^{n+1})^\kappa}\right)^\alpha \nu_t(\mathrm{d}z)\mathrm{d}t\\
&= \varepsilon^{-\alpha}T^{-\alpha\kappa}\sum_{n=0}^\infty (2^{n+1})^{\alpha\kappa}\int_{T(1-\frac{1}{2^n})}^{T(1-\frac{1}{2^{n+1}})}\int_{|z|\leqslant 1} |z|^\alpha\nu_t(\mathrm{d}z)\mathrm{d}t\\
&\leqslant \varepsilon^{-\alpha}T^{1-\alpha\kappa}\sum_{n=0}^\infty (2^{n+1})^{\alpha\kappa-1}\|\mathbb{1}_{\{|z|\leqslant 1\}}|z|^\alpha \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}\\
&= \varepsilon^{-\alpha}\frac{T^{1-\alpha\kappa}}{2^{1-\alpha\kappa}-1}\|\mathbb{1}_{\{|z|\leqslant 1\}}|z|^\alpha \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}.
\end{aligned}$$

Step 2. Combining *Step 1* with [26, Ch.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

$\mathcal{N}_{(3.7)}(\varepsilon,\kappa) \leqslant 1 + \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} N_Z(\mathrm{d}t,\mathrm{d}z)$ by (5.10), we have

$$
\begin{aligned}
&\|\mathcal{N}_{(3.7)}(\varepsilon,\kappa)\|_{L_2(\mathbb{P})} \leqslant 1 + \left\| \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} N_Z(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})}\\
&\leqslant 1 + \left\| \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} (N_Z-\pi_Z)(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})} + \left\| \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} \pi_Z(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})}\\
&= 1 + \left\| \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} \pi_Z(\mathrm{d}t,\mathrm{d}z)\right\|_{L_1(\mathbb{P})}^{\frac12} + \left\| \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}} \pi_Z(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})}\\
&\leqslant 1 + \sqrt{C_{(5.13)}} + C_{(5.13)},
\end{aligned}
$$

where one uses [26, Ch.II, Theorem 1.33(a)] to derive the equality. □

5.2.1. *Proof of Proposition 3.13.* Denote $\kappa := \frac{1-\theta}{2} \in [0,\frac12)$. We first consider the particular case when $\mathbb{Q}=\mathbb{P}$, $r=\infty$ and $q=2$. By Definition 3.9(1),

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

Thus,

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

In (5.13), substituting $\alpha=2$ and using $\inf_{n\geqslant 1}\sqrt{n}\varepsilon_n > 0$, we obtain

$$
C_{(5.13)} = T\|\mathbb{1}_{\{|z|>1\}}\star\nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} + \varepsilon_n^{-2}\frac{T^{1-2\kappa}}{2^{1-2\kappa}-1}\|\mathbb{1}_{\{|z|\leqslant 1\}}z^2\star\nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} \leqslant cn
$$

for a constant $c>0$ independent of n. Using (5.12) gives the desired conclusion.

In the next step we assume that $\mathbb{Q}\ll\mathbb{P}$ is a probability measure 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

$$
\|\#\tau_n \sqcup \rho(\varepsilon_n,\kappa)\|_{L_q(\mathbb{Q})} \leqslant \|\#\tau_n \sqcup \rho(\varepsilon_n,\kappa)\|_{L_2(\mathbb{P})}\,\|\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}\|_{L_r(\mathbb{P})}^{\frac1q},
$$

and hence (3.16) follows. □

5.2.2. *Proof of Theorem 3.10.* Again, we denote $\kappa := \frac{1-\theta}{2}\in[0,\frac12)$.

(1) *Step 1.* We handle the correction term in (3.8) and the corresponding error process. For $\varepsilon>0$, by the same arguments as in the proof of *Step 1* of Proposition 5.3, one has

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

where the finiteness holds due to (5.11). This allows us to decompose

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

where

$$
Z^{\leqslant} := \int_0^\cdot\int_{\mathbb{R}_0} z\mathbb{1}_{\{|z|\leqslant\varepsilon(T-u)^\kappa\}}(N_Z-\pi_Z)(\mathrm{d}u,\mathrm{d}z),
$$

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

Recall ϑ^τ in Definition 3.9. Since (5.11) holds in our context, applying Proposition 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 $t \in [0,T]$,

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

By the representation of Z in (2.4), one can decompose

$$\begin{aligned} \mathrm{d}S_t = \sigma(S_{t-})\mathrm{d}Z_t &= \sigma(S_{t-})\left(\mathrm{d}Z^{\mathrm{c}}_t + V_t\mathrm{d}t + \int_{\mathbb{R}_0} z(N_Z-\pi_Z)(\mathrm{d}t,\mathrm{d}z)\right)\\ &= \sigma(S_{t-})\left(\mathrm{d}Z^{\mathrm{c}}_t + V_t\mathrm{d}t + \mathrm{d}Z^{\leqslant}_t + \mathrm{d}Z^{>}_t - \mathrm{d}\gamma^{>}_t\right). \end{aligned}$$

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

$$\begin{aligned} &E^{\mathrm{adj}}(\vartheta,\tau|\varepsilon,\kappa) = \int_0^{\cdot}(\vartheta_{u-}-\vartheta^\tau_u)\mathrm{d}S_u - \sum_{\rho_i(\varepsilon,\kappa)\in[0,\cdot]\cap[0,T)} \left(\vartheta_{\rho_i(\varepsilon,\kappa)-} - \vartheta^\tau_{\rho_i(\varepsilon,\kappa)}\right)\Delta S_{\rho_i(\varepsilon,\kappa)}\\ &= \int_0^{\cdot}(\vartheta_{u-}-\vartheta^\tau_u)\sigma(S_{u-})(\mathrm{d}Z^{\mathrm{c}}_u + V_u\mathrm{d}u + \mathrm{d}Z^{\leqslant}_u + \mathrm{d}Z^{>}_u - \mathrm{d}\gamma^{>}_u) - \int_0^{\cdot}(\vartheta_{u-}-\vartheta^\tau_u)\sigma(S_{u-})\mathrm{d}Z^{>}_u\\ &= \int_0^{\cdot}(\vartheta_{u-}-\vartheta^\tau_u)\sigma(S_{u-})(\mathrm{d}Z^{\mathrm{c}}_u + V_u\mathrm{d}u + \mathrm{d}Z^{\leqslant}_u)\\ &\qquad - \int_0^{\cdot}(\vartheta_{u-}-\vartheta^\tau_u)\sigma(S_{u-})\int_{\mathbb{R}} z\mathbb{1}_{\{|z|>\varepsilon(T-u)^\kappa\}}\nu_u(\mathrm{d}z)\mathrm{d}u. \end{aligned} \tag{5.14}$$

Let us define the error processes $E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)$ induced from the "small jumps" part and $E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)$ involved with the "drift" part by

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

The triangle inequality applied to (5.14) gives

$$\|E^{\mathrm{adj}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant \|E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} + \|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})}. \tag{5.15}$$

Step 2. We investigate the right-hand side of (5.15).

Step 2.1. We consider $E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)$. Since $\Phi \in \mathcal{SM}_2(\mathbb{P})$ by assumption, Proposition 7.1(2) implies that $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$. Then, Lemma 2.3 asserts

$$\|E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \sim_{c_{(5.16)}} \|E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})} + |\Delta E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)|_{\overline{\Phi}}. \tag{5.16}$$

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 (5.9) holds for all $(\omega,t)\in\Omega_0\times[0,T)$. This implies that, for all $0\leqslant s<t<T$ and $\omega\in\Omega_0$,

$$|\vartheta_t-\vartheta_s|\sigma(S_t)\leqslant\sqrt{2}c_{(5.9)}(T-t)^{-\kappa}\Phi_t.$$

Due to (2.5), one has $\pi_Z(\omega,\{t\}\times\mathbb{R}_0)=0$ for any $(\omega,t)\in\Omega\times[0,T]$. Then, it holds that

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

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

$$\begin{aligned}|\Delta E_t^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)|&=\left|(\vartheta_{t-}-\vartheta_t^\tau)\sigma(S_{t-})\Delta Z_t^{\leqslant}\right|\leqslant\sqrt{2}c_{(5.9)}(T-t)^{-\kappa}\Phi_{t-}\varepsilon(T-t)^\kappa\\&=\sqrt{2}c_{(5.9)}\varepsilon\Phi_{t-}\leqslant\sqrt{2}c_{(5.9)}\varepsilon\overline{\Phi}_t.\end{aligned}$$

According to the definition of $|\cdot|_{\overline{\Phi}}$ given in (2.1), one then gets

$$|\Delta E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)|_{\overline{\Phi}}\leqslant\sqrt{2}c_{(5.9)}\varepsilon. \tag{5.17}$$

Let us continue with $\|E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})}$. We apply the conditional Itô isometry for the martingale component and apply Hölder's inequality for the finite variation component of $E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)$ to derive that, for $a\in[0,T)$, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\left[|E_T^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)-E_a^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)|^2\right]\\&\leqslant2\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T|\vartheta_{u-}-\vartheta_u^\tau|^2\sigma(S_{u-})^2\left(\mathrm{d}\langle Z^{\mathrm{c}}\rangle_u+\int_{\mathbb{R}}\mathbb{1}_{\{|z|\leqslant\varepsilon(T-u)^\kappa\}}z^2\nu_u(\mathrm{d}z)\mathrm{d}u\right)\right]\\&\qquad+2\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T|\vartheta_{u-}-\vartheta_u^\tau|^2\sigma(S_{u-})^2\mathrm{d}u\int_a^TV_u^2\mathrm{d}u\right]\\&\leqslant2(K_{(2.6)}^2+V_{(2.6)}^2)\mathbb{E}^{\mathcal{F}_a}\left[\langle\vartheta,\tau\rangle_T-\langle\vartheta,\tau\rangle_a\right]\\&\leqslant2(K_{(2.6)}^2+V_{(2.6)}^2)c_{(5.8)}^2\|\tau\|_\theta\overline{\Phi}_a^2,\end{aligned}\tag{5.18}$$

where $\langle\vartheta,\tau\rangle$ is given in (5.6), and where we use the fact that a càdlàg function has at most countably many discontinuities to obtain the second inequality. Combining (5.16) and (5.17) with (5.18) yields

$$\|E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})}\leqslant c_{(5.16)}\left(\sqrt{2}c_{(5.9)}\varepsilon+\sqrt{2(K_{(2.6)}^2+V_{(2.6)}^2)}c_{(5.8)}\sqrt{\|\tau\|_\theta}\right). \tag{5.19}$$

Step 2.2. We consider $E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)$. Since $E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)$ is continuous, it holds that

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

Now, for any $a\in[0,T)$, we use Hölder's inequality to get, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\left[|E_T^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)-E_a^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)|^2\right]\\&\leqslant\mathbb{E}^{\mathcal{F}_a}\left[\left(\int_a^T\left|\int_{\mathbb{R}}z\mathbb{1}_{\{|z|>\varepsilon(T-u)^\kappa\}}\nu_u(\mathrm{d}z)\right|^2\mathrm{d}u\right)\left(\int_a^T|\vartheta_{u-}-\vartheta_u^\tau|^2\sigma(S_{u-})^2\mathrm{d}u\right)\right]\\&=:\mathbb{E}^{\mathcal{F}_a}\left[\mathrm{I}_{(5.20)}\mathrm{II}_{(5.20)}\right].\end{aligned}\tag{5.20}$$

For the first factor $\mathrm{I}_{(5.20)}$, since $\|\mathbb{1}_{\{|z|\leqslant 1\}}|z|^{\alpha}\star\nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}<\infty$ by (3.10), one has, a.s.,

$$\begin{aligned}\frac{1}{2}\mathrm{I}_{(5.20)} &\leqslant \int_a^T\left|\int_{|z|>1\vee(\varepsilon(T-u)^{\kappa})} z\nu_u(\mathrm{d}z)\right|^2\mathrm{d}u+\int_a^T\left|\int_{1\geqslant|z|>\varepsilon(T-u)^{\kappa}}|z|^{\alpha}|z|^{1-\alpha}\nu_u(\mathrm{d}z)\right|^2\mathrm{d}u\\ &\leqslant (T-a)\|\mathbb{1}_{\{|z|>1\}}|z|\star\nu\|^2_{L_\infty(\mathbb{P}\otimes\lambda)}+\varepsilon^{2(1-\alpha)}\int_a^T\left|\int_{|z|\leqslant 1}|z|^{\alpha}\nu_u(\mathrm{d}z)\right|^2(T-u)^{2\kappa(1-\alpha)}\mathrm{d}u\\ &\leqslant (T-a)\|\mathbb{1}_{\{|z|>1\}}|z|\star\nu\|^2_{L_\infty(\mathbb{P}\otimes\lambda)}+\|\mathbb{1}_{\{|z|\leqslant 1\}}|z|^{\alpha}\star\nu\|^2_{L_\infty(\mathbb{P}\otimes\lambda)}\frac{(T-a)^{2\kappa(1-\alpha)+1}}{2\kappa(1-\alpha)+1}\varepsilon^{2(1-\alpha)}\\ &\leqslant c^2_{(5.21)}(1+\varepsilon^{2(1-\alpha)}),\end{aligned}\tag{5.21}$$

where $c_{(5.21)}=c_{(5.21)}(\alpha,\kappa,T,\nu)>0$ and one notices that $2\kappa(1-\alpha)+1>0$. For the second factor $\mathrm{II}_{(5.20)}$, we apply Proposition 5.2 to obtain, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\left[\mathrm{II}_{(5.20)}\right]=\mathbb{E}^{\mathcal{F}_a}[\langle\vartheta,\tau\rangle_T-\langle\vartheta,\tau\rangle_a]\leqslant c^2_{(5.8)}\|\tau\|_\theta\Phi_a^2\leqslant c^2_{(5.8)}\|\tau\|_\theta\overline{\Phi}_a^2.$$

Hence,

$$\begin{aligned}\|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} &\leqslant \sqrt{2}c_{(5.8)}c_{(5.21)}\sqrt{1+\varepsilon^{2(1-\alpha)}}\sqrt{\|\tau\|_\theta}\\ &\leqslant \sqrt{2}c_{(5.8)}c_{(5.21)}(1+\varepsilon^{1-\alpha})\sqrt{\|\tau\|_\theta}.\end{aligned}\tag{5.22}$$

Step 3. We plug (5.19) and (5.22) into (5.15) to derive (3.11).

(2) If (3.12) holds, then $\mathrm{I}_{(5.20)}$ is upper bounded by $Tc^2_{(3.12)}$. Hence,

$$\|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})}\leqslant\sqrt{T}c_{(3.12)}c_{(5.8)}\sqrt{\|\tau\|_\theta}.\tag{5.23}$$

Combining (5.19), (5.23) with (5.15) yields (3.13). □

6. Itô's chaos expansion and proof of Theorem 4.2

6.1. **Exponential Lévy processes.** Let X be a Lévy process with characteristics (γ,σ,ν) as in Subsection 4.1. It is known that the ordinary exponential $S=\mathrm{e}^X$ can be represented as the *Doléans–Dade exponential* (or *stochastic exponential*) $\mathcal{E}(Z)$ of another Lévy process Z (see, e.g., [1, Theorem 5.1.6]). This means that $S=\mathcal{E}(Z)$ and

$$\mathrm{d}S_t=S_{t-}\mathrm{d}Z_t,\quad S_0=1.\tag{6.1}$$

Remark 6.1. (a) The path relation of X and Z is given by

$$Z_t=X_t+\frac{\sigma^2t}{2}+\sum_{0\leqslant s\leqslant t}\left(\mathrm{e}^{\Delta X_s}-1-\Delta X_s\right),\quad\forall t\in[0,T]\ \text{ a.s.},\tag{6.2}$$

which implies $\Delta Z=\mathrm{e}^{\Delta X}-1$.

(b) For the triplet $(\gamma_Z,\sigma_Z,\nu_Z)$ of Z, using [1, Theorem 5.1.6] (with the truncation function $x\mathbb{1}_{\{|x|\leqslant 1\}}$ instead of $x\mathbb{1}_{\{|x|<1\}}$) yields that $\sigma_Z=\sigma$ and

$$\nu_Z=\int_{\mathbb{R}}\mathbb{1}_{\{\mathrm{e}^x-1\in\cdot\}}\nu(\mathrm{d}x),\quad\gamma_Z=\gamma+\frac{\sigma^2}{2}+\int_{\mathbb{R}}\left((\mathrm{e}^x-1)\mathbb{1}_{\{|\mathrm{e}^x-1|\leqslant 1\}}-x\mathbb{1}_{\{|x|\leqslant 1\}}\right)\nu(\mathrm{d}x).$$

Consequently, since $\lim_{x\to 0}\frac{\mathrm{e}^x-1}{x}=1$, it holds for any $\alpha\in[0,2]$ that

$$\int_{|z|\leqslant 1}|z|^{\alpha}\nu_Z(\mathrm{d}z)<\infty\Leftrightarrow\int_{|x|\leqslant 1}|x|^{\alpha}\nu(\mathrm{d}x)<\infty.\tag{6.3}$$

(c) Let $\mathbb{F}^Z = (\mathcal{F}_t^Z)_{t\in[0,T]}$ be the augmented natural filtration induced by Z. Then, we can deduce from (6.2) that $\mathcal{F}_t^Z = \mathcal{F}_t^X$ for all $t \in [0,T]$.

Let $q \in [0,\infty)$. Since $\nu_Z((-\infty,-1]) = 0$, by change of variables we get

$$\int_{|z|>1} |z|^q \nu_Z(\mathrm{d}z) = \int_{z>1} z^q \nu_Z(\mathrm{d}z) = \int_{\mathrm{e}^x-1>1} (\mathrm{e}^x-1)^q \nu(\mathrm{d}x).$$

Using $|x+y|^q \leqslant (1\vee 2^{q-1})(|x|^q + |y|^q)$ and applying [33, Theorem 25.3] yield

$$\begin{aligned}\mathbb{E}|Z_t|^q < \infty, \forall t>0 \Leftrightarrow \int_{|z|>1} |z|^q \nu_Z(\mathrm{d}z) < \infty \Leftrightarrow \int_{|x|>1} \mathrm{e}^{qx}\nu(\mathrm{d}x) < \infty \\ \Leftrightarrow \mathbb{E}\mathrm{e}^{qX_t} < \infty, \forall t>0. \end{aligned} \tag{6.4}$$

6.2. **Itô's chaos expansion.** We present briefly the Malliavin calculus for Lévy processes by means of Itô's chaos expansion. For further details, the reader can refer to [35, 30, 1] and the references therein.

We define the σ-finite measures μ on $\mathcal{B}(\mathbb{R})$ and $\mathbb{m}$ on $\mathcal{B}([0,T]\times\mathbb{R})$ by setting

$$\mu(\mathrm{d}x) := \sigma^2\delta_0(\mathrm{d}x) + x^2\nu(\mathrm{d}x) \quad \text{and} \quad \mathbb{m} := \lambda\otimes\mu,$$

where δ_0 is the Dirac measure at zero. For $B \in \mathcal{B}([0,T]\times\mathbb{R})$ with $\mathbb{m}(B) < \infty$, the random measure M is defined by

$$M(B) := \sigma\int_{\{t\in[0,T]:(t,0)\in B\}} \mathrm{d}W_t + L_2(\mathbb{P})\text{-}\lim_{n\to\infty}\int_{B\cap([0,T]\times\{\frac{1}{n}<|x|<n\})} x\widetilde{N}(\mathrm{d}t,\mathrm{d}x),$$

where W is the standard Brownian motion and $\widetilde{N}(\mathrm{d}t,\mathrm{d}x) := N(\mathrm{d}t,\mathrm{d}x) - \mathrm{d}t\nu(\mathrm{d}x)$ is the compensated Poisson random measure appearing in the Lévy–Itô decomposition of X (see [1, Theorem 2.4.16]).

Set $L_2(\mu^0) = L_2(\mathbb{m}^0) := \mathbb{R}$, and for $n \geqslant 1$ we denote

$$\begin{aligned} L_2(\mu^{\otimes n}) &:= L_2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu^{\otimes n}), \\ L_2(\mathbb{m}^{\otimes n}) &:= L_2(([0,T]\times\mathbb{R})^n, \mathcal{B}(([0,T]\times\mathbb{R})^n), \mathbb{m}^{\otimes n}). \end{aligned}$$

For $n \geqslant 1$, the multiple integral $I_n\colon L_2(\mathbb{m}^{\otimes n}) \to L_2(\mathbb{P})$ is defined by a standard approximation argument, where the multiple integral of a simple function is as follows: For

$$\xi_n^m := \sum_{k=1}^m a_k \mathbb{1}_{B_1^k\times\cdots\times B_n^k},$$

where $a_k \in \mathbb{R}$, $B_i^k \in \mathcal{B}([0,T]\times\mathbb{R})$ with $\mathbb{m}(B_i^k) < \infty$ and $B_i^k \cap B_j^k = \emptyset$ for $k = 1,\dots,m$, $i,j = 1,\dots,n$, $i \neq j$ and $m \geqslant 1$, we define

$$I_n(\xi_n^m) := \sum_{k=1}^m a_k M(B_1^k)\cdots M(B_n^k).$$

According to [24, Theorem 2], we have the following Itô chaos expansion

$$L_2(\Omega, \mathcal{F}_T^X, \mathbb{P}) = \bigoplus_{n=0}^\infty \{I_n(\xi_n) : \xi_n \in L_2(\mathbb{m}^{\otimes n})\},$$

where $I_0(\xi_0) := \xi_0 \in \mathbb{R}$. For $n \geqslant 1$, the symmetrization $\tilde{\xi}_n$ of a $\xi_n \in L_2(\mathbb{m}^{\otimes n})$ is

$$\tilde{\xi}_n((t_1,x_1),\dots,(t_n,x_n)) := \frac{1}{n!}\sum_\pi \xi_n((t_{\pi(1)},x_{\pi(1)}),\dots,(t_{\pi(n)},x_{\pi(n)})),$$

where the sum is taken over all permutations π of $\{1,\dots,n\}$. It then turns out from the definition of I_n that $I_n(\xi_n) = I_n(\tilde{\xi}_n)$ a.s. By the Itô chaos decomposition, $\xi \in L_2(\mathbb{P})$ if and only if there are $\xi_n \in L_2(\mathbb{m}^{\otimes n})$ so that $\xi = \sum_{n=0}^\infty I_n(\xi_n)$ a.s., and this expansion is unique if every ξ_n is symmetric, i.e. if $\xi_n = \tilde{\xi}_n$. Furthermore, $\|\xi\|^2_{L_2(\mathbb{P})} = \sum_{n=0}^\infty n!\|\tilde{\xi}_n\|^2_{L_2(\mathbb{m}^{\otimes n})}$.

Definition 6.2. The Malliavin–Sobolev space $\mathbb{D}_{1,2}$ consists of all $\xi = \sum_{n=0}^\infty I_n(\xi_n) \in L_2(\mathbb{P})$ satisfying

$$\|\xi\|^2_{\mathbb{D}_{1,2}} := \sum_{n=0}^\infty (n+1)!\|\tilde{\xi}_n\|^2_{L_2(\mathbb{m}^{\otimes n})} < \infty.$$

The *Malliavin derivative operator* $D\colon \mathbb{D}_{1,2} \to L_2(\mathbb{P}\otimes\mathbb{m})$, where $L_2(\mathbb{P}\otimes\mathbb{m}) := L_2(\Omega\times[0,T]\times\mathbb{R}, \mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}), \mathbb{P}\otimes\mathbb{m})$, is defined for $\xi = \sum_{n=0}^\infty I_n(\xi_n) \in \mathbb{D}_{1,2}$ by

$$D_{t,x}\xi := \sum_{n=1}^\infty nI_{n-1}(\tilde{\xi}_n((t,x),\cdot)), \quad (\omega,t,x) \in \Omega\times[0,T]\times\mathbb{R}.$$

Proposition 6.3 ([27]). *Let $t \in (0,T]$ and a Borel function $f\colon \mathbb{R}\to\mathbb{R}$ such that $f(X_t) \in L_2(\mathbb{P})$. Then, $f(X_t) \in \mathbb{D}_{1,2}$ if and only if the following two assertions hold:*

(a) *when $\sigma > 0$, f has a weak derivative[4] f'_w on $\mathbb{R}$ with $f'_w(X_t) \in L_2(\mathbb{P})$,*

(b) *the map $(s,x) \mapsto \frac{f(X_t+x)-f(X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x)$ belongs to $L_2(\mathbb{P}\otimes\mathbb{m})$.*

Furthermore, if $f(X_t) \in \mathbb{D}_{1,2}$, then for $\mathbb{P}\otimes\mathbb{m}$-a.e. $(\omega,s,x) \in \Omega\times[0,T]\times\mathbb{R}$ one has

$$D_{s,x}f(X_t) = f'_w(X_t)\mathbb{1}_{[0,t]\times\{0\}}(s,x) + \frac{f(X_t+x)-f(X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x),$$

where we set, by convention, $f'_w := 0$ whenever $\sigma = 0$.

This proposition was established in [27, Corollary 3.1 in the second article of this thesis] and it provides an equivalent condition such that a functional of X_t belongs to $\mathbb{D}_{1,2}$. For X being the Brownian motion, see [29, Proposition V.2.3.1], and for X without a Brownian component, see [16, Lemma 3.2].

6.3. Preparation for the proof of Theorem 4.2.

Since we shall work simultaneously with the two Lévy processes X and Z (under $\mathbb{P}$) for which it holds $\mathrm{e}^X = \mathcal{E}(Z)$ as introduced in Subsection 6.1, we agree on the following convention to avoid confusions and determine clearly the referred process.

Convention 6.4. For $Y \in \{X,Z\}$, the notations γ_Y, σ_Y, ν_Y, N_Y, μ_Y, $\mathbb{m}_Y$, M_Y, I^Y, D^Y, $\mathbb{D}^Y_{1,2}$ introduced in Subsection 4.1 and Subsection 6.2 are assigned to Y.

The following lemma shows that X and Z generate the same Malliavin–Sobolev space.

Lemma 6.5. *One has $\mathbb{D}^X_{1,2} = \mathbb{D}^Z_{1,2}$.*

Proof. We define a bijection $\varrho\colon \mathbb{R}\to(-1,\infty)$ by

$$\varrho(x) := \mathrm{e}^x - 1, \quad x \in \mathbb{R}.$$

[4]A locally integrable function h is called a *weak derivative* (unique up to a λ-null set) of a locally integrable function f on $\mathbb{R}$ if $\int_\mathbb{R} f(x)\phi'(x)\mathrm{d}x = -\int_\mathbb{R} h(x)\phi(x)\mathrm{d}x$ for any smooth function ϕ with compact support in $\mathbb{R}$. If such an h exists, then we denote $f'_w := h$.

It is clear that $\varrho(x) = 0 \Leftrightarrow x = 0$. In the sequel, we agree on the convention $\frac{e^0-1}{0} := 1$ and $\frac{\ln(0+1)}{0} := 1$. The relation between ν_X and ν_Z (see Remark 6.1) implies that, for any Borel function $w \geqslant 0$,

$$\int_{(-1,\infty)} w(z)\mu_Z(\mathrm{d}z) = \int_{\mathbb{R}} w(\varrho(x)) \left|\frac{\varrho(x)}{x}\right|^2 \mu_X(\mathrm{d}x). \tag{6.5}$$

Fix $n \geqslant 1$. Let us define the operator

$$\begin{aligned} \Psi_n \colon L_2(\mathbb{m}_X^{\otimes n}) &\to L_2(\mathbb{m}_Z^{\otimes n}) \\ \xi_n^X &\mapsto \xi_n^Z \end{aligned}$$

by setting, for $((t_1, z_1), \ldots, (t_n, z_n)) \in ([0,T] \times (-1,\infty))^n$, that

$$\xi_n^Z((t_1,z_1),\ldots,(t_n,z_n)) := \xi_n^X((t_1,\varrho^{-1}(z_1)),\ldots,(t_n,\varrho^{-1}(z_n))) \prod_{i=1}^n \frac{\varrho^{-1}(z_i)}{z_i}, \tag{6.6}$$

and let $\xi_n^Z := 0$ otherwise. We now show that Ψ_n is well-defined. Denote

$$\mathbb{m}_X^{\otimes n}(\mathrm{d}t, \mathrm{d}x) := \mathbb{m}_X(\mathrm{d}t_1, \mathrm{d}x_1) \cdots \mathbb{m}_X(\mathrm{d}t_n, \mathrm{d}x_n).$$

By an induction argument using (6.5), together with Fubini's theorem, one has

$$\begin{aligned} \|\xi_n^Z\|^2_{L_2(\mathbb{m}_Z^{\otimes n})} &= \int_{([0,T]\times\mathbb{R})^n} \left| \xi_n^Z((t_1,\varrho(x_1)),\ldots,(t_n,\varrho(x_n))) \prod_{i=1}^n \frac{\varrho(x_i)}{x_i}\right|^2 \mathbb{m}_X^{\otimes n}(\mathrm{d}t,\mathrm{d}x) \\ &= \int_{([0,T]\times\mathbb{R})^n} \left| \xi_n^X((t_1,x_1),\ldots,(t_n,x_n)) \prod_{i=1}^n \frac{x_i}{\varrho(x_i)}\right|^2 \left|\prod_{i=1}^n \frac{\varrho(x_i)}{x_i}\right|^2 \mathbb{m}_X^{\otimes n}(\mathrm{d}t,\mathrm{d}x) \\ &= \|\xi_n^X\|^2_{L_2(\mathbb{m}_X^{\otimes n})}, \end{aligned}$$

which ensures $\xi_n^Z \in L_2(\mathbb{m}_Z^{\otimes n})$, and thus $I_n^Z(\xi_n^Z)$ exists as an element in $L_2(\mathbb{P})$. Furthermore, as a by-product of the arguments above, the operator Ψ_n is linear and bounded, thus it is continuous.

We next show for any $\xi_n^X \in L_2(\mathbb{m}_X^{\otimes n})$ that, a.s.,

$$I_n^X(\xi_n^X) = I_n^Z(\Psi_n(\xi_n^X)) = I_n^Z(\xi_n^Z). \tag{6.7}$$

We prove (6.7) only for $n = 1$ since it follows for $n \geqslant 2$ in the same way. Let $(a,b] \subset [0,T]$, and let $B \in \mathcal{B}((-1,\infty))$ with $0 \notin \overline{B}$. Then, $0 \notin \overline{\varrho^{-1}(B)}$. We derive from (6.2) that $\Delta Z = \varrho(\Delta X)$, and hence, a.s.,

$$\begin{aligned} \int_{[0,T]\times\mathbb{R}_0} \mathbb{1}_{(a,b]\times B}(s,x) x N_X(\mathrm{d}s,\mathrm{d}x) &= \sum_{\substack{a<s\leqslant b\\ \Delta X_s \in B}} \Delta X_s = \sum_{\substack{a<s\leqslant b\\ \Delta Z_s \in \varrho(B)}} \varrho^{-1}(\Delta Z_s) \\ &= \int_{(a,b]\times\varrho(B)} \varrho^{-1}(z) N_Z(\mathrm{d}s,\mathrm{d}z) = \int_{[0,T]\times\mathbb{R}_0} \Psi_1(\mathbb{1}_{(a,b]\times B})(s,z) z N_Z(\mathrm{d}s,\mathrm{d}z). \end{aligned}$$

As a consequence, the expected values of both sides are equal, and hence, a.s.,

$$I_1^X(\mathbb{1}_{(a,b]\times B}) = I_1^Z(\Psi_1(\mathbb{1}_{(a,b]\times B})).$$

Recall that the Gaussian components of X and Z coincide pathwise. Hence, due to the denseness in $L_2(\mathbb{m}_X)$ of the linear hull of $\{\mathbb{1}_{(a,b]\times\{0\}}, \mathbb{1}_{(a,b]\times B} : (a,b] \subset [0,T], B \in \mathcal{B}(\mathbb{R}), 0 \notin \overline{B}\}$, together with the continuity of I_1^X, I_1^Z and Ψ_1, we deduce that, a.s.,

$$I_1^X(\xi_1^X) = I_1^Z(\Psi_1(\xi_1^X)) = I_1^Z(\xi_1^Z).$$

Finally, let $\xi \in \mathbb{D}^X_{1,2}$ and suppose that $\xi = \sum_{n=0}^{\infty} I^X_n(\tilde{\xi}^X_n)$, where $\tilde{\xi}^X_n \in L_2(\mathbb{m}^{\otimes n}_X)$ are symmetric. By the definition of Ψ_n, the function $\Psi_n(\tilde{\xi}^X_n)$ is also symmetric. Since $\mathcal{F}^Z_T = \mathcal{F}^X_T$ as showed in Remark 6.1, the uniqueness of chaos expansion and (6.7) lead to, a.s.,

$$\xi = \mathbb{E}\xi + \sum_{n=1}^{\infty} I^X_n(\tilde{\xi}^X_n) = \mathbb{E}\xi + \sum_{n=1}^{\infty} I^Z_n(\tilde{\xi}^Z_n). \tag{6.8}$$

Since $\|\tilde{\xi}^X_n\|^2_{L_2(\mathbb{m}^{\otimes n}_X)} = \|\tilde{\xi}^Z_n\|^2_{L_2(\mathbb{m}^{\otimes n}_Z)}$, it implies that $\xi \in \mathbb{D}^Z_{1,2}$, and hence, $\mathbb{D}^X_{1,2} \subseteq \mathbb{D}^Z_{1,2}$.

By exchanging the role of ϱ and ϱ^{-1}, together with the fact that $\nu_X = \nu_Z \circ (\varrho^{-1})^{-1}$, the converse inclusion $\mathbb{D}^Z_{1,2} \subseteq \mathbb{D}^X_{1,2}$ follows. Therefore, $\mathbb{D}^X_{1,2} = \mathbb{D}^Z_{1,2}$ as desired. □

We use Assumption 4.1 from now until the end of this section. Recall that Z is an $L_2(\mathbb{P})$-martingale with zero mean, hence one can write $\mathrm{d}Z_t = \int_{\mathbb{R}} M_Z(\mathrm{d}t, \mathrm{d}z)$.

We approach the GKW decomposition of $g(S_T) \in L_2(\mathbb{P})$ by means of chaos expansion with respect to the *Lévy process* Z in the way introduced in [15] as follows. First, it is known that (see, e.g., [15, Definiton 1 and Lemma 1]), a.s.,

$$S_T = 1 + \sum_{n=1}^{\infty} I^Z_n\left(\frac{\mathbb{1}^{\otimes n}_{[0,T]\times\mathbb{R}}}{n!}\right),$$

where the kernels in the chaos expansion of S_T do not depend on the time variables. According to [2, Theorem 4], this property is preserved for $g(S_T) \in L_2(\mathbb{P})$. Namely, a.s.,

$$g(S_T) = \sum_{n=0}^{\infty} I^Z_n\left(\tilde{g}_n \mathbb{1}^{\otimes n}_{[0,T]}\right), \tag{6.9}$$

where $\tilde{g}_n \in \tilde{L}_2(\mu^{\otimes n}_Z)$. For each $n \geqslant 1$, define the function $\tilde{h}_{n-1} \in \tilde{L}_2(\mu^{\otimes(n-1)}_Z)$ by

$$\tilde{h}_{n-1}(z_1, \dots, z_{n-1}) := \int_{\mathbb{R}} \tilde{g}_n(z_1, \dots, z_{n-1}, z)\frac{\mu_Z(\mathrm{d}z)}{\mu_Z(\mathbb{R})}. \tag{6.10}$$

Definition 6.6. (1) Let $\varphi^g = (\varphi^g_t)_{t\in[0,T)}$ be the càdlàg version of the $L_2(\mathbb{P})$-martingale

$$\left(\tilde{h}_0 + \sum_{n=1}^{\infty}(n+1)I^Z_n\left(\tilde{h}_n \mathbb{1}^{\otimes n}_{[0,t]}\right)\right)_{t\in[0,T)},$$

where the infinite sum is taken in $L_2(\mathbb{P})$.

(2) Define the process $\vartheta^g \in \mathrm{CL}([0,T))$ by setting $\vartheta^g := \varphi^g/S$.

Lemma 6.7. *Let $g(S_T) \in L_2(\mathbb{P})$. Then ϑ^g_- is a MVH strategy corresponding to $g(S_T)$.*

Proof. We use the functions $\tilde{g}_n, \tilde{h}_n$ defined in (6.9)–(6.10). Since each element in $\tilde{L}_2(\mathbb{m}^{\otimes n}_Z)$ is symmetric, we only need to define it on $((t_1, z_1), \dots, (t_n, z_n))$ with $0 < t_1 < \cdots < t_n < T$. Thus, for $n \geqslant 2$, we define $\tilde{k}_n \in \tilde{L}_2(\mathbb{m}^{\otimes n}_Z)$ by

$$\tilde{k}_n((t_1, z_1), \dots, (t_n, z_n)) := \tilde{h}_{n-1}(z_1, \dots, z_{n-1}) \quad \text{for } 0 < t_1 < \cdots < t_n < T,$$

and set $\tilde{k}_1(t, z) := \tilde{h}_0$. According to the argument in [15, Eqs. (7)–(10)], it holds that the stochastic integral $\int_0^T \varphi^g_{t-}\mathrm{d}Z_t$ is well-defined and

$$\int_0^T \varphi^g_{t-}\mathrm{d}Z_t = \sum_{n=1}^{\infty} I^Z_n(\tilde{k}_n).$$

Let $L^g = (L^g_t)_{t\in[0,T]}$ be the càdlàg version of the martingale closed by

$$L^g_T := g(S_T) - \mathbb{E}g(S_T) - \int_0^T \varphi^g_{t-}\mathrm{d}Z_t.$$

Then, $g(S_T)$ can be re-written as

$$g(S_T) = \mathbb{E}g(S_T) + \int_0^T \varphi^g_{t-}\mathrm{d}Z_t + L^g_T = \mathbb{E}g(S_T) + \sum_{n=1}^{\infty} I^Z_n(\tilde{k}_n) + L^g_T.$$

We now show that L^g is strongly orthogonal to Z. For $t \in (0,T]$, one has, a.s.,

$$L^g_t = \sum_{n=1}^{\infty} I^Z_n\left(\tilde{g}_n \mathbb{1}^{\otimes n}_{[0,t]}\right) - \sum_{n=1}^{\infty} I^Z_n\left(\tilde{k}_n \mathbb{1}^{\otimes n}_{[0,t]}\right) = \sum_{n=1}^{\infty} I^Z_n\left((\tilde{g}_n - \tilde{k}_n)\mathbb{1}^{\otimes n}_{[0,t]}\right).$$

Since $Z_t = \int_0^t \mathrm{d}Z_s = \int_0^t\int_{\mathbb{R}} M_Z(\mathrm{d}s,\mathrm{d}z)$, one has for any $t\in(0,T]$ and $n \geqslant 1$ that, a.s.,

$$\begin{aligned}\left\langle I^Z_n\left((\tilde{g}_n - \tilde{k}_n)\mathbb{1}^{\otimes n}_{[0,\cdot]}\right), Z\right\rangle_t &= n\int_0^t\int_{\mathbb{R}} I^Z_{n-1}\left((\tilde{g}_n(\cdot,z) - \tilde{k}_n(\cdot,(s,z)))\mathbb{1}^{\otimes(n-1)}_{[0,s]}\right)\mu_Z(\mathrm{d}z)\mathrm{d}s\\ &= n\int_0^t I^Z_{n-1}\left(\int_{\mathbb{R}}(\tilde{g}_n(\cdot,z) - \tilde{k}_n(\cdot,(s,z)))\mu_Z(\mathrm{d}z)\mathbb{1}^{\otimes(n-1)}_{[0,s]}\right)\mathrm{d}s\\ &= 0,\end{aligned}$$

where one can see that the second equality holds by testing with multiple integrals. Since the infinite sum in the chaos representation of L^g_t is taken in $L_2(\mathbb{P})$, we conclude that $\langle L^g, Z\rangle = 0$. Hence, it follows from $\vartheta^g_{t-}\mathrm{d}S_t = \varphi^g_{t-}\mathrm{d}Z_t$ that $g(S_T) = \mathbb{E}g(S_T) + \int_0^T \vartheta^g_{t-}\mathrm{d}S_t + L^g_T$ is the GKW decomposition of $g(S_T)$. □

6.4. **Proof of Theorem 4.2.** We verify that the process ϑ^g in Definition 6.6 satisfies the requirements. The assertion (1) and the martingale property of $\vartheta^g S$ are clear by the definition of ϑ^g and Lemma 6.7. For the latter part of (2), since φ^g and S are martingales adapted to the quasi-left continuous filtration $\mathbb{F}^X$, it implies that $\varphi^g_t = \varphi^g_{t-}$ a.s. and $S_t = S_{t-}$ a.s. for each $t\in[0,T)$ (see [31]). Therefore, $\vartheta^g_t = \vartheta^g_{t-}$ a.s. for each $t \in [0,T)$.

(3) Recall from Lemma 6.5 that $\mathbb{D}^X_{1,2} = \mathbb{D}^Z_{1,2}$. We have in Definition 6.6 and Lemma 6.7 the strategy given as chaos expansion with respect to Z. In order to get the explicit representation (4.2), we change it into a representation with respect to X where we can use Proposition 6.3.

Step 1. Let $\xi \in \mathbb{D}^Z_{1,2}$ have the expansion (6.8). We first write the Malliavin derivative of ξ as the element in $L_2(\mathbb{P}\otimes \mathbbm{m}_Z)$ and then integrate it with respect to $\mathbbm{m}_Z$ to obtain, a.s.,

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}} \left(D^Z_{s,z}\xi\right)\mathbbm{m}_Z(\mathrm{d}s,\mathrm{d}z)\\ &= \int_0^T\int_{\mathbb{R}}\left(L_2(\mathbb{P}\otimes\mathbbm{m}_Z)\text{-}\lim_{N\to\infty}\sum_{n=1}^{N} nI^Z_{n-1}(\tilde{\xi}^Z_n((s,z),\cdot))\right)\mathbbm{m}_Z(\mathrm{d}s,\mathrm{d}z)\\ &= L_2(\mathbb{P})\text{-}\lim_{N\to\infty}\sum_{n=1}^{N}\int_0^T\int_{\mathbb{R}} nI^Z_{n-1}(\tilde{\xi}^Z_n((s,z),\cdot))\mathbbm{m}_Z(\mathrm{d}s,\mathrm{d}z) \qquad (6.11)\\ &= L_2(\mathbb{P})\text{-}\lim_{N\to\infty}\sum_{n=1}^{N}\int_0^T\int_{\mathbb{R}} nI^Z_{n-1}(\tilde{\xi}^Z_n((s,\mathrm{e}^x-1),\cdot))\left|\frac{\mathrm{e}^x-1}{x}\right|^2\mathbbm{m}_X(\mathrm{d}s,\mathrm{d}x)\end{aligned}$$

$$= L_2(\mathbb{P})\text{-}\lim_{N\to\infty}\sum_{n=1}^{N}\int_0^T\int_{\mathbb{R}} nI_{n-1}^X(\tilde{\xi}_n^X((s,x),\cdot))\frac{\mathrm{e}^x-1}{x}\mathbb{m}_X(\mathrm{d}s,\mathrm{d}x) \tag{6.12}$$

$$= \int_0^T\int_{\mathbb{R}}\left(L_2(\mathbb{P}\otimes\mathbb{m}_X)\text{-}\lim_{N\to\infty}\sum_{n=1}^{N} nI_{n-1}^X(\tilde{\xi}_n^X((s,x),\cdot))\right)\frac{\mathrm{e}^x-1}{x}\mathbb{m}_X(\mathrm{d}s,\mathrm{d}x) \tag{6.13}$$

$$= \int_0^T\int_{\mathbb{R}}\left(D_{s,x}^X\xi\right)\frac{\mathrm{e}^x-1}{x}\mathbb{m}_X(\mathrm{d}s,\mathrm{d}x), \tag{6.14}$$

where one uses the fact that $\mathbb{m}_Z([0,T]\times\mathbb{R}) = \int_0^T\int_{\mathbb{R}}\left|\frac{\mathrm{e}^x-1}{x}\right|^2\mathbb{m}_X(\mathrm{d}s,\mathrm{d}x) < \infty$ to derive (6.11) and (6.13). In order to achieve (6.12), we apply the definition of Ψ_{n-1} in (6.6) and then use (6.7) with the convention that Ψ_0 is the identical map on $\mathbb{R}$ as follows

$$\begin{aligned}
&\tilde{\xi}_n^Z((s,\mathrm{e}^x-1),(s_1,z_1),\ldots,(s_{n-1},z_{n-1}))\\
&= \tilde{\xi}_n^X((s,x),(s_1,\ln(z_1+1)),\ldots,(s_{n-1},\ln(z_{n-1}+1)))\prod_{i=1}^{n-1}\frac{\ln(z_i+1)}{z_i}\frac{x}{\mathrm{e}^x-1}\\
&= \frac{x}{\mathrm{e}^x-1}\left(\Psi_{n-1}\left(\tilde{\xi}_n^X((s,x),\cdot)\right)((s_1,z_1),\ldots,(s_{n-1},z_{n-1}))\right).
\end{aligned}$$

Step 2. For $x\in\mathbb{R}, t\in(0,T)$, we define

$$f(x) := g(\mathrm{e}^x) \quad\text{and}\quad F(t,x) := G(t,\mathrm{e}^x).$$

It turns out that $F(t,X_t) = \mathbb{E}^{\mathcal{F}_t}[f(X_T)]$ a.s. We then derive from [18, Lemma D.1] that $F(t,X_t)\in\mathbb{D}_{1,2}^X$. Applying Proposition 6.3, we obtain

$$D_{s,x}^X F(t,X_t) = \partial_x F(t,X_t)\mathbb{1}_{[0,t]\times\{0\}}(s,x) + \frac{F(t,X_t+x)-F(t,X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x) \tag{6.15}$$

for $\mathbb{P}\otimes\mathbb{m}_X$-a.e. $(\omega,s,x)\in\Omega\times[0,T]\times\mathbb{R}$. We multiply both sides of (6.15) with $\frac{\mathrm{e}^x-1}{x}$ and then integrate them with respect to $\mathbb{m}_X$ to obtain, a.s.,

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}}\left(D_{s,x}^X F(t,X_t)\frac{\mathrm{e}^x-1}{x}\right)\mathbb{m}_X(\mathrm{d}s,\mathrm{d}x)\\
&= t\int_{\mathbb{R}}\left(\partial_x F(t,X_t)\mathbb{1}_{\{x=0\}} + \frac{F(t,X_t+x)-F(t,X_t)}{x}\frac{\mathrm{e}^x-1}{x}\mathbb{1}_{\{x\neq0\}}\right)\mu_X(\mathrm{d}x)\\
&= t\left(\sigma^2 S_t\partial_y G(t,S_t) + \int_{\mathbb{R}}(G(t,\mathrm{e}^x S_t)-G(t,S_t))(\mathrm{e}^x-1)\nu_X(\mathrm{d}x)\right).
\end{aligned} \tag{6.16}$$

On the other hand, for the representation of $g(S_T)$ given in (6.9), taking the conditional expectation of $g(S_T)$ with respect to $\mathcal{F}_t$ yields, a.s.,

$$G(t,S_t) = \mathbb{E}^{\mathcal{F}_t}[g(S_T)] = \sum_{n=0}^{\infty} I_n^Z\left(\tilde{g}_n\mathbb{1}_{[0,t]}^{\otimes n}\right).$$

Since $G(t,S_t)\in\mathbb{D}_{1,2}^Z$, we write the chaos representation of the Malliavin derivative of $G(t,S_t)$ with respect to the underlying process Z as in Definition 6.2, and then, integrate that with respect to the measure $\mathbb{m}_Z$ to obtain, a.s.,

$$\int_0^T\int_{\mathbb{R}}\left(D_{s,z}^Z G(t,S_t)\right)\mathbb{m}_Z(\mathrm{d}s,\mathrm{d}z) = \int_0^T\int_{\mathbb{R}}\left(\sum_{n=1}^{\infty} nI_{n-1}^Z\left(\tilde{g}_n(\cdot,z)\mathbb{1}_{[0,t]}^{\otimes(n-1)}\mathbb{1}_{[0,t]}(s)\right)\right)\mathbb{m}_Z(\mathrm{d}s,\mathrm{d}z)$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} n I_{n-1}^{Z}\left(\int_0^T\int_{\mathbb{R}} \tilde{g}_n(\cdot,z)\mathbb{1}_{[0,t]}^{\otimes(n-1)}\mathbb{1}_{[0,t]}(s)\mathrm{m}_Z(\mathrm{d}s,\mathrm{d}z)\right)\\
&= t\sum_{n=1}^{\infty} n I_{n-1}^{Z}\left(\left(\int_{\mathbb{R}} \tilde{g}_n(\cdot,z)\mu_Z(\mathrm{d}z)\right)\mathbb{1}_{[0,t]}^{\otimes(n-1)}\right)\\
&= t c_{(4.2)}^2 \vartheta_t^g S_t, \qquad (6.17)
\end{aligned}
$$

where the last equality comes from (6.10), Definition 6.6, and $\mu_Z(\mathbb{R}) = c_{(4.2)}^2$. Applying *Step 1* for $\xi = F(t, X_t) = G(t, S_t)$, we derive from (6.14) that, a.s.,

$$
\int_0^T\int_{\mathbb{R}} \left(D_{s,x}^X F(t,X_t)\frac{\mathrm{e}^x-1}{x}\right)\mathrm{m}_X(\mathrm{d}s,\mathrm{d}x) = \int_0^T\int_{\mathbb{R}} \left(D_{s,z}^Z G(t,S_t)\right)\mathrm{m}_Z(\mathrm{d}s,\mathrm{d}z). \qquad (6.18)
$$

Combining (6.16), (6.17) with (6.18), we get (4.2).

7. Technical results I: Regularity of the weight processes $\overline{\Phi}$ and $\Phi(\eta)$

We recall $\overline{\Phi}$ from (3.9) and Definition 2.1.

Proposition 7.1. (1) *Let $p, q, r \in (0,\infty)$ with $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then for any $\Phi, \Psi \in \mathrm{CL}^+([0,T])$,*

$$
\|\Phi\Psi\|_{\mathcal{SM}_r(\mathbb{P})} \leqslant \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}\|\Psi\|_{\mathcal{SM}_q(\mathbb{P})}.
$$

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

$$
\|\overline{\Phi}\|_{\mathcal{SM}_p(\mathbb{P})} \leqslant \begin{cases} 3\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} + 1 & \text{if } p \in [1,\infty) \\ (3\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}^p + 1)^{\frac{1}{p}} & \text{if } p \in (0,1). \end{cases}
$$

Proof. Assertion (1) is given in [18, Proposition A.2]. We now prove (2). Let $a \in [0,T)$ be arbitrary. For $p \in [1,\infty)$, applying the conditional Minkovski inequality yields, a.s.,

$$
\begin{aligned}
\left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T]}\overline{\Phi}_t^p\right]\right)^{\frac{1}{p}} &\leqslant \left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T]}\Phi_t^p\right]\right)^{\frac{1}{p}} + \left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{s\in[0,T]}|\Delta\Phi_s|^p\right]\right)^{\frac{1}{p}}\\
&\leqslant \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}\Phi_a + \sup_{s\in[0,a]}|\Delta\Phi_s| + \left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in(a,T]}|\Delta\Phi_t|^p\right]\right)^{\frac{1}{p}}\\
&\leqslant \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}\Phi_a + \sup_{s\in[0,a]}|\Delta\Phi_s| + 2\left(\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in(a,T]}\Phi_t^p\right]\right)^{\frac{1}{p}}\\
&\leqslant (3\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} + 1)\overline{\Phi}_a.
\end{aligned}
$$

For $p \in (0,1)$, we use the same argument as in the previous case where one applies the inequality $|x+y|^p \leqslant |x|^p + |y|^p$ for $x, y \in \mathbb{R}$ to obtain, a.s.,

$$
\mathbb{E}^{\mathcal{F}_a}\left[\sup_{t\in[a,T]}\overline{\Phi}_t^p\right] \leqslant (3\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}^p + 1)\overline{\Phi}_a^p.
$$

Hence, the desired conclusion follows. □

Recall the Lévy process X with characteristic triplet (γ,σ,ν) and exponent ψ mentioned in Subsection 4.1. Recall $\Phi(\eta)$ from (4.3) and $S = \mathrm{e}^X$.

Proposition 7.2. *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,*

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

Proof. The first step considers the particular case when S is a martingale, and the general case is handled in the second step.

Step 1. Assume that S is a $\mathbb{P}$-martingale. By (6.4), the assumption $\int_{|x|>1} \mathrm{e}^{qx}\nu(\mathrm{d}x) < \infty$ implies $\mathrm{e}^{X_t} \in L_q(\mathbb{P})$ for all $t > 0$. Denote $c_q := (\frac{q}{q-1})^q$ and define $M = (M_t)_{t\in[0,T]}$ by

$$M_t := \sup\nolimits_{u\in[0,t]} \mathrm{e}^{X_t - X_u}.$$

We show that M is a positive $L_q(\mathbb{P})$-submartingale. The adaptedness and positivity are clear. Pick a $t \in (0, T]$. Since $(X_t - X_{t-u})_{u\in[0,t]}$ is *càglàd* (left-continuous with right limits) and $(X_u)_{u\in[0,t]}$ is càdlàg, and both processes have the same finite-dimensional distribution, applying Doob's maximal inequality yields

$$\begin{aligned}\mathbb{E}M_t^q &= \mathbb{E}\Big[\sup\nolimits_{u\in[0,t]} \mathrm{e}^{q(X_t-X_u)}\Big] = \mathbb{E}\Big[\sup\nolimits_{u\in[0,t]} \mathrm{e}^{q(X_t-X_{t-u})}\Big] \\ &= \mathbb{E}\Big[\sup\nolimits_{u\in[0,t]} \mathrm{e}^{qX_u}\Big] \leqslant c_q \mathbb{E}\mathrm{e}^{qX_t} < \infty.\end{aligned} \tag{7.1}$$

For $0 \leqslant s \leqslant t \leqslant T$ one has, a.s.,

$$\mathbb{E}^{\mathcal{F}_s}[M_t] \geqslant \mathbb{E}^{\mathcal{F}_s}\Big[\sup\nolimits_{u\in[0,s]} \mathrm{e}^{X_t-X_u}\Big] = \sup\nolimits_{u\in[0,s]} \mathrm{e}^{X_s-X_u}\mathbb{E}\mathrm{e}^{X_t-X_s} = M_s,$$

where we use $\mathbb{E}\mathrm{e}^{X_t-X_s} = \mathbb{E}S_{t-s} = 1$.

We observe that the process $\Phi(\eta)$ can be re-written as

$$\Phi(\eta)_t = \mathrm{e}^{\eta X_t} \sup\nolimits_{s\in[0,t]} \mathrm{e}^{(1-\eta)(X_t-X_s)} = \mathrm{e}^{\eta X_t} M_t^{1-\eta}.$$

Let us fix $\eta \in (0, 1)$ and $a \in [0, T]$. For $\mathrm{e}^{\eta X} = (\mathrm{e}^{\eta X_t})_{t\in[0,T]}$, applying Doob's maximal inequality and Jensen's inequality we obtain that, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\Big[\sup\nolimits_{t\in[a,T]}(\mathrm{e}^{\eta X_t})^{\frac{q}{\eta}}\Big] &= \mathrm{e}^{qX_a}\mathbb{E}\Big[\sup\nolimits_{t\in[a,T]} \mathrm{e}^{q(X_t-X_a)}\Big] \leqslant c_q\mathrm{e}^{qX_a}\mathbb{E}\mathrm{e}^{q(X_T-X_a)} \\ &= c_q\mathrm{e}^{qX_a}\mathbb{E}\mathrm{e}^{qX_{T-a}} \leqslant c_q\mathrm{e}^{qX_a}\mathbb{E}\mathrm{e}^{qX_T},\end{aligned}$$

which implies

$$\|\mathrm{e}^{\eta X}\|_{\mathcal{SM}_{q/\eta}(\mathbb{P})} \leqslant (c_q\mathbb{E}\mathrm{e}^{qX_T})^{\frac{\eta}{q}}.$$

For $M^{1-\eta} = (M_t^{1-\eta})_{t\in[0,T]}$, one has that, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\Big[\sup\nolimits_{t\in[a,T]}(M_t^{1-\eta})^{\frac{q}{1-\eta}}\Big] &= \mathbb{E}^{\mathcal{F}_a}\Big[\sup\nolimits_{t\in[a,T]} M_t^q\Big] \leqslant c_q\mathbb{E}^{\mathcal{F}_a}\big[M_T^q\big] \\ &\leqslant c_q\mathbb{E}^{\mathcal{F}_a}\Big[\sup\nolimits_{s\in[0,a]} \mathrm{e}^{q(X_T-X_s)}\Big] + c_q\mathbb{E}^{\mathcal{F}_a}\Big[\sup\nolimits_{s\in[a,T]} \mathrm{e}^{q(X_T-X_s)}\Big] \\ &= c_q \sup\nolimits_{s\in[0,a]} \mathrm{e}^{q(X_a-X_s)}\mathbb{E}\mathrm{e}^{q(X_T-X_a)} + c_q\mathbb{E}\Big[\sup\nolimits_{s\in[a,T]} \mathrm{e}^{q(X_T-X_s)}\Big] \\ &\leqslant 2c_q \sup\nolimits_{s\in[0,a]} \mathrm{e}^{q(X_a-X_s)}\mathbb{E}\Big[\sup\nolimits_{s\in[a,T]} \mathrm{e}^{q(X_T-X_s)}\Big] \\ &\leqslant \Big(2c_q\mathbb{E}\Big[\sup\nolimits_{s\in[0,T]} \mathrm{e}^{q(X_T-X_s)}\Big]\Big) M_a^q \\ &\leqslant \big(2c_q^2\mathbb{E}\mathrm{e}^{qX_T}\big) M_a^q,\end{aligned}$$

where the conditional Doob maximal inequality is applied for the positive sub-martingale M to obtain the first inequality, and the last one comes from (7.1). Hence,

$$\|M^{1-\eta}\|_{\mathcal{SM}_{q/(1-\eta)}(\mathbb{P})} \leqslant (2c_q^2\mathbb{E}\mathrm{e}^{qX_T})^{\frac{1-\eta}{q}}.$$

Applying Proposition 7.1(1) with $\frac{1}{q} = \frac{1}{q/\eta} + \frac{1}{q/(1-\eta)}$, we obtain

$$\|\Phi(\eta)\|_{\mathcal{SM}_q(\mathbb{P})} \leqslant \|e^{\eta X}\|_{\mathcal{SM}_{q/\eta}(\mathbb{P})} \|M^{1-\eta}\|_{\mathcal{SM}_{q/(1-\eta)}(\mathbb{P})} \leqslant 2^{\frac{1-\eta}{q}} \left(\frac{q}{q-1}\right)^2 \|S_T\|_{L_q(\mathbb{P})} < \infty,$$

which asserts $\Phi(\eta) \in \mathcal{SM}_q(\mathbb{P})$. When $\eta = 0$ or $\eta = 1$, the desired conclusion is straightforward as $\Phi(0) = M$, $\Phi(1) = e^X$.

Step 2. In the general case, we define

$$\widetilde{S}_t := e^{t\psi(-i)} S_t.$$

Then, it is known that $\widetilde{S}$ is a martingale under $\mathbb{P}$. Some standard calculations yield

$$e^{-T|\psi(-i)|}\widetilde{\Phi}(\eta)_t \leqslant \Phi(\eta)_t \leqslant e^{T|\psi(-i)|}\widetilde{\Phi}(\eta)_t,$$

where $\widetilde{\Phi}(\eta)_t := \widetilde{S}_t \sup_{u\in[0,t]}(\widetilde{S}_u^{\eta-1})$. Applying *Step 1* for $\mathbb{P}$-martingale $\widetilde{S}$ we derive that $\widetilde{\Phi}(\eta) \in \mathcal{SM}_q(\mathbb{P})$. Hence, for $a \in [0, T]$, one has, a.s.,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_a}\left[\sup\nolimits_{t\in[a,T]} \Phi(\eta)_t^q\right] &\leqslant e^{qT|\psi(-i)|}\mathbb{E}^{\mathcal{F}_a}\left[\sup\nolimits_{t\in[a,T]} \widetilde{\Phi}(\eta)_t^q\right] \\
&\leqslant e^{qT|\psi(-i)|}\|\widetilde{\Phi}(\eta)\|^q_{\mathcal{SM}_q(\mathbb{P})}\widetilde{\Phi}(\eta)_a^q \\
&\leqslant e^{2qT|\psi(-i)|}2^{1-\eta}\left(\frac{q}{q-1}\right)^{2q}\|\widetilde{S}_T\|^q_{L_q(\mathbb{P})}\Phi(\eta)_a^q \\
&\leqslant e^{T|\psi(-i)|(2q+1)}2^{1-\eta}\left(\frac{q}{q-1}\right)^{2q}\|S_T\|^q_{L_q(\mathbb{P})}\Phi(\eta)_a^q,
\end{aligned}$$

which proves the desired conclusion. □

8. Technical results II: Gradient type estimates for a Lévy semigroup on Hölder spaces

This section provides some gradient type estimates in the Lévy setting for proving Theorem 4.6, and they might also be of independent interest.

Let us introduce some notations. For a non-empty and open set $U \subseteq \mathbb{R}$ and for $n \geqslant 1$, let $C^n(U)$ denote the family of n times continuously differentiable functions on U, and set $C^\infty(U) := \cap_{n\geqslant 1} C^n(U)$. The space $C_c^\infty(U)$ consists of all $f \in C^\infty(U)$ with compact support in U. When $U = \mathbb{R}$, we let $C_0^\infty(\mathbb{R})$ denote the family of all $f \in C^\infty(\mathbb{R})$ with $\lim_{|x|\to\infty} f^{(n)}(x) = 0$ for all $n \geqslant 0$.

For $s \in \mathbb{R}$, we define the weighted Lebesgue measure λ_s on $\mathcal{B}(\mathbb{R})$ by setting

$$\lambda_s(\mathrm{d}x) := e^{sx}\mathrm{d}x,$$

and let $\lambda_\infty(\mathrm{d}x) := \lambda_0(\mathrm{d}x) = \mathrm{d}x$ be the usual Lebesgue measure.

8.1. **Some integral estimates for Hölder functions.**

For a Borel function g and a random variable Y such that $\mathbb{E}|g(ye^Y)| < \infty$ for all $y > 0$, we define

$$G(y) := \mathbb{E}g(ye^Y), \quad y > 0.$$

For later use, we establish in this part some estimates for $|G(z) - G(y)|$, where g is a Hölder continuous function or a bounded Borel function.

The first result deals with $g \in \mathring{W}^{1,q}(\mathbb{R}_+)$ (see Definition 4.3).

Proposition 8.1. *Let $q \in [1,\infty]$ and $\eta := 1 - \frac{1}{q} \in [0,1]$. If $g \in \mathring{W}^{1,q}(\mathbb{R}_+)$ and if Y has a density $p \in L_r(\mathbb{R}, \lambda_{\eta r})$ for some $r \in [1, \frac{1}{\eta}]$, then for any $z, y > 0$,*

$$|G(z) - G(y)| \leqslant \left(|g|_{\mathring{W}^{1,q}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} \right) |z-y|^{\eta} \, |\ln z - \ln y|^{1-\frac{1}{r}} \, ,$$

where we set, by convention, $0^0 := 1$, and set $\eta r := 1$ if $\eta = 0$, $r = \infty$.

Proof. Assume that $g(y) - g(x) = \int_x^y h(u)\mathrm{d}u$ for $h \in L_q(\mathbb{R}_+)$. Let $q', r' \in [1,\infty]$ be such that $\frac{1}{q} + \frac{1}{q'} = 1$ and $\frac{1}{r} + \frac{1}{r'} = 1$. Since $r \in [1, \frac{1}{\eta}] = [1, q']$, it implies that $1 \leqslant q \leqslant r' \leqslant \infty$. Denote $B := \frac{q}{r'} \in [0,1]$ and $A := 1 - B$, where $B := 1$ if $q = r' = \infty$. Then, with the sign function $\operatorname{sgn}(x) := \mathbb{1}_{\{x \geqslant 0\}} - \mathbb{1}_{\{x<0\}}$, we have that

$$\begin{aligned}
&|G(z) - G(y)| = \left| \int_{\mathbb{R}} (g(z\mathrm{e}^x) - g(y\mathrm{e}^x)) p(x) \mathrm{d}x \right| \\
&\leqslant \int_{\mathbb{R}} |g(z\mathrm{e}^x) - g(y\mathrm{e}^x)|^A |g(z\mathrm{e}^x) - g(y\mathrm{e}^x)|^B p(x)\mathrm{d}x \\
&\leqslant |g|^A_{C^{0,\eta}(\mathbb{R}_+)} |z-y|^{A\eta} \int_{\mathbb{R}} \mathrm{e}^{A\eta x} |g(z\mathrm{e}^x) - g(y\mathrm{e}^x)|^B p(x)\mathrm{d}x \\
&= |g|^A_{C^{0,\eta}(\mathbb{R}_+)} |z-y|^{A\eta} \int_{\mathbb{R}} \mathrm{e}^{(A-1)\eta x} |g(z\mathrm{e}^x) - g(y\mathrm{e}^x)|^B p(x)\mathrm{e}^{\eta x}\mathrm{d}x \\
&\leqslant |g|^A_{C^{0,\eta}(\mathbb{R}_+)} |z-y|^{A\eta} \left(\int_{\mathbb{R}} \mathrm{e}^{-xB\eta r'} |g(z\mathrm{e}^x) - g(y\mathrm{e}^x)|^{Br'}\mathrm{d}x \right)^{\frac{1}{r'}} \left(\int_{\mathbb{R}} |p(x)\mathrm{e}^{\eta x}|^r \mathrm{d}x \right)^{\frac{1}{r}} \\
&= |g|^A_{C^{0,\eta}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} y^{B\eta} |z-y|^{A\eta} \left(\int_0^\infty u^{-q\eta-1} |g(uz/y) - g(u)|^q \mathrm{d}u \right)^{\frac{1}{r'}} \\
&= |g|^A_{C^{0,\eta}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} y^{B\eta} |z-y|^{A\eta} \left(\int_0^\infty \left| \int_1^{\frac{z}{y}} h(ux)\mathrm{d}x \right|^q \mathrm{d}u \right)^{\frac{1}{r'}} \\
&\leqslant |g|^A_{C^{0,\eta}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} y^{B\eta} |z-y|^{A\eta} \left| \frac{z}{y} - 1 \right|^{\frac{q-1}{r'}} \left(\int_0^\infty \operatorname{sgn}(z-y) \int_1^{\frac{z}{y}} |h(ux)|^q \mathrm{d}x\mathrm{d}u \right)^{\frac{1}{r'}} \\
&= |g|^A_{C^{0,\eta}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} |z-y|^{1-\frac{1}{q}} \, |\ln z - \ln y|^{\frac{1}{r'}} \, |g|^B_{\mathring{W}^{1,q}(\mathbb{R}_+)} \\
&\leqslant |g|_{\mathring{W}^{1,q}(\mathbb{R}_+)} \|p\|_{L_r(\mathbb{R},\lambda_{\eta r})} |z-y|^{\eta} \, |\ln z - \ln y|^{1-\frac{1}{r}} \, ,
\end{aligned}$$

where we apply Hölder's inequality for the third and fourth inequality, and the last one comes from $|g|_{C^{0,\eta}(\mathbb{R}_+)} \leqslant |g|_{\mathring{W}^{1,q}(\mathbb{R}_+)}$. □

The following result is formulated for a bounded or Hölder continuous g.

Proposition 8.2. *Let $\eta \in [0,1]$ and $g \in C^{0,\eta}(\mathbb{R}_+)$. If Y has a density $p \in C^1(\mathbb{R}) \cap L_1(\mathbb{R}, \lambda_\eta)$ with the derivative $p' \in L_1(\mathbb{R}) \cap L_1(\mathbb{R}, \lambda_\eta)$, then for all $z, y > 0$,*

$$|G(z) - G(y)| \leqslant \left(|g|_{C^{0,\eta}(\mathbb{R}_+)} \|p'\|^{1-\eta}_{L_1(\mathbb{R})} \inf_{\kappa > 0} \left| \int_{\mathbb{R}} |\mathrm{e}^x - \kappa| |p'(x)| \mathrm{d}x \right|^{\eta} \right) \frac{|z^\eta - y^\eta|}{\eta}, \qquad (8.1)$$

where we set, by convention, $\frac{|z^0 - y^0|}{0} := \lim_{\eta \downarrow 0} \frac{|z^\eta - y^\eta|}{\eta} = |\ln z - \ln y|$ when $\eta = 0$.

Proof. The assumption $p \in L_1(\mathbb{R}, \lambda_\eta)$ means that $\mathbb{E}\mathrm{e}^{\eta Y} < \infty$, and hence $\mathbb{E}|g(y\mathrm{e}^Y)| < \infty$ for all $y > 0$. Let us pick a constant $\kappa' > 0$ arbitrarily. By a change of variables,

$$G(z) - G(y) = \mathbb{E}g(z\mathrm{e}^Y) - \mathbb{E}g(y\mathrm{e}^Y)$$

$$
\begin{aligned}
&= \int_0^\infty g(u)(p(\ln u - \ln z) - p(\ln u - \ln y))\frac{\mathrm{d}u}{u} \\
&= \int_0^\infty (g(u) - g(\kappa'))(p(\ln u - \ln z) - p(\ln u - \ln y))\frac{\mathrm{d}u}{u}.
\end{aligned}
$$

Since $p \in C^1(\mathbb{R})$, the fundamental theorem of calculus gives

$$
\begin{aligned}
&G(z) - G(y) \\
&= \int_0^\infty (g(u) - g(\kappa'))\left((\ln y - \ln z)\int_0^1 p'(\ln u - \ln y + r(\ln y - \ln z))\mathrm{d}r\right)\frac{\mathrm{d}u}{u} \\
&= (\ln y - \ln z)\int_0^\infty (g(u) - g(\kappa'))\left(\int_0^1 p'(\ln u - \ln y + r(\ln y - \ln z))\mathrm{d}r\right)\frac{\mathrm{d}u}{u}.
\end{aligned}
$$

Since $|g(u) - g(\kappa')| \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}|u - \kappa'|^\eta$, where $0^0 := 1$, we have

$$
\begin{aligned}
&|G(z) - G(y)| \\
&\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\,|\ln z - \ln y|\int_0^\infty |u - \kappa'|^\eta \int_0^1 |p'(\ln u - \ln y + r(\ln y - \ln z))|\mathrm{d}r\frac{\mathrm{d}u}{u} \\
&= |g|_{C^{0,\eta}(\mathbb{R}_+)}\,|\ln z - \ln y|\int_0^1 \left(\int_0^\infty |u - \kappa'|^\eta |p'(\ln u - \ln y + r(\ln y - \ln z))|\frac{\mathrm{d}u}{u}\right)\mathrm{d}r \\
&= |g|_{C^{0,\eta}(\mathbb{R}_+)}\,|\ln z - \ln y|\int_0^1 \left(\int_{\mathbb{R}} |\mathrm{e}^x y^{1-r} z^r - \kappa'|^\eta |p'(x)|\mathrm{d}x\right)\mathrm{d}r. \qquad (8.2)
\end{aligned}
$$

If $\eta = 0$, then (8.1) is obvious in the view of (8.2). Let us now consider $\eta \in (0, 1]$. Thanks to (8.2), G is locally Lipschitz on $\mathbb{R}_+$, which implies the absolute continuity of G on any compact interval of $\mathbb{R}_+$. Consequently, G is differentiable λ-a.e. on $\mathbb{R}_+$. Let $y > 0$ be such that $G'(y)$ exists and is finite. We divide both sides of (8.2) by $|z - y|$ and then let $z \to y$, where the dominated convergence theorem is applicable on the right-hand side due to $p' \in L_1(\mathbb{R}) \cap L_1(\mathbb{R}, \lambda_\eta)$, to derive that, for all $\kappa' > 0$,

$$
|G'(y)| \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)} y^{-1}\int_{\mathbb{R}} |y\mathrm{e}^x - \kappa'|^\eta |p'(x)|\mathrm{d}x.
$$

Hence, for any $\kappa > 0$, we obtain by choosing $\kappa' = y\kappa$ that

$$
|G'(y)| \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)} y^{\eta-1}\int_{\mathbb{R}} |\mathrm{e}^x - \kappa|^\eta |p'(x)|\mathrm{d}x.
$$

Now, for $z, y > 0$, using the fundamental theorem of (Lebesgue integral) calculus yields

$$
\begin{aligned}
|G(z) - G(y)| &= \left|\int_y^z G'(u)\mathrm{d}u\right| \leqslant \operatorname{sgn}(z - y)\int_y^z |G'(u)|\mathrm{d}u \\
&\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\operatorname{sgn}(z - y)\int_y^z u^{\eta-1}\mathrm{d}u \int_{\mathbb{R}} |\mathrm{e}^x - \kappa|^\eta |p'(x)|\mathrm{d}x \\
&= |g|_{C^{0,\eta}(\mathbb{R}_+)}\frac{|z^\eta - y^\eta|}{\eta}\int_{\mathbb{R}} |\mathrm{e}^x - \kappa|^\eta |p'(x)|\mathrm{d}x \\
&\leqslant \left(|g|_{C^{0,\eta}(\mathbb{R}_+)}\|p'\|_{L_1(\mathbb{R})}^{1-\eta}\left|\int_{\mathbb{R}} |\mathrm{e}^x - \kappa||p'(x)|\mathrm{d}x\right|^\eta\right)\frac{|z^\eta - y^\eta|}{\eta},
\end{aligned}
$$

where one applies Hölder's inequality with $\frac{1}{1/\eta} + \frac{1}{1/(1-\eta)} = 1$ to obtain the last estimate. By taking the infimum over $\kappa > 0$, (8.1) follows. $\square$

8.2. **Hölder estimates for a Lévy semigroup.** Let $X = (X_t)_{t\geqslant 0}$ be a Lévy process with characteristic triplet (γ, σ, ν) and exponent ψ as in Subsection 4.1. Let us define

$$\mathcal{D}_{\exp} := \{g\colon \mathbb{R}_+ \to \mathbb{R} \text{ Borel} : \ \mathbb{E}|g(y\mathrm{e}^{X_t})| < \infty \text{ for all } y > 0, t \geqslant 0\}.$$

It is clear that $\mathcal{D}_{\exp}$ depends on the distribution of X. For example, if $\int_{|x|>1} \mathrm{e}^{rx}\nu(\mathrm{d}x) < \infty$ for some $r \in \mathbb{R}$, then any Borel function g with $\sup_{y>0}(1+y)^{-r}|g(y)| < \infty$ belongs to $\mathcal{D}_{\exp}$ because of (6.4). For $t \geqslant 0$, define the mapping $P_t\colon \mathcal{D}_{\exp} \to \mathcal{D}_{\exp}$ by

$$P_t g(y) := \mathbb{E}g(y\mathrm{e}^{X_t}).$$

Since $P_{t+s} = P_t \circ P_s$ for any $s, t \geqslant 0$, the family $(P_t)_{t\geqslant 0}$ is a semigroup on $\mathcal{D}_{\exp}$.

To be able to estimate the integral term of the MVH strategy formula (4.2), we aim to establish an estimate for

$$|P_t g(z) - P_t g(y)|,$$

where g is bounded or Hölder continuous.

The following lemma provides an estimate for the $L_1(\mathbb{R})$-norm of derivatives of transition densities.

Lemma 8.3. *Let X be a Lévy process with characteristic exponent ψ. If*

$$0 < \liminf_{|u|\to\infty} |u|^{-\alpha}\mathrm{Re}\psi(u) \leqslant \limsup_{|u|\to\infty} |u|^{-\alpha}\mathrm{Re}\psi(u) < \infty \tag{8.3}$$

for some $\alpha \in (0,2)$, then X has transition densities $(p_t)_{t>0} \subset C_0^\infty(\mathbb{R})$ such that

$$\sup\nolimits_{t\in(0,T]} t^{\frac{1}{\alpha}} \|\partial_x p_t\|_{L_1(\mathbb{R})} < \infty, \quad T > 0.$$

Proof. See the proof of [18, Theorem 8.21]. □

Since we aim to apply Proposition 8.2, and in order to handle the quantity involving the infimum in (8.1), we provide in Lemma 8.4 below estimates under assumptions which are typically satisfied in applications.

Lemma 8.4. *For some $t > 0$ such that X_t has a differentiable density p_t on $\mathbb{R}$, we define*

$$K_t := \inf_{\kappa>0} \int_{\mathbb{R}} |\mathrm{e}^x - \kappa|\, |\partial_x p_t(x)|\,\mathrm{d}x \in [0,\infty].$$

(1) *If $\sigma > 0$, then $K_t \leqslant \frac{1}{\sigma\sqrt{t}}\|\mathrm{e}^{X_t} - 1\|_{L_2(\mathbb{P})}$ for all $t > 0$.*
(2) *If there is an $m_t \in \mathbb{R}$ such that p_t is non-decreasing on $(-\infty, m_t)$ and non-increasing on (m_t, ∞), then $K_t \leqslant \mathbb{E}\mathrm{e}^{X_t}$.*

Proof. (1) Denote $J := X - \sigma W$. Let $p_t^{\sigma W}$ be the density of σW_t. Then, the independence of σW and J implies $p_t = p_t^{\sigma W} * \mathbb{P}_{J_t}$ for all $t > 0$. Choosing $\kappa = 1$ yields

$$\begin{aligned} K_t &\leqslant \int_{\mathbb{R}} |\mathrm{e}^x - 1|\, |\partial_x p_t(x)|\,\mathrm{d}x \leqslant \frac{1}{\sigma^2 t}\int_{\mathbb{R}}\int_{\mathbb{R}} |\mathrm{e}^x - 1|\,|x-y|\,p_t^{\sigma W}(x-y)\mathbb{P}_{J_t}(\mathrm{d}y)\mathrm{d}x \\ &= \frac{1}{\sigma^2 t}\int_{\mathbb{R}}\int_{\mathbb{R}} |\mathrm{e}^{x+y} - 1|\,|x|\,p_t^{\sigma W}(x)\mathrm{d}x\mathbb{P}_{J_t}(\mathrm{d}y) = \frac{1}{\sigma^2 t}\mathbb{E}\left|\sigma W_t\left(\mathrm{e}^{\sigma W_t + J_t} - 1\right)\right| \\ &\leqslant \frac{1}{\sigma t}\|W_t\|_{L_2(\mathbb{P})}\|\mathrm{e}^{X_t} - 1\|_{L_2(\mathbb{P})} = \frac{\|\mathrm{e}^{X_t} - 1\|_{L_2(\mathbb{P})}}{\sigma\sqrt{t}}. \end{aligned}$$

(2) We may assume that $\mathbb{E}\mathrm{e}^{X_t} < \infty$, otherwise the inequality is obvious. By the monotonicity of p_t, one has $p_t(x) \to 0$ as $|x| \to \infty$, and for $x > m_t + 1$,

$$\mathrm{e}^x p_t(x) \leqslant \mathrm{e}^x \int_{x-1}^{x} p_t(u)\mathrm{d}u \leqslant \mathrm{e}^x \int_{x-1}^{\infty} p_t(u)\mathrm{d}u \to 0 \quad \text{as } x \to \infty,$$

where the limit holds due to $\mathbb{E}e^{X_t} < \infty$. Now, choosing $\kappa = e^{m_t}$ and using integration by parts, together with $\lim_{|x|\to\infty} e^x p_t(x) = 0$, we have

$$\begin{aligned} K_t &\leqslant \int_{\mathbb{R}} |e^x - e^{m_t}|\,|\partial_x p_t(x)|\,dx \\ &= \int_{-\infty}^{m_t} (e^{m_t} - e^x)\partial_x p_t(x)dx + \int_{m_t}^{\infty} (e^x - e^{m_t})(-\partial_x p_t(x))dx \\ &= \int_{\mathbb{R}} e^x p_t(x)dx = \mathbb{E}e^{X_t}. \end{aligned}$$

□

Proposition 8.5 is an extension of [18, Theorem 8.9] to the exponential Lévy setting. Because of the weighted setting caused by the exponential Lévy process, it seems that the interpolation techniques using in [18, Theorem 8.9] cannot be applied, at least in a straightforward way. We recall the classes $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ of stable-like Lévy measures from Definition 4.4.

Proposition 8.5. *Let $g \in C^{0,\eta}(\mathbb{R}_+)$ with $\eta \in [0,1]$. Then, for $T \in (0,\infty)$ there exists a constant $c_{(8.4)} > 0$ such that for any $z > 0, y > 0$ and any $t \in (0,T]$ one has*

$$|P_t g(z) - P_t g(y)| \leqslant c_{(8.4)} U_t(y,z), \tag{8.4}$$

where the cases for $U_t(y,z)$ are provided as follows:

(1) *If $\sigma > 0$ and $\int_{|x|>1} e^{2x}\nu(dx) < \infty$, then $U_t(y,z) = \left(t^{\frac{\eta-1}{2}} \frac{|z^\eta - y^\eta|}{\eta}\right) \wedge |z-y|^\eta$.*

(2) *When $\sigma = 0$ and $\int_{|x|>1} e^x \nu(dx) < \infty$:*

   (a) *If $\nu \in \mathcal{S}_1(\alpha)$ for some $\alpha \in (0,2)$, then $U_t(y,z) = \left(t^{\frac{\eta-1}{\alpha}} \frac{|z^\eta - y^\eta|}{\eta}\right) \wedge |z-y|^\eta$.*

   (b) *If $\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in (0,2)$ and $g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$, then*

$$U_t(y,z) = \left(t^{\frac{\eta-1}{\alpha}} |\ln z - \ln y|^{1-\eta}|z-y|^\eta\right) \wedge |z-y|^\eta.$$

Here, we set $0^0 := 1$ and $\frac{|z^0 - y^0|}{0} := \lim_{\eta\downarrow 0} \frac{|z^\eta - y^\eta|}{\eta} = |\ln z - \ln y|$ by convention.

Proof. For $r \in \mathbb{R}$, since $e^{-t\psi(-ir)} = \mathbb{E}e^{rX_t} < \infty$ for all $t > 0$ if and only if $\int_{|x|>1} e^{rx}\nu(dx) < \infty$, it follows from the integrability conditions for ν in items (1) and (2) that $C^{0,\eta}(\mathbb{R}_+) \subseteq \mathcal{D}_{\exp}$ for any $\eta \in [0,1]$. Let $T \in (0,\infty)$. Then, the Hölder continuity of g implies that, for any $t \in [0,T]$ and $z > 0$, $y > 0$,

$$|P_t g(z) - P_t g(y)| \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)} \mathbb{E}e^{\eta X_t}|z-y|^\eta \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)} e^{T|\psi(-i\eta)|}|z-y|^\eta. \tag{8.5}$$

(1) Set $J := X - \sigma W$. Let $p_t^{\sigma W}$ (resp. p_t) be the probability density of σW_t (resp. X_t). For $t \in (0,T]$, since $p_t = p_t^{\sigma W} * \mathbb{P}_{J_t}$, one has

$$\|\partial_x p_t\|_{L_1(\mathbb{R})} = \|\partial_x p_t^{\sigma W} * \mathbb{P}_{J_t}\|_{L_1(\mathbb{R})} \leqslant \|\partial_x p_t^{\sigma W}\|_{L_1(\mathbb{R})} = \sqrt{2/(\pi\sigma^2 t)}.$$

It is clear that $p_t \in L_1(\mathbb{R}, \lambda_\eta)$, and similar computations as in the proof of Lemma 8.4(1) show $\partial_x p_t \in L_1(\mathbb{R}, \lambda_\eta)$. Hence, the assumptions for p_t required in Proposition 8.2 are satisfied. Furthermore, we have $e^{-t\psi(-i)} = \mathbb{E}e^{X_t} < \infty$ and $e^{-t\psi(-2i)} = \mathbb{E}e^{2X_t} < \infty$ for all $t \in (0,T]$, and hence

$$\mathbb{E}|e^{X_t} - 1|^2 = \mathbb{E}e^{2X_t} - 2\mathbb{E}e^{X_t} + 1 = e^{-t\psi(-2i)} - 2e^{-t\psi(-i)} + 1,$$

which implies

$$c_{(8.6)}^2 := \sup\nolimits_{t\in(0,T]} \left(t^{-1}\mathbb{E}|e^{X_t} - 1|^2\right) < \infty. \tag{8.6}$$

Then, for $\eta \in [0,1]$, $t \in (0,T]$, $z > 0$, $y > 0$, combining (8.1) with Lemma 8.4(1) yields

$$|P_t g(z) - P_t g(y)| \leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)} \|\partial_x p_t\|_{L_1(\mathbb{R})}^{1-\eta} \frac{c_{(8.6)}^{\eta}}{\sigma^{\eta}} \frac{|z^{\eta} - y^{\eta}|}{\eta} \leqslant c_{(8.7)} t^{\frac{\eta-1}{2}} \frac{|z^{\eta} - y^{\eta}|}{\eta}, \quad (8.7)$$

where $c_{(8.7)} := |g|_{C^{0,\eta}(\mathbb{R}_+)} c_{(8.6)}^{\eta} \frac{1}{\sigma} (\frac{2}{\pi})^{(1-\eta)/2}$. Then, (8.5) and (8.7) imply the assertion.

(2a) Let $\nu = \nu_1 + \nu_2$ for ν_1, ν_2 as in Definition 4.4. Assume that X^1 and X^2 are independent Lévy processes defined on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ with characteristics $(0,0,\nu_1)$ and $(\gamma, 0, \nu_2)$ respectively. Then, X and $X^1 + X^2$ have the same finite-dimensional distribution. Since $\int_{|x|>1} \mathrm{e}^x \nu_i(\mathrm{d}x) \leqslant \int_{|x|>1} \mathrm{e}^x \nu(\mathrm{d}x) < \infty$, $i = 1,2$, it implies that $\widetilde{\mathbb{E}} \mathrm{e}^{X_t^1} < \infty$ and $\widetilde{\mathbb{E}} \mathrm{e}^{X_t^2} < \infty$ for all $t \in (0,T]$.

Because of the conditions imposed on ν_1, it is straightforward to check that (8.3) is satisfied for the characteristic exponent of X^1. According to Lemma 8.3, X^1 has transition densities $(p_t^1)_{t>0} \subset C_0^{\infty}(\mathbb{R})$ with $\partial_x^n p_t^1 \in \cap_{1 \leqslant s \leqslant \infty} L_s(\mathbb{R})$ for all $n \geqslant 0$, $t \in (0,T]$ and there is a constant $c_{(8.8)} > 0$ such that

$$\|\partial_x p_t^1\|_{L_1(\mathbb{R})} \leqslant c_{(8.8)} t^{-\frac{1}{\alpha}}, \quad t \in (0,T]. \quad (8.8)$$

Since X^1 is selfdecomposable (see [33, Sec.53]), applying [33, Theorem 53.1] yields that $\widetilde{\mathbb{P}}_{X_t^1}$ is unimodal for all $t \in (0,T]$. Let m_t be a mode of $\widetilde{\mathbb{P}}_{X_t^1}$ so that the density p_t^1 of X_t^1 is non-decreasing on $(-\infty, m_t)$ and non-increasing on (m_t, ∞). Lemma 8.4(2) gives

$$\inf_{\kappa>0} \int_{\mathbb{R}} |\mathrm{e}^x - \kappa| |\partial_x p_t^1(x)| \mathrm{d}x \leqslant \widetilde{\mathbb{E}} \mathrm{e}^{X_t^1}, \quad t \in (0,T].$$

A similar argument as in the proof of Lemma 8.4(2) yields $\partial_x p_t^1 \in L_1(\mathbb{R}, \lambda_\eta)$. Hence, for $t \in (0,T]$ and $z > 0$, $y > 0$, using the independence of X^1 and X^2, together with Proposition 8.2, we get

$$\begin{aligned}
&|P_t g(z) - P_t g(y)| = |\mathbb{E} g(z\mathrm{e}^{X_t}) - \mathbb{E} g(y\mathrm{e}^{X_t})| = |\widetilde{\mathbb{E}} g(z\mathrm{e}^{X_t^2} \mathrm{e}^{X_t^1}) - \widetilde{\mathbb{E}} g(y\mathrm{e}^{X_t^2} \mathrm{e}^{X_t^1})| \\
&\leqslant \widetilde{\mathbb{E}} \left[\left(|g|_{C^{0,\eta}(\mathbb{R}_+)} \|\partial_x p_t^1\|_{L_1(\mathbb{R})}^{1-\eta} \inf_{\kappa>0} \left| \int_{\mathbb{R}} |\mathrm{e}^x - \kappa| |\partial_x p_t^1(x)| \mathrm{d}x \right|^{\eta} \right) \frac{|(z\mathrm{e}^{X_t^2})^{\eta} - (y\mathrm{e}^{X_t^2})^{\eta}|}{\eta} \right] \\
&\leqslant \left(|g|_{C^{0,\eta}(\mathbb{R}_+)} \|\partial_x p_t^1\|_{L_1(\mathbb{R})}^{1-\eta} |\widetilde{\mathbb{E}} \mathrm{e}^{X_t^1}|^{\eta} \right) \frac{|z^{\eta} - y^{\eta}|}{\eta} \widetilde{\mathbb{E}} \mathrm{e}^{\eta X_t^2} \\
&\leqslant |\mathbb{E} \mathrm{e}^{X_t}|^{\eta} |g|_{C^{0,\eta}(\mathbb{R}_+)} c_{(8.8)}^{1-\eta} t^{\frac{\eta-1}{\alpha}} \frac{|z^{\eta} - y^{\eta}|}{\eta} \\
&\leqslant c_{(8.9)} t^{\frac{\eta-1}{\alpha}} \frac{|z^{\eta} - y^{\eta}|}{\eta}, \qquad (8.9)
\end{aligned}$$

where $c_{(8.9)} := \mathrm{e}^{\eta T |\psi(-\mathrm{i})|} |g|_{C^{0,\eta}(\mathbb{R}_+)} c_{(8.8)}^{1-\eta}$. Combining (8.9) with (8.5) yields the assertion.

(2b) The assumption $\nu \in \mathcal{S}_2(\alpha)$ means that (8.3) is satisfied. Hence, X has transition densities $(p_t)_{t>0} \subset C_0^{\infty}(\mathbb{R})$ with

$$\sup_{t \in (0,T]} t^{\frac{1}{\alpha}} \|\partial_x p_t\|_{L_1(\mathbb{R})} < \infty.$$

Since we aim to apply Proposition 8.1 with $r = \frac{1}{\eta}$, let us first estimate $\|p_t\|_{L_r(\mathbb{R},\lambda_1)}$ for $r = \frac{1}{\eta}$ and $\eta \in [0,1]$. For the case $\eta \in (0,1]$, we have

$$\|p_t\|_{L_r(\mathbb{R},\lambda_1)}^r = \int_{\mathbb{R}} |p_t(x)|^r \mathrm{e}^x \mathrm{d}x \leqslant \|p_t\|_{L_\infty(\mathbb{R})}^{r-1} \mathbb{E} \mathrm{e}^{X_t} \leqslant \mathrm{e}^{T|\psi(-\mathrm{i})|} \|p_t\|_{L_\infty(\mathbb{R})}^{r-1}.$$

Since $p_t \in C_0^\infty(\mathbb{R})$, it holds that $\|p_t\|_{L_\infty(\mathbb{R})} \leqslant \|\partial_x p_t\|_{L_1(\mathbb{R})}$. Hence, there exists a constant $c_{(8.10)} > 0$ such that

$$\|p_t\|_{L_r(\mathbb{R},\lambda_1)} \leqslant c_{(8.10)} t^{\frac{\eta-1}{\alpha}}, \quad t \in (0,T]. \tag{8.10}$$

Since $\|p_t\|_{L_\infty(\mathbb{R},\lambda_1)} = \|p_t\|_{L_\infty(\mathbb{R})}$, inequality (8.10) also holds for the case $\eta = 0$, $r = \infty$.

Now we apply Proposition 8.1 with $r = \frac{1}{\eta}$ to obtain that, for $t \in (0,T]$, $z > 0$, $y > 0$,

$$\begin{aligned} |P_t g(z) - P_t g(y)| &\leqslant |g|_{\mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)} \|p_t\|_{L_r(\mathbb{R},\lambda_1)} |z-y|^\eta |\ln z - \ln y|^{1-\eta} \\ &\leqslant |g|_{\mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)} c_{(8.10)} t^{\frac{\eta-1}{\alpha}} |\ln z - \ln y|^{1-\eta} |z-y|^\eta. \end{aligned} \tag{8.11}$$

Combining (8.11) with (8.5), we derive the desired conclusion. □

8.3. **Estimate for the gradient in the GKW decomposition.** Motivated by the formula (4.2), for a Lévy measure ℓ and a Borel function g let us write symbolically

$$\mathit{\Gamma}_\ell(t,y) := \sigma^2 \partial_y P_t g(y) + \int_{\mathbb{R}} \frac{P_t g(\mathrm{e}^x y) - P_t g(y)}{y} (\mathrm{e}^x - 1)\ell(\mathrm{d}x) \tag{8.12}$$

for $(t,y) \in \mathbb{R}_+^2$, where we set $\partial_y P_t g(y) := 0$ 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 various contexts (e.g., see [38]).

Proposition 8.6(3)–(4) below are variants of [18, Theorem 8.12] in the exponential Lévy setting. Here, the exponent of the time variable t in the estimates we obtain is the same as in [18, Theorem 8.12]. Again, we recall $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ from Definition 4.4.

Proposition 8.6. *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 constant $c_{(8.13)} > 0$ such that*

$$|\mathit{\Gamma}_\ell(t,y)| \leqslant c_{(8.13)} V_t y^{\eta-1}, \quad \forall (t,y) \in (0,T] \times \mathbb{R}_+, \tag{8.13}$$

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 $\int_{|x|\leqslant 1} |x|^{\eta+1}\ell(\mathrm{d}x) < \infty$, then $V_t = 1$.*
3. *If $\sigma = 0$ and if the following two conditions hold:*
   (a) *$\nu \in \mathcal{S}_1(\alpha)$ for some $\alpha \in (0,2)$ and $\int_{|x|>1} \mathrm{e}^{x}\nu(\mathrm{d}x) < \infty$,*
   (b) *there is a $\beta \in (1+\eta, 2]$ such that*
   $$0 < \sup_{r \in (0,1]} r^\beta \int_{|x|\leqslant 1} \left(\left|\frac{x}{r}\right|^2 \wedge \left|\frac{x}{r}\right|^{\eta+1} \right) \ell(\mathrm{d}x) < \infty, \tag{8.14}$$
   then one has $V_t = t^{\frac{\eta+1-\beta}{\alpha}}$.
4. *If $\sigma = 0$ and $g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$, and if the following two conditions hold:*
   (a) *$\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in (0,2)$ and $\int_{|x|>1} \mathrm{e}^{x}\nu(\mathrm{d}x) < \infty$,*
   (b) *there is a $\beta \in (1+\eta, 2]$ such that (8.14) is satisfied,*
   then one has $V_t = t^{\frac{\eta+1-\beta}{\alpha}}$.

Here, the constant $c_{(8.13)}$ may depend on β in items (3) and (4).

Remark 8.7. Since $|\frac{x}{r}|^2 \wedge |\frac{x}{r}|^{\eta+1} \leqslant |\frac{x}{r}|^\beta$ for $\beta \in (1+\eta, 2]$, a sufficient condition for (8.14) is that $0 < \int_{|x|\leqslant 1} |x|^\beta \ell(\mathrm{d}x) < \infty$.

Proof of Proposition 8.6. In the sequel, we use the following inequality without mentioning it again:

$$\frac{|\mathrm{e}^{\eta x}-1|}{\eta} \leqslant \mathrm{e}^{\eta}|x|, \quad \forall |x|\leqslant 1, \eta\in[0,1],$$

where $\frac{|\mathrm{e}^{0x}-1|}{0} := \lim_{\eta\downarrow 0}\frac{|\mathrm{e}^{\eta x}-1|}{\eta} = |x|$. Let us fix $T\in(0,\infty)$.

(1) Since $\sigma>0$ and $\int_{|x|>1}\mathrm{e}^{2x}\nu(\mathrm{d}x)<\infty$, Proposition 8.5(1) implies that

$$|P_t g(z)-P_t g(y)| \leqslant c_{(8.4)}\left(\left(t^{\frac{\eta-1}{2}}\frac{|z^\eta-y^\eta|}{\eta}\right)\wedge|z-y|^\eta\right) \tag{8.15}$$

for all $z>0$, $y>0$, $t\in(0,T]$. Moreover, since $P_t g\in C^\infty(\mathbb{R}_+)$ due to $\sigma>0$, we divide both side of (8.15) by $|z-y|$ and then let $z\to y$ to obtain that

$$|\partial_y P_t g(y)| \leqslant c_{(8.4)} t^{\frac{\eta-1}{2}} y^{\eta-1}, \quad \forall (t,y)\in(0,T]\times\mathbb{R}_+.$$

Hence, we separate $\int_{\mathbb{R}} = \int_{|x|\leqslant 1}+\int_{|x|>1}$ and apply (8.15) with $z=y\mathrm{e}^x$ to obtain

$$\begin{aligned}|\mathit{\Gamma}_\ell(t,y)| &\leqslant c_{(8.4)}\left(\sigma^2+\int_{|x|\leqslant 1}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}|\mathrm{e}^x-1|\ell(\mathrm{d}x)\right)t^{\frac{\eta-1}{2}}y^{\eta-1}\\ &\qquad + c_{(8.4)}y^{\eta-1}\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x).\end{aligned} \tag{8.16}$$

Since $0<\sigma^2+\int_{|x|\leqslant 1}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}|\mathrm{e}^x-1|\ell(\mathrm{d}x)\leqslant \sigma^2+\mathrm{e}^{\eta+1}\int_{|x|\leqslant 1}|x|^2\ell(\mathrm{d}x)<\infty$ and $\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)<\infty$, together with $\inf_{t\in(0,T]}t^{\frac{\eta-1}{2}}>0$, the second term on the right-hand side of (8.16) can be upper bounded by the first term up to a positive constant. Hence, the desired conclusion follows.

(2) One has $\mathrm{e}^{-t\psi(-\eta\mathrm{i})} = \mathbb{E}\mathrm{e}^{\eta X_t}<\infty$ for $t>0$. The Hölder continuity of g implies that $|P_t g(\mathrm{e}^x y)-P_t g(y)|\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\mathbb{E}\mathrm{e}^{\eta X_t}|\mathrm{e}^x-1|^\eta y^\eta$, and hence

$$\begin{aligned}|\mathit{\Gamma}_\ell(t,y)| &\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\mathbb{E}\mathrm{e}^{\eta X_t}y^{\eta-1}\int_{\mathbb{R}}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\\ &\leqslant |g|_{C^{0,\eta}(\mathbb{R}_+)}\mathrm{e}^{T|\psi(-\eta\mathrm{i})|}\left(\mathrm{e}^{\eta+1}\int_{|x|\leqslant 1}|x|^{\eta+1}\ell(\mathrm{d}x)+\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\right)y^{\eta-1},\end{aligned}$$

which implies the assertion.

(3) Let $t\in(0,T]$ and $y>0$. We separate $\int_{\mathbb{R}} = \int_{|x|\leqslant 1}+\int_{|x|>1}$, and then apply Proposition 8.5(2a) with $z=y\mathrm{e}^x$ to obtain

$$\begin{aligned}|\mathit{\Gamma}_\ell(t,y)| &\leqslant c_{(8.4)}y^{\eta-1}\Bigg(\int_{|x|\leqslant 1}\left(\left(t^{\frac{\eta-1}{\alpha}}\frac{|\mathrm{e}^{\eta x}-1|}{\eta}\right)\wedge|\mathrm{e}^x-1|^\eta\right)|\mathrm{e}^x-1|\ell(\mathrm{d}x)\\ &\qquad\qquad + \int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\Bigg)\\ &\leqslant c_{(8.4)}y^{\eta-1}\Bigg(\mathrm{e}^{\eta+1}t^{\frac{\eta+1}{\alpha}}\int_{|x|\leqslant 1}\left(\left|\frac{x}{t^{1/\alpha}}\right|^2\wedge\left|\frac{x}{t^{1/\alpha}}\right|^{\eta+1}\right)\ell(\mathrm{d}x)\\ &\qquad\qquad + \int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\Bigg)\\ &\leqslant c_{(8.4)}y^{\eta-1}\left(c_{(8.17)}t^{\frac{\eta+1-\beta}{\alpha}}+\int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\right),\end{aligned} \tag{8.17}$$

where $c_{(8.17)} := \mathrm{e}^{\eta+1}(T^{\frac{\beta-2}{\alpha}} \vee T^{\frac{\beta-\eta-1}{\alpha}}) \sup_{r\in(0,1]} r^{\beta} \int_{|x|\leqslant 1}(|\frac{x}{r}|^2 \wedge |\frac{x}{r}|^{\eta+1})\ell(\mathrm{d}x) \in (0,\infty)$ by (8.14). Since $\inf_{(t,\beta)\in(0,T]\times(1+\eta,2]} t^{\frac{\eta+1-\beta}{\alpha}} > 0$, the desired conclusion follows from (8.17).

(4) Let $t \in (0,T]$ and $y > 0$. We apply Proposition 8.5(2b) with $z = y\mathrm{e}^x$ and use the same argument as in the proof of item (3) to obtain

$$\begin{aligned} |\varGamma_\ell(t,y)| &\leqslant c_{(8.4)} y^{\eta-1}\Bigg(\int_{|x|\leqslant 1}\Big(\Big(t^{\frac{\eta-1}{\alpha}}|x|^{1-\eta}|\mathrm{e}^x-1|^{\eta}\Big)\wedge|\mathrm{e}^x-1|^{\eta}\Big)|\mathrm{e}^x-1|\ell(\mathrm{d}x) \\ &\qquad\qquad + \int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\Bigg) \\ &\leqslant c_{(8.4)} y^{\eta-1}\left(c_{(8.17)} t^{\frac{\eta+1-\beta}{\alpha}} + \int_{|x|>1}|\mathrm{e}^x-1|^{\eta+1}\ell(\mathrm{d}x)\right). \end{aligned}$$

Again, a similar argument as in the one after inequality (8.17) yields the assertion. □

Acknowledgment. The author would like to sincerely thank his Ph.D. supervisors, Christel Geiss and Stefan Geiss, for helpful discussions.

References

1. D. Applebaum, *Lévy processes and stochastic calculus (2nd ed.)*, Cambridge Univ. Press, Cambridge, 2009.
2. F. Baumgartner, S. Geiss, Permutation invariant functionals of Lévy processes, *Trans. Amer. Math. Soc.* 369, 2017, 8607–8641.
3. F. Benth, G. Di Nunno, A. Løkka, B. Øksendal, F. Proske, Explicit representation of the minimal variance portfolio in markets driven by Lévy processes, *Math. Finance* 13(1), 2003, 55–72.
4. R. Blumenthal, R. Getoor, Sample functions of stochastic processes with stationary independent increments, *J. Math. Mech.* 10, 1961, 493–516.
5. M. Brodén, P. Tankov, Tracking errors from discrete hedging in exponential Lévy model, *Int. J. Theor. Appl. Finance* 14(6), 2011, 803–837.
6. T. Choulli, L. Krawczyk, C. Stricker, $\mathcal{E}$-martingales and their applications in mathematical finance, *Ann. Probab.* 26(2), 1998, 853–876.
7. R. Cont, P. Tankov, *Financial modeling with jump processes*, Chapman & Hall/CRC Press, 2003.
8. R. Cont, P. Tankov, E. Voltchkova, Hedging with options in models with jumps, In: *Stochastic Analysis and Applications – The Abel Symposium 2005*, Springer, Berlin, 2007.
9. S. Dereich, F. Heidenreich, A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations, *Stochastic Process. Appl.* 121, 2011, 1565–1587.
10. F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, C. Stricker, Weighted norm inequalities and hedging in incomplete markets, *Finance Stoch.* 1, 1997, 181–227.
11. C. Dellacherie, P.-A. Meyer, *Probabilities and potential B: Theory of martingales*, North-Holland Publishing Co., 1982.
12. M. Fukasawa, Asymptotically efficient discrete hedging. In: *Stochastic Analysis with Financial Applications. Progress in Probability*, 65, 331–346, Springer, Basel, 2011.
13. M. Fukasawa, Efficient discretization of stochastic integrals, *Finance Stoch.* 18, 2014, 175–208.
14. C. Geiss, S. Geiss, On approximation of a class of stochastic integrals and interpolation, *Stochastics Stochastics Rep.* 76(4), 2004, 339–362.
15. C. Geiss, S. Geiss, E. Laukkarinen, A note on Malliavin fractional smoothness for Lévy processes and approximation, *Potential Anal.* 39, 2013, 203–230.
16. C. Geiss, A. Steinicke, Malliavin derivative of random functions and applications to Lévy driven BSDEs, *Electron. J. Probab.* 21(10), 2016, 28 pp.
17. S. Geiss, Weighted BMO and discrete time hedging within the Black-Scholes model, *Probab. Theory Related Fields* 132, 2005, 13–38.
18. S. Geiss, N. T. Thuan, On Riemann–Liouville operators, BMO, gradient estimates in the Lévy–Itô space, and approximation, 2020. arXiv:2009.00899.
19. S. Geiss, A. Toivola, On fractional smoothness and L_p-approximation on the Gaussian space, *Ann. Probab.* 43(2), 2015, 605–638.

20. S. Geiss, A. Toivola, Weak convergence of error processes in discretization of stochastic integrals and Besov spaces, *Bernoulli* 15(4), 925–954, 2009.
21. E. Gobet, E. Temam, Discrete time hedging errors for options with irregular pay-offs, *Finance Stoch.* 5(3), 2001, 357–367.
22. E. Gobet, N. Landon, Almost sure optimal hedging strategy, *Ann. Appl. Probab.* 24(4), 2014, 1652–1690.
23. F. Hubalek, J. Kallsen, L. Krawczyk, Variance-optimal hedging for processes with stationary independent increments, *Ann. Appl. Probab.* 16(2), 2006, 853–885.
24. K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, *Trans. Amer. Math. Soc.* 81, 1956, 253–263.
25. J. Jacob, S. Méléard, P. Protter, Explicit form and robustness of martingale representations, *Ann. Probab.* 28(4), 2000, 1747–1780.
26. J. Jacod, A. Shiryaev, *Limit theorems for stochastic processes (2nd ed.)*, Springer, Berlin, 2003.
27. E. Laukkarinen, *On Malliavin calculus and approximation of stochastic integrals for Lévy process*, Ph.D. thesis, University of Jyväskylä, 2013.
28. A. Løkka, Martingale representation of functionals of Lévy processes, *Stoch. Anal. Appl.* 22(4), 2004, 867–892.
29. P. Malliavin, H. Airault, L. Kay, G. Letac, *Integration and probability*, Springer, New York, 1995.
30. G. Di Nunno, B. Øksendal, F. Proske, *Malliavin calculus for Lévy processes with applications to finance*, Springer-Verlag Berlin Heidelberg, 2009.
31. P. Protter, *Stochastic integration and differential equations (2nd ed., ver. 2.1)*, Springer-Verlag, Berlin, 2005.
32. M. Rosenbaum, P. Tankov, Asymptotically optimal discretization of hedging strategies with jumps, *Ann. Appl. Probab.* 24(3), 2014, 1002–1048.
33. K. Sato, *Lévy processes and infinitely divisible distributions (2nd ed.)*, Cambridge Univ. Press, Cambridge, 2013.
34. M. Schweizer, A guided tour through quadratic hedging approaches, In: *Option pricing, interest rates and risk management*, Cambridge Univ. Press, Cambridge, 2001.
35. J. Solé, F. Utzet, J. Vives, Chaos expansion and Malliavin calculus for Lévy processes, In: *Stochastic Analysis and Applications – The Abel Symposium 2005*, Springer, Berlin, 2007.
36. P. Tankov, *Pricing and hedging in exponential Lévy models: review of recent results*, Paris-Princeton Lecture Notes in Mathematical Finance, Springer, 2010.
37. P. Tankov, E. Voltchkova, Asymptotic analysis of hedging errors in models with jumps, *Stochastic Process. Appl.* 119, 2009, 2004–2027.
38. N. T. Thuan, Explicit Föllmer–Schweizer decomposition and discrete-time hedging in exponential Lévy models. arXiv:2009.04328.
39. R. Zhang, *Couverture approchée des options Européennes*, Ph.D. thesis, École Nationale des Ponts et Chaussées, Paris, 1998.

[1]Department of Mathematics and Statistics, P.O.Box 35, FI-40014 University of Jyväskylä, Finland

Email address: `thuan.t.nguyen@jyu.fi`

[2]Department of Mathematics, Vinh University, 182 Le Duan, Vinh, Nghe An, Vietnam

EXPLICIT FÖLLMER–SCHWEIZER DECOMPOSITION AND DISCRETE-TIME HEDGING IN EXPONENTIAL LÉVY MODELS

NGUYEN TRAN THUAN

Abstract. In a financial market driven by an exponential Lévy process, an explicit representation is shown for the Föllmer–Schweizer decomposition of European type options, implying a closed-form expression of the corresponding local risk-minimizing strategies. Using a jump-adjusted approximation scheme, the error caused by discretizing the local risk-minimizing strategies is investigated in dependence of properties of the Lévy measure, the regularity of the pay-off function and the chosen random discretization times. The rate of this error as the number of expected discretization times increases is measured in weighted BMO spaces, implying also L_p-estimates. Moreover, the effect of a change of measure satisfying a reverse Hölder inequality is addressed.

1. Introduction

This article is concerned with hedging problems in financial markets driven by exponential Lévy processes. We investigate two problems corresponding to two typical types of risks for hedging an option. The first one comes from the incompleteness of the market. We consider the semimartingale setting and aim to determine an explicit form for the Föllmer–Schweizer decomposition of European type options which provides directly a *closed form* for the local risk-minimizing strategies (a similar closed form expression in the martingale setting has been established in [8, 19, 36, 37]). The second type of risk is due to the impossibility of continuously rebalancing a hedging portfolio which leads to the discrete-time hedging. The discretization error we measure in weighted bounded mean oscillation spaces from which one can achieve good distributional tail estimates such as a pth-order polynomial decay, $p \in (2, \infty)$.

Let us introduce some notations to state the main results. Let $T \in (0, \infty)$ be a fixed time horizon and $X = (X_t)_{t\in[0,T]}$ a Lévy process defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of X which satisfies the usual conditions (right continuity and completeness). Assume that $\mathcal{F} = \mathcal{F}_T$. Let $\sigma \geqslant 0$ be the coefficient of the standard Brownian component and ν the Lévy measure of X (see (2.1)). We assume that the underlying discounted price process is modelled by the exponential $S = \mathrm{e}^X$.

1.1. Explicit Föllmer–Schweizer (FS) decomposition.

Because models with jumps typically correspond to incomplete markets, in general there is no hedging strategy which is self-financing and replicates an option at maturity. Hence, one has to look for certain strategies that minimize some types of risk. In the current work, we choose the quadratic hedging approach which is a popular method to deal with the problem in

Date: November 7, 2020.

2010 *Mathematics Subject Classification.* Primary: 60H05, 41A25; Secondary: 60G51, 91G99.

Key words and phrases. Approximation of stochastic intgral, Föllmer–Schweizer decomposition, Lévy process, Local risk-minimizing, Weighted bounded mean oscillation.

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

models with jumps. We refer the reader to the survey article [34] for this approach. Two typical types of quadratic hedging strategies are the *local risk-minimizing* (LRM) strategies and the *mean-variance hedging* (MVH) strategies. Roughly speaking, the LRM strategy is mean-self-financing, replicates an option at maturity and minimizes the riskiness of the cost process locally in time, while the MVH strategy is self-financing and minimizes the *global* hedging error in the mean square sense. Both types of those strategies are intimately related to the so-called *FS decomposition*. Namely, in our (exponential Lévy) setting, the FS decomposition gives directly the LRM strategy, and the MVH strategy can be determined based on this decomposition. This article discusses the FS decomposition and focuses on the LRM strategies only.

Assume that S is square integrable so that it is a semimartingale satisfying the *structure condition*, and that the *mean-variance trade-off process* of S is deterministic and bounded (see Remark 4.3). Then, the FS decomposition of an $H \in L_2(\mathbb{P})$ is of the form

$$H = H_0 + \int_0^T \vartheta_t^H \mathrm{d}S_t + L_T^H, \tag{1.1}$$

where $H_0 \in \mathbb{R}$, ϑ^H is an admissible integrand (specified in (4.2)), and L^H is an $L_2(\mathbb{P})$-martingale starting at zero which is orthogonal to the martingale part of S. The integrand ϑ^H is called the LRM strategy of H, and it is unique up to a $\mathbb{P}\otimes\lambda$-null set where λ is the Lebesgue measure. A key tool to study the FS decomposition is the *minimal (signed) local martingale measure* for S (see [33]), and we denote this signed measure by $\mathbb{P}^*$ from now on. Recently, [6, Theorem 4.3] indicated that under a regularity condition for $\mathbb{P}^*$, we can determine the LRM strategy ϑ^H based on the martingale representation of H with respect to $\mathbb{P}^*$.

There are many works interested in finding an explicit representation for the FS decomposition and the LRM strategy in the semimartingale framework (see, e.g., [2, 16, 17, 20, 36]). In the exponential Lévy setting and in the case of a European type option $H = g(S_T)$, Hubalek et al. [17] assumed that the function g can be represented as an integral transform of finite complex measures from which one can determine a closed form for the LRM strategy. The key idea of this approach is the separation of the function g and the underlying price process S by using a kind of inverse Fourier transform. An advantage of this method is that one gains much flexibility for choosing the underlying Lévy process where there is no extra regularity required for the driving process S except some mild integrability.

As our first main result, Theorem 1.1 below provides a closed form for the LRM strategy ϑ^H of an $H = g(S_T)$. To obtain this result, except of some mild integrability conditions, we neither assume any regularity for the payoff function g nor require any extra condition for the small jump behavior of X. However, the price one has to pay is the condition that $\mathbb{P}^*$ exists as a true probability measure (see Assumption 4.5) which leads to a constraint for the characteristics of X. This result might be regarded as a counterpart of [17, Proposition 3.1] in which only the square integrability is required for S while the function g are supposed to be the integral transform of finite complex measures. The notation $\mathbb{E}^*$ below means the expectation with respect to $\mathbb{P}^*$.

Theorem 1.1. *Assume that X is not a.s. deterministic and $S = \mathrm{e}^X$ is square $\mathbb{P}$-integrable. Under Assumption 4.5, if $g\colon (0,\infty) \to \mathbb{R}$ is a Borel function with $\mathbb{E}^*|g(yS_t)| < \infty$ for all $(t, y) \in [0, T] \times (0,\infty)$ and $g(S_T) \in L_2(\mathbb{P}) \cap L_2(\mathbb{P}^*)$, then the following assertions hold:*

(1) *The LRM strategy ϑ^H corresponding to $H = g(S_T)$ is of the form*

$$\vartheta^H_t = \frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 \partial_y G^*(t, S_{t-}) + \int_{\mathbb{R}} \frac{G^*(t, \mathrm{e}^x S_{t-}) - G^*(t, S_{t-})}{S_{t-}} (\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right) \quad (1.2)$$

for $\mathbb{P}\otimes\lambda$-a.e. $(\omega, t) \in \Omega \times [0, T]$, where $\|(\sigma,\nu)\| := \sigma^2 + \int_{\mathbb{R}}(\mathrm{e}^x - 1)^2\nu(\mathrm{d}x) \in (0,\infty)$, $G^*(t, y) := \mathbb{E}^* g(yS_{T-t})$, and we set $\partial_y G^* := 0$ when $\sigma = 0$ by convention.

(2) *There exists a process $\tilde{\vartheta}^g$ which is adapted and càdlàg on $[0, T)$, satisfies $\tilde{\vartheta}^g_- = \vartheta^H$ for $\mathbb{P}\otimes\lambda$-a.e. $(\omega, t) \in \Omega \times [0, T)$, and $\tilde{\vartheta}^g S$ is a $\mathbb{P}^*$-martingale.*

According to Theorem 1.1(2), $\tilde{\vartheta}^g_-$ is also a LRM strategy of $H = g(S_T)$, and one can determine it at every time $t \in [0, T)$ as showed in Remark 4.6 below. Furthermore, the càdlàg property of $\tilde{\vartheta}^g$ is useful to design some Riemann-type approximations for $\int_0^T \tilde{\vartheta}^g_{t-}\mathrm{d}S_t$. For example, an approximation scheme based on tracking jumps of $\tilde{\vartheta}^g$ has been constructed in [30]. We also employ the càdlàg version of the LRM strategy for the discrete-time hedging problem in Section 5. Such a path regularity for the integrand in the martingale setting was also studied in [24].

Some formulas resembling (1.2) have been established in [19, Formula (2.12)], [8, Formula (4.1)], [36, Formula (45)], or in [37, Formula (4.2)]. But in fact they are different. The formulas in [19, 8, 36, 37] were obtained by projecting H orthogonally down to the space of stochastic integrals driven by a (local) martingale, while the formula (1.2) is derived from the FS decomposition which is a different orthogonal decomposition in the semimartingale framework.

The proof of Theorem 1.1 is provided in Section 4, and the main tool we use is Proposition 1.2 where the square integrability of e^X is not necessarily assumed. We denote by W the standard Brownian motion and by $\widetilde{N}$ the compensated Poisson random measure appearing in the Lévy–Itô decomposition of X (see, e.g., [1, Theorem 2.4.16]).

Proposition 1.2. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a Borel function such that $\mathbb{E}|f(x + X_t)| < \infty$ for all $(t, x) \in [0, T] \times \mathbb{R}$. If $f(X_T) \in L_2(\mathbb{P})$, then*

$$\mathbb{E}\int_0^T |\sigma\partial_x F(t, X_{t-})|^2 \mathrm{d}t + \mathbb{E}\int_0^T\int_{\mathbb{R}} |F(t, X_{t-} + x) - F(t, X_{t-})|^2 \nu(\mathrm{d}x)\mathrm{d}t < \infty$$

and, a.s.,

$$\begin{aligned} f(X_T) = \mathbb{E}f(X_T) &+ \int_0^T \sigma\partial_x F(t, X_{t-})\mathrm{d}W_t \\ &+ \int_0^T\int_{\mathbb{R}\setminus\{0\}} (F(t, X_{t-} + x) - F(t, X_{t-}))\widetilde{N}(\mathrm{d}t, \mathrm{d}x), \end{aligned} \quad (1.3)$$

where $F(t, x) := \mathbb{E}f(x + X_{T-t})$ for $(t, x) \in [0, T] \times \mathbb{R}$, and we set $\partial_x F := 0$ if $\sigma = 0$.

Proposition 1.2 provides a martingale representation for functionals of X_T in which the integrands with respect to the Brownian part and the jump part are determined explicitly. Its proof is given in Section 3 by using Malliavin calculus. We also remark here that (1.3) is a Clark–Ocone type formula but $f(X_T)$ is not necessarily differentiable in the Malliavin sense.

Proposition 1.2 extends [8, Proposition 7] in which the function f has a polynomial growth and X satisfies a certain condition. A similar representation to (1.3) in a general framework (with different assumptions from ours) can be found in the proof of [19, Theorem 2.4]. On the other hand, when $f(X_T)$ is Malliavin differentiable then one can use the Clark–Ocone formula (see, e.g., [2, 3, 23]) to obtain its explicit martingale representation. However, the Malliavin differentiability of $f(X_T)$ fails to hold in many contexts. For example, if $f(x) = \mathbb{1}_{[K,\infty)}(x)$ for some $K \in \mathbb{R}$, and if X is of infinite variation and X_T has a density satisfying a mild condition, then $f(X_T)$ is not Malliavin differentiable (see [22, Theorem 6(b)]).

1.2. **Discrete-time hedging in weighted bounded mean oscillation (BMO) spaces.** We investigate the discrete-time approximation problem for stochastic integrals driving by the exponential Lévy process S. Let $E = (E_t)_{t\in[0,T]}$ be the error given by

$$E_t := \int_0^t \vartheta_{u-} \mathrm{d}S_u - A_t, \quad t \in [0,T],$$

where ϑ is an admissible integrand and $A = (A_t)_{t\in[0,T]}$ is an approximation scheme for the stochastic integral. In mathematical finance, the stochastic integral can be interpreted as the theoretical hedging portfolio which is continuously readjusted. However, in practice one can only rebalance the portfolio finitely many times, and this leads to a discretization of the stochastic integral, represented by A.

In case that $A = A^{\mathrm{Rm}}$ is the Riemann approximation process, the caused error $E = E^{\mathrm{Rm}}$ and its convergence rate have been investigated in the L_2-sense in several works. When S is assumed to be a martingale, the error was examined in [5, 11]. The error was also considered in a more general setting in [30] where the driving process is a local martingale with jumps. In general, the L_2-approach for the error yields a second-order polynomial decay for its distributional tail by Markov's inequality.

In the second part of this article, we aim to improve the distributional tail estimate for the approximation error by means of the *weighted bounded mean oscillation* (weighted BMO) approach. Moreover, the driving process S is not necessarily a (local) martingale but a semimartingale. To do this, we use the approximation scheme introduced in [37], the so-called *jump adjusted method* which was constructed by tracking jumps of the driving process S. Moreover, we show how the theory of weighted BMO spaces can be used to obtain L_p-estimates, $p \in (2,\infty)$, for the corresponding error. This approach also allows a change of the underlying measure which leaves the error estimates unchanged provided the change of measure satisfies a reverse Hölder inequality (see Proposition 5.3). The latter is frequently encountered in mathematical finance, and it is particularly useful here to switch the approximation problem between the martingale setting and the semimartingale setting.

The main results of the second part are Theorems 5.7 and 5.12 below. In Theorem 5.7, we provide several estimates for the error measured in weighted BMO-norms and describe a situation so that the L_p-estimate can be achieved for $p \in (2,\infty)$. Theorem 5.12 serves as an application of Theorem 5.7 where we consider the approximation problem for the stochastic integral term in (1.1) and the chosen integrand is the LRM strategy of a European type option. The results show how the interplay between the regularity of payoff functions and the small jumps intensity of the underlying Lévy process affects the convergence rate.

1.3. **Structure of the article.** We introduce the notation and recall Malliavin–Sobolev spaces and exponential Lévy processes in Section 2. The proof of Proposition 1.2 is contained in Section 3. Section 4 is devoted to prove Theorem 1.1. Section 5 presents the discrete-time hedging problem with the weighted BMO-approach for exponential Lévy models. Some technical results used in this article are given in Appendix A.

2. Preliminaries

2.1. **General notations.** Denote $\mathbb{R}_+ := (0,\infty)$ 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 $A,B \geqslant 0$ and $c \geqslant 1$, by $A \sim_c B$ we mean $\frac{1}{c}A \leqslant B \leqslant cA$. Subindexing a symbol by a label indicates the place where that symbol appears (e.g., $c_{(5.1)}$ refers to formula (5.1)).

Let $\mathcal{B}(\mathbb{R})$ be the Borel σ-algebra on $\mathbb{R}$. The Lebesgue measure on $\mathcal{B}(\mathbb{R})$ 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 space $L_p(A)$ consists of all p-order integrable Borel functions on A with respect to λ, where the essential supremum is taken when $p = \infty$. For a measure μ on $\mathcal{B}(\mathbb{R})$, its support is defined by

$$\operatorname{supp}\mu := \{x \in \mathbb{R} : \mu((x-\varepsilon, x+\varepsilon)) > 0, \forall \varepsilon > 0\}.$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\xi\colon \Omega \to \mathbb{R}$ a random variable. Denote by $\mathbb{P}_\xi$ the push-forward measure of $\mathbb{P}$ with respect to ξ. If ξ is integrable (non-negative), 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 set $L_p(\mathbb{P}) := L_p(\Omega, \mathcal{F}, \mathbb{P})$.

For a non-empty and open interval $U \subseteq \mathbb{R}$, let $C^\infty(U)$ denote the family of all functions f which have derivatives of all orders on U.

2.2. **Notation for stochastic processes.** Let $T > 0$ be a fixed finite time horizon, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be 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. The conditions imposed on $\mathbb{F}$ allow us to assume that every martingale adapted to this filtration is *càdlàg* (right continuous with left limits). We use the following notations and conventions where

$$\mathbb{I} = [0,T] \quad \text{or} \quad \mathbb{I} = [0,T).$$

- For processes $X = (X_t)_{t\in\mathbb{I}}$ and $Y = (Y_t)_{t\in\mathbb{I}}$, we write $X = Y$ to indicate that $X_t = Y_t$ for all $t \in \mathbb{I}$ a.s., and similarly when the relation "$=$" is replaced by some other standard relations such as "$\leqslant$", "$\geqslant$", etc.
- For a càdlàg process $X = (X_t)_{t\in\mathbb{I}}$, the process $X_- = (X_{t-})_{t\in\mathbb{I}}$ is defined by setting $X_{0-} := X_0$ and $X_{t-} := \lim_{0<s\uparrow t} X_s$ for $t \in \mathbb{I}\backslash\{0\}$. We set $\Delta X := X - X_-$.
- $\mathrm{CL}(\mathbb{I})$ denotes the family of all càdlàg and $\mathbb{F}$-adapted processes.
- $\mathrm{CL}_0(\mathbb{I})$ (resp. $\mathrm{CL}^+(\mathbb{I})$) consists of all $X \in \mathrm{CL}(\mathbb{I})$ with $X_0 = 0$ a.s. (resp. $X \geqslant 0$).
- For $p \in [1,\infty]$ and $X \in \mathrm{CL}([0,T])$, we set $\|X\|_{L_p(\mathbb{P})} := \|\sup_{t\in[0,T]}|X_t|\|_{L_p(\mathbb{P})}$.
- $\mathcal{P}$ is the predictable σ-algebra[1] on $\Omega \times [0,T]$ and $\widetilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{B}(\mathbb{R})$.

We recall some notions regarding semimartingales on the finite time interval $[0,T]$.

- A process $M \in \mathrm{CL}([0,T])$ is called a local (resp. locally square integrable) martingale if there is a sequence of non-decreasing stopping times $(\rho_n)_{n\geqslant 1}$ taking values in $[0,T]$ such that $\mathbb{P}(\rho_n < T) \to 0$ as $n \to \infty$ and the stopped process $M^{\rho_n} = (M_{t\wedge\rho_n})_{t\in[0,T]}$ is a martingale (resp. square integrable martingale) for all $n \geqslant 1$. Let $\mathcal{M}_2^0(\mathbb{P})$ be the space of all square integrable $\mathbb{P}$-martingales $M = (M_t)_{t\in[0,T]}$ with $M_0 = 0$ a.s.
- A process $S \in \mathrm{CL}([0,T])$ is called a semimartingale if S can be written as a sum of a local martingale and a process of finite variation a.s. The *quadratic covariation* of two semimartingales S and R is denoted by $[S,R]$. The predictable $\mathbb{Q}$-compensator of $[S,R]$, if it exists, is denoted by $\langle S,R\rangle^{\mathbb{Q}}$, where $\mathbb{Q}$ is a probability measure. We will omit the reference measure if there is no risk of confusion.

[1] $\mathcal{P}$ is the σ-algebra generated by $\{A \times \{0\} : A \in \mathcal{F}_0\} \cup \{A \times (s,t] : 0 \leqslant s < t \leqslant T, A \in \mathcal{F}_s\}$.

– Let M, N be locally square integrable martingales under a probability measure $\mathbb{Q}$. Then, M and N are said to be $\mathbb{Q}$-*orthogonal* if $[M,N]$ is a local martingale under $\mathbb{Q}$, or equivalently, $\langle M,N\rangle^{\mathbb{Q}}=0$.

2.3. **Lévy process and Itô's chaos expansion.** Let $X=(X_t)_{t\in[0,T]}$ be a real-valued Lévy process on a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$, i.e. $X_0=0$, X has independent and stationary increments and X has càdlàg paths. Let $\mathbb{F}=(\mathcal{F}_t)_{t\in[0,T]}$ denote the augmented natural filtration generated by X. From now on, we assume that $\mathcal{F}=\mathcal{F}_T$. According to the Lévy–Khintchine formula (see, e.g., [31, Theorem 8.1]), the *characteristic exponent* ψ of X, which is defined by

$$\mathbb{E}\mathrm{e}^{\mathrm{i}uX_t}=\mathrm{e}^{-t\psi(u)},\quad u\in\mathbb{R},t\in[0,T],$$

is of the form

$$\psi(u)=-\mathrm{i}\gamma u+\frac{\sigma^2u^2}{2}-\int_{\mathbb{R}}\left(\mathrm{e}^{\mathrm{i}ux}-1-\mathrm{i}ux\mathbb{1}_{\{|x|\leqslant 1\}}\right)\nu(\mathrm{d}x),\quad u\in\mathbb{R}. \tag{2.1}$$

Here, $\gamma\in\mathbb{R}$, while $\sigma\geqslant 0$ is the coefficient of the Brownian component, and $\nu\colon\mathcal{B}(\mathbb{R})\to[0,\infty]$ is a Lévy measure (i.e. $\nu(\{0\}):=0$ and $\int_{\mathbb{R}}(x^2\wedge 1)\nu(\mathrm{d}x)<\infty$). The triplet (γ,σ,ν) is also called the characteristics of X. To indicate explicitly the characteristics of X under $\mathbb{P}$, we write

$$(X|\mathbb{P})\sim(\gamma,\sigma,\nu)\quad\text{or}\quad(X|\mathbb{P})\sim\psi.$$

We present briefly the Malliavin calculus for Lévy processes by means of Itô's chaos expansion which is the main tool to prove Proposition 1.2. For further details, we refer to [35, 27, 28, 1] and the references therein. Define the σ-finite measures μ on $\mathcal{B}(\mathbb{R})$ and $\mathbb{m}$ on $\mathcal{B}([0,T]\times\mathbb{R})$ by setting

$$\mu(\mathrm{d}x):=\sigma^2\delta_0(\mathrm{d}x)+x^2\nu(\mathrm{d}x)\quad\text{and}\quad\mathbb{m}:=\lambda\otimes\mu,$$

where δ_0 is the Dirac measure at zero. For $B\in\mathcal{B}([0,T]\times\mathbb{R})$ with $\mathbb{m}(B)<\infty$, the random measure M is defined by

$$M(B):=\sigma\int_{\{t\in[0,T]:(t,0)\in B\}}\mathrm{d}W_t+L_2(\mathbb{P})\text{-}\lim_{n\to\infty}\int_{B\cap([0,T]\times\{\frac{1}{n}<|x|<n\})}x\widetilde{N}(\mathrm{d}t,\mathrm{d}x),$$

where W is the standard Brownian motion and $\widetilde{N}$ is the compensated Poisson random measure appearing in the Lévy–Itô decomposition of X (see, e.g., [1, Theorem 2.4.16]).

Set $L_2(\mu^0)=L_2(\mathbb{m}^0):=\mathbb{R}$. For $n\geqslant 1$, we denote

$$\begin{aligned}L_2(\mu^{\otimes n})&:=L_2(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),\mu^{\otimes n}),\\ L_2(\mathbb{m}^{\otimes n})&:=L_2(([0,T]\times\mathbb{R})^n,\mathcal{B}(([0,T]\times\mathbb{R})^n),\mathbb{m}^{\otimes n}).\end{aligned}$$

The multiple integral $I_n\colon L_2(\mathbb{m}^{\otimes n})\to L_2(\mathbb{P})$ is defined in the sense of Itô [18] by using an approximation argument, where it is given for simple functions as follows: For

$$\xi_n^m:=\sum_{k=1}^m a_k\mathbb{1}_{B_1^k\times\cdots\times B_n^k},$$

where $a_k\in\mathbb{R}$, $B_i^k\in\mathcal{B}([0,T]\times\mathbb{R})$ with $\mathbb{m}(B_i^k)<\infty$ and $B_i^k\cap B_j^k=\emptyset$ for $k=1,\dots,m$, $i,j=1,\dots,n$, $i\neq j$ and $m\geqslant 1$, we define

$$I_n(\xi_n^m):=\sum_{k=1}^m a_kM(B_1^k)\cdots M(B_n^k).$$

Then, [18, Theorem 2] asserts the following Itô chaos expansion

$$L_2(\mathbb{P}) = \bigoplus_{n=0}^{\infty} \{I_n(\xi_n) : \xi_n \in L_2(\mathbb{m}^{\otimes n})\},$$

where $I_0(\xi_0) := \xi_0 \in \mathbb{R}$. For $n \geqslant 1$, the symmetrization $\tilde{\xi}_n$ of a $\xi_n \in L_2(\mathbb{m}^{\otimes n})$ is

$$\tilde{\xi}_n((t_1, x_1), \ldots, (t_n, x_n)) := \frac{1}{n!} \sum_{\pi} \xi_n((t_{\pi(1)}, x_{\pi(1)}), \ldots, (t_{\pi(n)}, x_{\pi(n)})),$$

where the sum is taken over all permutations π of $\{1, \ldots, n\}$, so that $I_n(\xi_n) = I_n(\tilde{\xi}_n)$ a.s. The Itô chaos decomposition verifies that $\xi \in L_2(\mathbb{P})$ if and only if there are $\xi_n \in L_2(\mathbb{m}^{\otimes n})$ such that $\xi = \sum_{n=0}^{\infty} I_n(\xi_n)$ a.s., and this expansion is unique if every ξ_n is symmetric, i.e. $\xi_n = \tilde{\xi}_n$. Furthermore, $\|\xi\|^2_{L_2(\mathbb{P})} = \sum_{n=0}^{\infty} n! \|\tilde{\xi}_n\|^2_{L_2(\mathbb{m}^{\otimes n})}$.

Definition 2.1. Let $\mathbb{D}_{1,2}$ be the Malliavin–Sobolev space of all $\xi = \sum_{n=0}^{\infty} I_n(\xi_n) \in L_2(\mathbb{P})$ such that

$$\|\xi\|^2_{\mathbb{D}_{1,2}} := \sum_{n=0}^{\infty} (n+1)! \|\tilde{\xi}_n\|^2_{L_2(\mathbb{m}^{\otimes n})} < \infty.$$

The *Malliavin derivative operator* $D \colon \mathbb{D}_{1,2} \to L_2(\mathbb{P} \otimes \mathbb{m})$, where $L_2(\mathbb{P} \otimes \mathbb{m}) := L_2(\Omega \times [0, T] \times \mathbb{R}, \mathcal{F} \otimes \mathcal{B}([0, T] \times \mathbb{R}), \mathbb{P} \otimes \mathbb{m})$, is defined for $\xi = \sum_{n=0}^{\infty} I_n(\xi_n) \in \mathbb{D}_{1,2}$ by

$$D_{t,x}\xi := \sum_{n=1}^{\infty} n I_{n-1}(\tilde{\xi}_n((t, x), \cdot)), \quad (\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}.$$

2.4. **Exponential Lévy processes.** Let X be a Lévy process with $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$. The stochastic exponential of X, denoted by $\mathcal{E}(X)$, is the càdlàg process that satisfies the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{E}(X) = \mathcal{E}(X)_{-}\mathrm{d}X, \quad \mathcal{E}(X)_0 = 1.$$

We apply [1, Theorem 5.1.6] with the truncation function $x\mathbb{1}_{\{|x| \leqslant 1\}}$ instead of $x\mathbb{1}_{\{|x|<1\}}$ to obtain that if $\mathcal{E}(X) > 0$, then there exists a Lévy process Y with $(Y|\mathbb{P}) \sim (\gamma_Y, \sigma_Y, \nu_Y)$ such that $\mathcal{E}(X) = \mathrm{e}^Y$, where $\sigma_Y = \sigma$ and

$$\nu_Y(B) = \int_{\mathbb{R}} \mathbb{1}_{\{\ln(1+x) \in B\}} \nu(\mathrm{d}x), \quad B \in \mathcal{B}(\mathbb{R}),$$
$$\gamma_Y = \gamma - \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left(\mathbb{1}_{\{|\ln(1+x)| \leqslant 1\}} \ln(1+x) - x\mathbb{1}_{\{|x| \leqslant 1\}} \right) \nu(\mathrm{d}x).$$

Conversely, there is a Lévy process Z with $(Z|\mathbb{P}) \sim (\gamma_Z, \sigma_Z, \nu_Z)$ such that $\mathrm{e}^X = \mathcal{E}(Z)$. Moreover, one has $\sigma_Z = \sigma$ and

$$\nu_Z(B) = \int_{\mathbb{R}} \mathbb{1}_{\{\mathrm{e}^x - 1 \in B\}} \nu(\mathrm{d}x), \quad B \in \mathcal{B}(\mathbb{R}),$$
$$\gamma_Z = \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left((\mathrm{e}^x - 1)\mathbb{1}_{\{|\mathrm{e}^x - 1| \leqslant 1\}} - x\mathbb{1}_{\{|x| \leqslant 1\}} \right) \nu(\mathrm{d}x).$$

3. Martingale representation with explicit integrands

This section is devoted to prove Proposition 1.2 by using Malliavin calculus. There are two key observations: first, the kernels in the chaos expansion of $f(X_T) \in L_2(\mathbb{P})$ do not depend on the time variables which implies the Malliavin differentiability of $\mathbb{E}_{\mathcal{F}_t}[f(X_T)]$ for any $t \in [0, T)$ (see Lemma 3.3); secondly, the Malliavin derivative of a functional of X_t, provided it is Malliavin differentiable, can be expressed in an explicit form (see Lemma 3.2).

In this section, we assume $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$. The following lemma is taken from [15, Example 8.18(1)].

Lemma 3.1 ([15]). *Assume $\sigma > 0$. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a Borel function with $\mathbb{E}|f(X_T)|^q < \infty$ for some $q > 1$. Then, $\mathbb{E}|f(x+X_{T-t})| < \infty$ for all $(t,x) \in [0,T]\times\mathbb{R}$, and the function $x \mapsto F(t,x) := \mathbb{E}f(x+X_{T-t})$ belongs to $C^\infty(\mathbb{R})$ for any $t \in [0,T)$. Furthermore,*

$$\mathbb{E}_{\mathcal{F}_s}[\partial_x F(t, X_t)] = \partial_x F(s, X_s) \quad a.s.$$

for any $0 \leqslant s < t < T$.

Lemma 3.2 below was obtained in [21, Corollary 3.1 in the second article of this thesis] and it provides an equivalent condition such that a functional of X_t belongs to $\mathbb{D}_{1,2}$. We refer to [25, Proposition V 2.3.1] when X is a Brownian motion and refer to [12, Lemma 3.2] when X has no Brownian component.

Lemma 3.2 ([21]). *Let $t \in (0,T]$ and a Borel function $f\colon \mathbb{R} \to \mathbb{R}$ with $f(X_t) \in L_2(\mathbb{P})$. Then, $f(X_t) \in \mathbb{D}_{1,2}$ if and only if the following two assertions hold:*

(a) *when $\sigma > 0$, f has a weak derivative[2] f'_w on $\mathbb{R}$ with $f'_w(X_t) \in L_2(\mathbb{P})$,*

(b) *the map $(s,x) \mapsto \frac{f(X_t+x)-f(X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x)$ belongs to $L_2(\mathbb{P}\otimes\mathbb{m})$.*

Furthermore, if $f(X_t) \in \mathbb{D}_{1,2}$, then for $\mathbb{P}\otimes\mathbb{m}$-a.e. $(\omega,s,x) \in \Omega\times[0,T]\times\mathbb{R}$ one has

$$D_{s,x} f(X_t) = f'_w(X_t)\mathbb{1}_{[0,t]\times\{0\}}(s,x) + \frac{f(X_t+x)-f(X_t)}{x}\mathbb{1}_{[0,t]\times\mathbb{R}_0}(s,x),$$

where we set, by convention, $f'_w := 0$ when $\sigma = 0$.

Lemma 3.3. *Let $f\colon \mathbb{R}\to\mathbb{R}$ be a Borel function with $f(X_T) \in L_2(\mathbb{P})$.*

(1) *There are symmetric $\tilde{f}_n \in L_2(\mu^{\otimes n})$ such that $f(X_T) = \sum_{n=0}^\infty I_n(\tilde{f}_n \mathbb{1}_{[0,T]}^{\otimes n})$ a.s.*

(2) *For $t \in [0,T)$, one has $\mathbb{E}_{\mathcal{F}_t}[f(X_T)] = \sum_{n=0}^\infty I_n(\tilde{f}_n\mathbb{1}_{[0,t]}^{\otimes n})$ a.s. and $\mathbb{E}_{\mathcal{F}_t}[f(X_T)] \in \mathbb{D}_{1,2}$.*

(3) *For $t \in (0,T)$, it holds*

$$\mathbb{E}\left(|\sigma\partial_x F(t,X_t)|^2 + \int_{\mathbb{R}} |F(t,X_t+x) - F(t,X_t)|^2\nu(\mathrm{d}x)\right) < \infty, \tag{3.1}$$

where $F(t,x) := \mathbb{E}f(x+X_{T-t})$ if $\sigma > 0$, and in the case $\sigma = 0$ we let $F(t,\cdot)$ be a Borel function such that $F(t,X_t) = \mathbb{E}_{\mathcal{F}_t}[f(X_T)]$ a.s. and set $\partial_x F := 0$.

Proof. Items (1) and (2) are due to [15, Lemma D.1]. For item (3), it is clear for the case $\sigma = 0$ that (3.1) is implied by Lemma 3.2. Let us turn to the case $\sigma > 0$. According to Lemma 3.1, one has $F(t,\cdot) \in C^\infty(\mathbb{R})$, and hence $(F(t,\cdot))'_w = \partial_x F(t,\cdot)$ a.e. with respect to the Lebesgue measure λ. Since the law of X_t is absolutely continuous with respect to λ, it holds that $(F(t,\cdot))'_w(X_t) = \partial_x F(t,X_t)$ a.s. Then, (3.1) follows from Lemma 3.2. □

We are now in a position to prove Proposition 1.2.

[2] A locally integrable function h is called a *weak derivative* of a locally integrable funtion f on $\mathbb{R}$ if $\int_{\mathbb{R}} f(x)\phi'(x)\mathrm{d}x = -\int_{\mathbb{R}} h(x)\phi(x)\mathrm{d}x$ for all smooth functions ϕ with compact support in $\mathbb{R}$. When such an h exists (unique up to a λ-null set), then we denote $f'_w := h$.

Proof of Proposition 1.2. For $(t,x) \in [0,T] \times \mathbb{R}$, denote

$$\Delta F(t,x) := \partial_x F(t, X_{t-}) \mathbb{1}_{\{x=0\}} + \frac{F(t, X_{t-}+x) - F(t, X_{t-})}{x} \mathbb{1}_{\{x \neq 0\}}, \tag{3.2}$$

where we recall that $\partial_x F := 0$ if $\sigma = 0$ by convention. The assumption $\mathbb{E}|f(x+X_t)| < \infty$ for all $(t,x) \in [0,T] \times \mathbb{R}$ implies that $(F(t, X_t+x) - F(t, X_t))_{t \in [0,T]}$ is a martingale for each $x \in \mathbb{R}$. Moreover, in the case $\sigma > 0$, the assumption $f(X_T) \in L_2(\mathbb{P})$ and Lemma 3.1 imply that $F(t,\cdot) \in C^\infty(\mathbb{R})$ for all $t \in [0,T)$ and $(\partial_x F(t, X_t))_{t \in [0,T)}$ is a martingale.

Step 1. We show that for any $t \in (0,T)$,

$$C(t) := \mathbb{E} \int_0^t \int_{\mathbb{R}} |\Delta F(s,x)|^2 \mathbb{m}(\mathrm{d}s, \mathrm{d}x) < \infty.$$

Observe that $(t,x) \mapsto F(t,x)$ is Borel measurable by Fubini's theorem. In addition, since X_- is predictable, we infer that $(\omega, t, x) \mapsto F(t, X_{t-}(\omega) + x)$ is $\widetilde{\mathcal{P}}$-measurable. Therefore, ΔF given in (3.2) is $\widetilde{\mathcal{P}}$-measurable.

Remark that $X_s = X_{s-}$ a.s. for each $s \in [0,T]$. Using Fubini's theorem and the martingale property, together with (3.1), we obtain for any $t \in (0,T)$ that

$$\begin{aligned} C(t) &= \mathbb{E} \int_0^t |\sigma \partial_x F(s, X_s)|^2 \mathrm{d}s + \mathbb{E} \int_0^t \int_{\mathbb{R}} |F(s, X_s + x) - F(s, X_s)|^2 \nu(\mathrm{d}x) \mathrm{d}s \\ &\leqslant t \left(\mathbb{E}|\sigma \partial_x F(t, X_t)|^2 + \mathbb{E} \int_{\mathbb{R}} |F(t, X_t + x) - F(t, X_t)|^2 \nu(\mathrm{d}x) \right) \\ &< \infty. \end{aligned}$$

Hence, the stochastic integral $\int_0^t \int_{\mathbb{R}} \Delta F(s,x) M(\mathrm{d}s, \mathrm{d}x)$ exists as an element in $L_2(\mathbb{P})$.

Step 2. Fix $t \in (0,T)$. We prove that, a.s.,

$$F(t, X_t) = \mathbb{E} f(X_T) + \int_0^t \int_{\mathbb{R}} \Delta F(s,x) M(\mathrm{d}s, \mathrm{d}x). \tag{3.3}$$

The representation (3.3) can be regarded as a consequence of the Clark–Ocone formula. However, this formula seems to be considered either when the Lévy process X is square integrable or when X has no Brownian component (i.e. $\sigma = 0$) (see, e.g., [3, 23, 27, 28, 35]). So, for the reader's convenience, we present here a complete proof for (3.3) where neither square integrability nor $\sigma = 0$ is assumed. Due to the denseness of the simple multiple stochastic integrals in $L_2(\mathbb{P})$ (see [10, Lemma 2.1]), in order to obtain (3.3) it is sufficient to check that

$$\mathbb{E}\left[I_m(k_m) F(t, X_t)\right] = \mathbb{E}\left[I_m(k_m) \int_0^t \int_{\mathbb{R}} \Delta F(s,x) M(\mathrm{d}s, \mathrm{d}x)\right] \tag{3.4}$$

for all $m \geqslant 1$ and all functions k_m of the form

$$k_m = \mathbb{1}_{B_1 \times \cdots \times B_m}, \tag{3.5}$$

where $B_i = (s_i, t_i] \times (a_i, b_i]$ in which $(a_i, b_i]$ are finite intervals and the time intervals $(s_i, t_i] \subset [0,t]$ satisfy $t_{i-1} \leqslant s_i$, $i = 2, \ldots, m$.

Since $F(t, X_t) \in \mathbb{D}_{1,2}$ by Lemma 3.3(2), applying Lemma 3.2 we have for $\mathbb{P} \otimes \mathbb{m}$-a.e. $(\omega, s, x) \in \Omega \times [0,T] \times \mathbb{R}$,

$$D_{s,x} F(t, X_t) = \partial_x F(t, X_t) \mathbb{1}_{[0,t] \times \{0\}}(s,x) + \frac{F(t, X_t + x) - F(t, X_t)}{x} \mathbb{1}_{[0,t] \times \mathbb{R}_0}(s,x). \tag{3.6}$$

Moreover, for each $(s,x) \in [0,t] \times \mathbb{R}$, the martingale property implies that, a.s.,

$$\mathbb{E}_{\mathcal{F}_s}\left[\partial_x F(t, X_t) \mathbb{1}_{\{x=0\}} + \frac{F(t, X_t + x) - F(t, X_t)}{x} \mathbb{1}_{\{x \neq 0\}}\right]$$

$$= \partial_x F(s, X_s)\mathbb{1}_{\{x=0\}} + \frac{F(s, X_s + x) - F(s, X_s)}{x}\mathbb{1}_{\{x\neq 0\}}$$
$$= \Delta F(s, x),$$

where the second equality comes from the fact that $X_s = X_{s-}$ a.s.

We let $f(X_T) = \sum_{n=0}^{\infty} I_n(\tilde{f}_n \mathbb{1}_{[0,T]}^{\otimes n})$ and $F(t, X_t) = \sum_{n=0}^{\infty} I_n(\tilde{f}_n \mathbb{1}_{[0,t]}^{\otimes n})$ as in Lemma 3.3(1) and (2) respectively, where $\tilde{f}_n \in L_2(\mu^{\otimes n})$ are symmetric. Let k_m be of the form as in (3.5). Since functions $\tilde{f}_n$ are symmetric, the left-hand side of (3.4) is computed as follows

$$\mathrm{LHS}_{(3.4)} = m! \int_{B_1\times\cdots\times B_m} \tilde{f}_m(x_1, \ldots, x_m)\mathbb{m}(\mathrm{d}s_1, \mathrm{d}x_1)\cdots\mathbb{m}(\mathrm{d}s_m, \mathrm{d}x_m). \tag{3.7}$$

For the right-hand side of (3.4), writing $I_m(k_m) = \int_{B_m} I_{m-1}(k_{m-1})M(\mathrm{d}s, \mathrm{d}x)$, where $k_{m-1} := \mathbb{1}_{B_1\times\cdots\times B_{m-1}}$, and using Fubini's theorem we obtain

$$\begin{aligned}
\mathrm{RHS}_{(3.4)} &= \mathbb{E}\int_{B_m} I_{m-1}(k_{m-1})\Delta F(s, x)\mathbb{m}(\mathrm{d}s, \mathrm{d}x)\\
&= \int_{B_m} \mathbb{E}\left[I_{m-1}(k_{m-1})\mathbb{E}_{\mathcal{F}_s}\left[\partial_x F(t, X_t)\mathbb{1}_{\{x=0\}} + \frac{F(t, X_t + x) - F(t, X_t)}{x}\mathbb{1}_{\{x\neq 0\}}\right]\right]\mathbb{m}(\mathrm{d}s, \mathrm{d}x)\\
&= \mathbb{E}\int_{B_m} I_{m-1}(k_{m-1})D_{s,x}F(t, X_t)\mathbb{m}(\mathrm{d}s, \mathrm{d}x) \qquad (3.8)\\
&= \mathbb{E}\int_{B_m} I_{m-1}(k_{m-1})\left(L_2(\mathbb{P}\otimes\mathbb{m})\text{-}\lim_{j\to\infty}\sum_{i=1}^{j} iI_{i-1}\left(\tilde{f}_i(\cdot, x)\mathbb{1}_{[0,t]}^{(i-1)}\mathbb{1}_{[0,t]}(s)\right)\right)\mathbb{m}(\mathrm{d}s, \mathrm{d}x)\\
&= m\int_{B_m} \mathbb{E}\left[I_{m-1}(k_{m-1})I_{m-1}\left(\tilde{f}_m(\cdot, x)\mathbb{1}_{[0,t]}^{(m-1)}\mathbb{1}_{[0,t]}(s)\right)\right]\mathbb{m}(\mathrm{d}s, \mathrm{d}x)\\
&= m!\int_{B_m}\int_{B_1\times\cdots\times B_{m-1}} \tilde{f}_m(x_1, \ldots, x_{m-1}, x)\mathbb{m}(\mathrm{d}s_1, \mathrm{d}x_1)\cdots\mathbb{m}(\mathrm{d}s_{m-1}, \mathrm{d}x_{m-1})\mathbb{m}(\mathrm{d}s, \mathrm{d}x). \qquad (3.9)
\end{aligned}$$

Here, one uses (3.6) and the fact that $I_{m-1}(k_{m-1})$ is $\mathcal{F}_s$-measurable for all $s \in (s_m, t_m]$ to obtain (3.8). Combining (3.7) with (3.9) yields (3.4).

Step 3. For any $t \in (0, T)$, Jensen's inequality implies that $\mathbb{E}|f(X_T)|^2 \geqslant \mathbb{E}|F(t, X_t)|^2$. Then, we apply *Step 2* and Itô's isometry to obtain

$$\mathbb{E}|f(X_T)|^2 \geqslant |\mathbb{E}f(X_T)|^2 + \mathbb{E}\int_0^t\int_{\mathbb{R}} |\Delta F(s, x)|^2\mathbb{m}(\mathrm{d}s, \mathrm{d}x).$$

Letting $t \uparrow T$, we infer that the stochastic integral $\int_0^T\int_{\mathbb{R}} \Delta F(s, x)M(\mathrm{d}s, \mathrm{d}x)$ exists as an element in $L_2(\mathbb{P})$ and equals to $L_2(\mathbb{P})$-$\lim_{t\uparrow T}\int_0^t\int_{\mathbb{R}} \Delta F(s, x)M(\mathrm{d}s, \mathrm{d}x)$. On the other hand, due to the martingale convergence theorem, $F(t, X_t) = \mathbb{E}_{\mathcal{F}_t}[f(X_T)] \to \mathbb{E}_{\mathcal{F}_{T-}}[f(X_T)]$ a.s. and in $L_2(\mathbb{P})$ as $t \uparrow T$, where $\mathcal{F}_{T-} := \sigma(\cup_{t<T}\mathcal{F}_t)$. Since $(\mathcal{F}_t)_{t\in[0,T]}$ is the augmented natural filtration of the Lévy process X, it holds that $\mathcal{F}_{T-} = \mathcal{F}_T$, and hence the desired conclusion follows. □

4. Closed form for the local risk-minimizing strategy

This section gives the proof of Theorem 1.1. First, let us fix the setting of this section.

Setting 4.1. Let $S = \mathrm{e}^X$ be the exponential of a Lévy process X with $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$. Assume that $\sigma^2 + \nu(\mathbb{R}) > 0$ and $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$.

The condition $\sigma^2 + \nu(\mathbb{R}) > 0$ is simply to exclude the trivial case that X is a.s. deterministic. The condition $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$ is equivalent to the square integrability of S (see [31, Theorem 25.3]).

By Itô's formula, one has

$$\begin{aligned} S &= 1 + \left(\int_0^\cdot \sigma S_{t-}\mathrm{d}W_t + \int_0^\cdot \int_{\mathbb{R}_0} S_{t-}(\mathrm{e}^x - 1)\widetilde{N}(\mathrm{d}t, \mathrm{d}x) \right) + \int_0^\cdot \gamma_S S_{t-}\mathrm{d}t \\ &=: 1 + S^{\mathrm{m}} + S^{\mathrm{fv}}, \end{aligned}$$

where S^{m} and S^{fv} respectively denote the martingale part and the predictable finite variation part in the representation of S, and where

$$\gamma_S := \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (\mathrm{e}^x - 1 - x\mathbb{1}_{\{|x|\leqslant 1\}})\nu(\mathrm{d}x). \tag{4.1}$$

Recall from Theorem 1.1 the notation

$$\|(\sigma, \nu)\| = \sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^x - 1)^2 \nu(\mathrm{d}x) \in (0, \infty).$$

4.1. **Föllmer–Schweizer (FS) decomposition.** We briefly present the FS decomposition of a random variable and the notion of the minimal local martingale measure which is the key tool to determine the FS decomposition. We refer the reader to [34] for a survey about these objects.

In this article, we follow [17, p.863] and use the family of *admissible strategies* as

$$\Sigma_S^{\mathrm{adm}}(\mathbb{P}) := \left\{ \vartheta \text{ predictable} : \mathbb{E}\int_0^T \vartheta_t^2 S_{t-}^2 \mathrm{d}t < \infty \right\}. \tag{4.2}$$

It turns out that if $\vartheta \in \Sigma_S^{\mathrm{adm}}(\mathbb{P})$, then

$$\begin{aligned} \mathbb{E}\int_0^T \vartheta_t^2 \mathrm{d}[S,S]_t &= \mathbb{E}\int_0^T \vartheta_t^2 \mathrm{d}[S^{\mathrm{m}}, S^{\mathrm{m}}]_t = \mathbb{E}\int_0^T \vartheta_t^2 \mathrm{d}\langle S^{\mathrm{m}}, S^{\mathrm{m}}\rangle_t \\ &= \|(\sigma,\nu)\|\mathbb{E}\int_0^T \vartheta_t^2 S_{t-}^2 \mathrm{d}t < \infty. \end{aligned} \tag{4.3}$$

The following definition is due to [34].

Definition 4.2. (1) An $H \in L_2(\mathbb{P})$ admits a *FS decomposition* if H can be written as

$$H = H_0 + \int_0^T \vartheta_t^H \mathrm{d}S_t + L_T^H,$$

where $H_0 \in \mathbb{R}$, $\vartheta^H \in \Sigma_S^{\mathrm{adm}}(\mathbb{P})$ and $L^H \in \mathcal{M}_2^0(\mathbb{P})$ is $\mathbb{P}$-orthogonal to S^{m}.

(2) The integrand ϑ^H is called the *local risk-minimizing* strategy of H.

Remark 4.3. In our context, S satisfies the *structure condition* and the *mean-variance trade-off process* $\widehat{K}$ of S in the sense of [34, p.553] is

$$\widehat{K}_t = \frac{\gamma_S^2}{\|(\sigma,\nu)\|} t,$$

which is uniformly bounded in $(\omega, t) \in \Omega \times [0, T]$. Hence, it is known that any $H \in L_2(\mathbb{P})$ admits a unique FS decomposition (see [26, Theorem 3.4]).

We continue with the notion of the minimal martingale measure.

Definition 4.4 ([33], Section 2)**.** Let $\mathcal{E}(U) \in \mathrm{CL}([0,T])$ be the stochastic exponential of U, i.e. $\mathrm{d}\mathcal{E}(U) = \mathcal{E}(U)_-\mathrm{d}U$ with $\mathcal{E}(U)_0 = 1$, where

$$U = -\frac{\gamma_S}{\|(\sigma,\nu)\|}\left(\sigma W + \int_0^{\cdot}\int_{\mathbb{R}_0}(\mathrm{e}^x - 1)\widetilde{N}(\mathrm{d}s, \mathrm{d}x)\right). \tag{4.4}$$

If $\mathcal{E}(U) > 0$, then the probability measure $\mathbb{P}^*$ defined by

$$\mathrm{d}\mathbb{P}^* := \mathcal{E}(U)_T \mathrm{d}\mathbb{P}$$

is called the *minimal martingale measure* for S.

Since U given in (4.4) is a Lévy process and belongs to $\mathcal{M}_2^0(\mathbb{P})$, it is known that $\mathcal{E}(U)$ is also an $L_2(\mathbb{P})$-martingale (see, e.g., [29, Ch.V, Theorem 67] or [11, Lemma 1]).

We now give a condition imposed on the characteristics of X such that $\mathbb{P}^*$ exists. Let $(U|\mathbb{P}) \sim (\gamma_U, \sigma_U, \nu_U)$ and denote

$$\alpha_U(x) := -\frac{\gamma_S(\mathrm{e}^x - 1)}{\|(\sigma,\nu)\|}, \quad x \in \mathbb{R}.$$

Then, it follows from (4.4) that

$$\gamma_U = -\int_{|\alpha_U(x)|>1}\alpha_U(x)\nu(\mathrm{d}x), \quad \sigma_U = \frac{|\gamma_S|\sigma}{\|(\sigma,\nu)\|}, \quad \nu_U = \nu\circ\alpha_U^{-1}. \tag{4.5}$$

Since

$$\mathcal{E}(U) > 0 \Leftrightarrow \Delta U > -1 \Leftrightarrow \nu_U((-\infty,-1]) = 0 \Leftarrow \gamma_S(\mathrm{e}^x - 1) < \|(\sigma,\nu)\|, \forall x \in \operatorname{supp}\nu,$$

the following assumption ensures the existence of $\mathbb{P}^*$:

Assumption 4.5. $\gamma_S(\mathrm{e}^x - 1) < \|(\sigma,\nu)\|$ for all $x \in \operatorname{supp}\nu$.

Remark that a sufficient condition for Assumption 4.5 is

$$0 \geqslant \gamma_S \geqslant -\|(\sigma,\nu)\|.$$

Assume that Assumption 4.5 holds true. Then, by an application of Girsanov's theorem (see, e.g., [9, Propositions 2 and 3]), X is also a Lévy process under $\mathbb{P}^*$ with $(X|\mathbb{P}^*) \sim (\gamma^*, \sigma^*, \nu^*)$, where

$$\begin{aligned}\gamma^* &= \gamma - \frac{\gamma_S}{\|(\sigma,\nu)\|}\left(\sigma + \int_{|x|\leqslant 1}x(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right),\\ \sigma^* &= \sigma \quad \text{and} \quad \nu^*(\mathrm{d}x) = \left(1 - \frac{\gamma_S(\mathrm{e}^x - 1)}{\|(\sigma,\nu)\|}\right)\nu(\mathrm{d}x).\end{aligned} \tag{4.6}$$

Moreover, if W^* and $\widetilde{N}^*$ are the standard Brownian motion and the compensated Poisson random measure of X under $\mathbb{P}^*$, then

$$W_t^* = W_t + \frac{\gamma_S\sigma}{\|(\sigma,\nu)\|}t, \tag{4.7}$$

$$\widetilde{N}^*(\mathrm{d}t, \mathrm{d}x) = \widetilde{N}(\mathrm{d}t, \mathrm{d}x) + \frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\mathrm{d}t. \tag{4.8}$$

In the sequel, let $\mathbb{E}^*$ (resp. $\mathbb{E}^*_{\mathcal{G}}$) denote the expectation (resp. conditional expectation given a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$) with respect to $\mathbb{P}^*$.

4.2. **Proof of Theorem 1.1.** Let $f(x) := g(\mathrm{e}^x)$ and $F^*(t,x) := \mathbb{E}^* f(x + X_{T-t})$ so that $G^*(t, \mathrm{e}^x) = F^*(t,x)$ for $(t,x) \in [0,T] \times \mathbb{R}$. We define

$$\Delta_J G^*(t,x) := G^*(t, \mathrm{e}^x S_{t-}) - G^*(t, S_{t-}), \quad (t,x) \in [0,T] \times \mathbb{R}.$$

(1) We present here a direct proof for this assertion, an alternative argument for more general settings can be found in [6, Proof of Theorem 4.3]. By assumption, $f(X_T) = g(S_T) \in L_2(\mathbb{P}^*)$ and $\mathbb{E}^*|f(x + X_t)| = \mathbb{E}^*|g(\mathrm{e}^x S_t)| < \infty$ for any $(t,x) \in [0,T] \times \mathbb{R}$, we apply Proposition 1.2 to obtain

$$K^* = \mathbb{E}^* g(S_T) + \int_0^\cdot \sigma S_{t-} \partial_y G^*(t, S_{t-}) \mathrm{d}W_t^* + \int_0^\cdot \int_{\mathbb{R}_0} \Delta_J G^*(t,x) \widetilde{N}^*(\mathrm{d}t, \mathrm{d}x), \tag{4.9}$$

where $K^* = (K_t^*)_{t \in [0,T]}$ is the càdlàg version of the $L_2(\mathbb{P}^*)$-martingale $(\mathbb{E}^*_{\mathcal{F}_t}[g(S_T)])_{t\in[0,T]}$, and where W^* and $\widetilde{N}^*$ are introduced in (4.7) and (4.8). Then, it holds that $\mathcal{E}(U)K^*$ is a martingale under $\mathbb{P}$. Since the $\mathbb{P}$-martingale U given in (4.4) satisfies that

$$\|\langle U, U \rangle_T\|_{L_\infty(\mathbb{P})} = \frac{\gamma_S^2 T}{\|(\sigma,\nu)\|^2} \left(\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^x - 1)^2 \nu(\mathrm{d}x) \right) < \infty,$$

it implies that $\mathcal{E}(U)$ is regular and satisfies (R_2) in the sense of [7, Proposition 3.7]. Since $K_T^* = g(S_T) \in L_2(\mathbb{P})$ by assumption, we apply [7, Theorem 4.9((i)⇔(ii))] to obtain

$$\mathbb{E}[K^*, K^*]_T < \infty.$$

Combining this with (4.9) yields

$$\mathbb{E} \int_0^T \sigma^2 |S_{t-} \partial_y G^*(t, S_{t-})|^2 \mathrm{d}t + \mathbb{E} \int_0^T \int_{\mathbb{R}_0} |\Delta_J G^*(t,x)|^2 N(\mathrm{d}t, \mathrm{d}x) = \mathbb{E}[K^*, K^*]_T < \infty.$$

Since $\mathrm{d}t\nu(\mathrm{d}x)$ is the predictable $\mathbb{P}$-compensator of $N(\mathrm{d}t, \mathrm{d}x)$, it implies that

$$\mathbb{E} \int_0^T \sigma^2 |S_{t-} \partial_y G^*(t, S_{t-})|^2 \mathrm{d}t + \mathbb{E} \int_0^T \int_{\mathbb{R}} |\Delta_J G^*(t,x)|^2 \nu(\mathrm{d}x) \mathrm{d}t < \infty. \tag{4.10}$$

Using Cauchy–Schwarz's inequality yields

$$\begin{aligned}
&\mathbb{E} \int_0^T \sigma^2 S_{t-}^2 |\partial_y G^*(t, S_{t-})| \mathrm{d}t + \mathbb{E} \int_0^T \int_{\mathbb{R}} |\Delta_J G^*(t,x) S_{t-} (\mathrm{e}^x - 1)| \nu(\mathrm{d}x) \mathrm{d}t \\
&\leqslant \sqrt{\mathbb{E} \int_0^T S_{t-}^2 \mathrm{d}t} \sqrt{\mathbb{E} \int_0^T |\sigma^2 S_{t-} \partial_y G^*(t, S_{t-})|^2 \mathrm{d}t} \\
&\quad + \sqrt{\int_{\mathbb{R}} (\mathrm{e}^x - 1)^2 \nu(\mathrm{d}x)} \sqrt{\mathbb{E} \int_0^T S_{t-}^2 \mathrm{d}t} \sqrt{\mathbb{E} \int_0^T \int_{\mathbb{R}} |\Delta_J G^*(t,x)|^2 \nu(\mathrm{d}x) \mathrm{d}t} \\
&< \infty.
\end{aligned} \tag{4.11}$$

On the other hand, the FS decomposition of $H = g(S_T)$ is

$$g(S_T) = H_0 + \int_0^T \vartheta_t^H \mathrm{d}S_t + L_T^H \tag{4.12}$$

where $H_0 \in \mathbb{R}$, $\vartheta^H \in \Sigma_S^{\mathrm{adm}}(\mathbb{P})$ and $L^H \in \mathcal{M}_2^0(\mathbb{P})$ is $\mathbb{P}$-orthogonal to the martingale component S^{m} of S. According to [34, Eq. (3.10)], it holds that L^H is a local $\mathbb{P}^*$-martingale. We remark that $\int_0^\cdot \vartheta_t^H \mathrm{d}S_t$ is also a local $\mathbb{P}^*$-martingale. Using Cauchy–Schwarz's inequality and (4.3), we obtain

$$\mathbb{E}^* \sqrt{[L^H, L^H]_T} \leqslant \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})} \sqrt{\mathbb{E}[L^H, L^H]_T} < \infty,$$

$$\mathbb{E}^*\sqrt{\int_0^T |\vartheta_t^H|^2 \mathrm{d}[S,S]_t} \leqslant \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{\mathbb{E}\int_0^T |\vartheta_t^H|^2 \mathrm{d}[S,S]_t} < \infty.$$

Hence, the Burkholder–Davis–Gundy inequality verifies that both L^H and $\int_0^\cdot \vartheta_t^H \mathrm{d}S_t$ are $\mathbb{P}^*$-martingales. Combining (4.9) with (4.12), we derive $H_0 = \mathbb{E}^* g(S_T)$ and

$$\int_0^\cdot \vartheta_t^H \mathrm{d}S_t + L^H = \int_0^\cdot \sigma S_{t-}\partial_y G^*(t, S_{t-})\mathrm{d}W_t^* + \int_0^\cdot\int_{\mathbb{R}_0} \Delta_J G^*(t,x)\widetilde{N}^*(\mathrm{d}t, \mathrm{d}x). \quad (4.13)$$

Recall that the martingale part of S is $S^{\mathrm{m}} = \int_0^\cdot \sigma S_{t-}\mathrm{d}W_t + \int_0^\cdot \int_{\mathbb{R}_0} S_{t-}(\mathrm{e}^x - 1)\widetilde{N}(\mathrm{d}t, \mathrm{d}x)$. Since $\langle L^H, S^{\mathrm{m}}\rangle^{\mathbb{P}} = 0$ by the definition of the FS decomposition, we take the predictable quadratic covariation on both sides of (4.13) with S^{m} under $\mathbb{P}$ and notice that the integrability condition (4.11) holds to obtain

$$\|(\sigma,\nu)\|\int_0^\cdot \vartheta_t^H S_{t-}^2 \mathrm{d}t = \int_0^\cdot \sigma^2 S_{t-}^2 \partial_y G^*(t, S_{t-})\mathrm{d}t + \int_0^\cdot\int_{\mathbb{R}} \Delta_J G^*(t,x) S_{t-}(\mathrm{e}^x-1)\nu(\mathrm{d}x)\mathrm{d}t,$$

which yields (1.2) as desired.

(2) It follows from Cauchy–Schwarz's inequality and (4.10) that

$$\begin{aligned}
&\mathbb{E}^*\int_0^T |\sigma^2 S_{t-}\partial_y G^*(t,S_{t-})|\mathrm{d}t + \mathbb{E}^*\int_0^T\int_{\mathbb{R}} |\Delta_J G^*(t,x)(\mathrm{e}^x-1)|\nu(\mathrm{d}x)\mathrm{d}t \\
&\leqslant \sqrt{T}\|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{\mathbb{E}\int_0^T |\sigma^2 S_{t-}\partial_y G^*(t,S_{t-})|^2 \mathrm{d}t} \\
&\qquad + \|\mathcal{E}(U)_T\|_{L_2(\mathbb{P})}\sqrt{T\int_{\mathbb{R}}(\mathrm{e}^x-1)^2\nu(\mathrm{d}x)}\sqrt{\mathbb{E}\int_0^T\int_{\mathbb{R}} |\Delta_J G^*(t,x)|^2\nu(\mathrm{d}x)\mathrm{d}t} \\
&< \infty.
\end{aligned} \quad (4.14)$$

By assumption, it is clear that $(G^*(t,\mathrm{e}^x S_t) - G^*(t,S_t))_{t\in[0,T]}$ is a $\mathbb{P}^*$-martingale for each $x\in\mathbb{R}$. In the case $\sigma>0$, due to $g(S_T)\in L_2(\mathbb{P}^*)$ and Lemma 3.1, $(S_t\partial_y G^*(t,S_t))_{t\in[0,T)}$ is also a $\mathbb{P}^*$-martingale. Hence, the function

$$[0,T)\ni t\mapsto \mathbb{E}^*|\sigma^2 S_t\partial_y G^*(t,S_t)| + \mathbb{E}^*\int_{\mathbb{R}}|G^*(t,\mathrm{e}^x S_t) - G^*(t,S_t)||\mathrm{e}^x-1|\nu(\mathrm{d}x)$$

is non-decreasing by the martingale property. In addition, noticing that $S_{t-} = S_t$ a.s. for each $t\in[0,T]$, we infer from (4.14) and Fubini's theorem that

$$\mathbb{E}^*|\sigma^2 S_t\partial_y G^*(t,S_t)| + \mathbb{E}^*\int_{\mathbb{R}}|G^*(t,\mathrm{e}^x S_t) - G^*(t,S_t)||\mathrm{e}^x-1|\nu(\mathrm{d}x) < \infty$$

for all $t\in[0,T)$. Therefore,

$$\left(\frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 S_t\partial_y G^*(t,S_t) + \int_{\mathbb{R}}(G^*(t,\mathrm{e}^x S_t) - G^*(t,S_t))(\mathrm{e}^x-1)\nu(\mathrm{d}x)\right)\right)_{t\in[0,T)}$$

is a $\mathbb{P}^*$-martingale for which one can find a càdlàg modification, denoted by φ^g. Then, the process $\tilde{\vartheta}^g$ defined by

$$\tilde{\vartheta}^g := \frac{\varphi^g}{S} \quad (4.15)$$

satisfies the desired requirements. □

Remark 4.6. Let $\tilde{\vartheta} \in \mathrm{CL}([0,T))$ be such that $\tilde{\vartheta} = \tilde{\vartheta}^g$ for $\mathbb{P} \otimes \lambda$-a.e. $(\omega, t) \in \Omega \times [0,T)$, where $\tilde{\vartheta}^g$ given in (4.15). Then, $\mathbb{P}(\tilde{\vartheta}_t = \tilde{\vartheta}^g_t,\ \forall t \in [0,T)) = 1$ due to the càdlàg property. Hence, $\tilde{\vartheta}_-$ is also a LRM strategy of $H = g(S_T)$, and it holds that, for any $t \in [0,T)$,

$$\tilde{\vartheta}_t = \frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 S_t \partial_y G^*(t, S_t) + \int_{\mathbb{R}} (G^*(t, \mathrm{e}^x S_t) - G^*(t, S_t))(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right) \quad \text{a.s.}$$

5. Discrete-time hedging in weighted bounded mean oscillation spaces

This section is a continuation of the work in [37] for the exponential Lévy models. First, we use the approximation scheme for stochastic integrals introduced in [37] and investigate the resulting error in weighted BMO spaces. Consequently, the L_p-estimates ($p \in (2,\infty)$) for the error are provided. Secondly, to illustrate the obtained results, we consider the stochastic integral term in the FS decomposition of a European type option. This integral can be interpreted as the hedgeable part of the option. Notice that we do not assume the (local) martingale property under the reference measure for the underlying price process.

5.1. Weighted bounded mean oscillation (BMO) spaces.

Let $\mathcal{S}([0,T])$ denote the family of all stopping times $\rho\colon \Omega \to [0,T]$, and set $\inf \emptyset := \infty$.

Definition 5.1 ([14, 15])**.** Let $p \in (0,\infty)$.

(1) For $\Phi \in \mathrm{CL}^+([0,T])$, we denote by $\mathrm{BMO}_p^\Phi(\mathbb{P})$ the space of all $Y \in \mathrm{CL}_0([0,T])$ with $\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} < \infty$, where

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

(2) (*Weight regularity*) Let $\mathcal{SM}_p(\mathbb{P})$ be the space of all $\Phi \in \mathrm{CL}^+([0,T])$ with $\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} < \infty$, where

$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} := \inf\left\{c \geqslant 0 : \mathbb{E}_{\mathcal{F}_a}\left[\sup\nolimits_{t\in[a,T]} \Phi_t^p\right] \leqslant c^p \Phi_a^p \text{ a.s.}, \forall a \in [0,T]\right\}.$$

The theory of non-weighted BMO-martingales (i.e. when $\Phi \equiv 1$ and Y is a martingale) can be found in [29, Ch.IV]. One remarks that the weighted BMO spaces above were introduced in [14] for general càdlàg processes which are not necessarily martingales.

Definition 5.2 ([14])**.** For $s \in (1,\infty)$, we denote by $\mathcal{RH}_s(\mathbb{P})$ the family of all probability measures $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} =: U \in L_s(\mathbb{P})$ and there exists a constant $c_{(5.1)} > 0$ such that U satisfies the following *reverse Hölder inequality*

$$\mathbb{E}_{\mathcal{F}_\rho}[U^s] \leqslant c_{(5.1)}^s |\mathbb{E}_{\mathcal{F}_\rho}[U]\,|^s \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T]). \tag{5.1}$$

We refer the reader to [14, 15] for further properties of those quantities. Proposition 5.3 below recalls some features of weighted BMO which are crucial for our applications, and their proofs can be found in [15, Proposition A.6] and [37, Proposition 2.5].

Proposition 5.3 ([15, 37])**.** *Let $p \in (0,\infty)$.*

(1) *There is a constant $c_1 = c(p) > 0$ such that $\|\cdot\|_{L_p(\mathbb{P})} \leqslant c_1 \|\Phi\|_{L_p(\mathbb{P})} \|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$.*

(2) *If $\Phi \in \mathcal{SM}_p(\mathbb{P})$, then for any $r \in (0,p]$ there is a constant $c_2 = c_2(r,p,\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}) > 0$ 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 exists a constant $c_3 = c(s,p) > 0$ such that $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{Q})} \leqslant c_3 \|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$.*

Remark 5.4. The benefit of Proposition 5.3(2) is as follows: If $p \in [2, \infty)$ (this is usually the case in applications), then one can choose $r = 2$ so that $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} \sim_{c_2} \|\cdot\|_{\mathrm{BMO}_2^\Phi(\mathbb{P})}$, and then we can still exploit some similar techniques as in the $L_2(\mathbb{P})$-theory to deal with $\|\cdot\|_{\mathrm{BMO}_2^\Phi(\mathbb{P})}$. Combining this observation with Proposition 5.3(1) yields the following estimate provided that $\Phi \in \mathcal{SM}_p(\mathbb{P})$, $p \in [2, \infty)$,

$$\|\cdot\|_{L_p(\mathbb{P})} \leqslant c_1 c_2 \|\Phi\|_{L_p(\mathbb{P})} \|\cdot\|_{\mathrm{BMO}_2^\Phi(\mathbb{P})}. \tag{5.2}$$

Proposition 5.3(3) gives a change of the underlying measure which might be of interest for further applications in mathematical finance.

5.2. **Jump adjusted approximation.** Let us recall from [37] the approximation scheme with the jump adjusted method. Roughly speaking, this method is constructed by adding suitable correction terms to the classical Riemann sum of the stochastic integral as soon as relatively large jumps of the driving process occur.

Time-nets. Let $\mathcal{T}_{\mathrm{det}}$ denote the family of all *deterministic* time-nets $\tau = (t_i)_{i=0}^n$, $0 = t_0 < t_1 < \cdots < t_n = T$, $n \geqslant 1$. The mesh size of $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$ associated with a parameter $\theta \in (0, 1]$ is defined by

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

Let $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n + 1$. By a short calculation we can find that $\|\tau_n\|_\theta \geqslant \frac{T^\theta}{n}$. Minimizing $\|\tau_n\|_\theta$ over $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau = n + 1$ leads us to the following adapted time-nets, which were exploited in [11, 13, 14, 15, 37]: For $\theta \in (0, 1]$ and $n \geqslant 1$, the *adapted time-net* $\tau_n^\theta = (t_{i,n}^\theta)_{i=0}^n$ is defined by

$$t_{i,n}^\theta := T\left(1 - \sqrt[\theta]{1 - i/n}\right), \quad i = 1, \ldots, n. \tag{5.3}$$

Then, a calculation gives

$$\frac{T^\theta}{n} \leqslant \|\tau_n^\theta\|_\theta \leqslant \frac{T^\theta}{\theta n}, \quad n \geqslant 1.$$

Jump adjusted approximation scheme. Let $S = \mathrm{e}^X$ be the exponential Lévy process and assume Setting 4.1. Let $\tilde{\vartheta} \in \mathrm{CL}([0, T))$ be such that $\mathbb{E}\int_0^T \tilde{\vartheta}_{t-}^2 S_{t-}^2 \mathrm{d}t < \infty$ (the tilde sign here indicates the càdlàg property of the process $(\tilde{\vartheta}_t)_{t \in [0,T)}$). For $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$, the Riemann approximation $A^{\mathrm{Rm}}(\tilde{\vartheta}, \tau)$ of $\int_0^T \tilde{\vartheta}_{t-} \mathrm{d}S_t$ is defined by

$$A_t^{\mathrm{Rm}}(\tilde{\vartheta}, \tau) := \sum_{i=1}^n \tilde{\vartheta}_{t_{i-1}-}(S_{t_i \wedge t} - S_{t_{i-1} \wedge t}), \quad t \in [0, T].$$

Before proceeding to the jump adjusted approximation, we need the following stopping times which capture the relative large jumps of S: For $\varepsilon > 0$ and $\kappa \geqslant 0$, we define the family of stopping times $\rho(\varepsilon, \kappa) = (\rho_i(\varepsilon, \kappa))_{i \geqslant 0}$ by setting $\rho_0(\varepsilon, \kappa) := 0$ and

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

By specializing $\kappa = 0$, the parameter ε can be regarded as the *jump size threshold*. When $\kappa > 0$, this threshold shrinks as $t \uparrow T$, and thus the parameter κ indices the *jump size decay rate*. The reason for using the decay function $\varepsilon(T - t)^\kappa$ is to compensate the growth of integrands.

Definition 5.5. Let $\varepsilon > 0$, $\kappa \in [0, \frac{1}{2})$ and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$.

(1) Let $\tau \sqcup \rho(\varepsilon,\kappa)$ denote the *combined time-net* constructed by combining τ with $\rho(\varepsilon,\kappa)$ and re-ordering their time-knots.
(2) For $t \in [0,T]$, we define

$$\begin{aligned}
\tilde{\vartheta}(\tau)_t &:= \sum_{i=1}^{n} \tilde{\vartheta}_{t_{i-1}-} \mathbb{1}_{(t_{i-1},t_i]}(t),\\
A_t^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa) &:= A_t^{\mathrm{Rm}}(\tilde{\vartheta},\tau) + \sum_{\rho_i(\varepsilon,\kappa)\in[0,t]\cap[0,T)} \left(\tilde{\vartheta}_{\rho_i(\varepsilon,\kappa)-} - \tilde{\vartheta}(\tau)_{\rho_i(\varepsilon,\kappa)}\right) \Delta S_{\rho_i(\varepsilon,\kappa)},\\
E_t^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa) &:= \int_0^t \tilde{\vartheta}_{u-}\mathrm{d}S_u - A_t^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa).
\end{aligned}$$

Denote

$$\mathcal{N}(\varepsilon,\kappa) := \inf\{i \geqslant 1 : \rho_i(\varepsilon,\kappa) = T\}.$$

We apply [37, Proposition 5.3] (with $\alpha = 2$) to conclude that $\mathcal{N}(\varepsilon,\kappa) < \infty$ a.s. for any $\varepsilon > 0$ and $\kappa \in [0,\frac{1}{2})$. Hence, the sum in the definition of $A^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa)$ is a finite sum a.s. By adjusting this sum on a set of probability zero, we may assume that $A^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa)$, and hence, $E^{\mathrm{adj}}(\tilde{\vartheta},\tau|\varepsilon,\kappa)$, belong to $\mathrm{CL}_0([0,T])$.

5.3. **Discrete-time approximation in weighted** BMO**: A general result.** Let us introduce the main assumption to obtain the approximation results.

Assumption 5.6. Let $S = \mathrm{e}^X$ with $(X|\mathbb{P}) \sim (\gamma,\sigma,\nu)$. Let $\tilde{\vartheta} \in \mathrm{CL}([0,T))$ and $\theta \in (0,1]$. Assume that

(i) $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$.
(ii) $\Delta\tilde{\vartheta}_t = 0$ a.s. for each $t \in [0,T)$.
(iii) There exists a random measure $\Upsilon\colon \Omega \times \mathcal{B}((0,T)) \to [0,\infty]$ such that $\Upsilon(\omega,(0,t]) < \infty$ for all $(\omega,t) \in \Omega \times (0,T)$, and such that for any $0 \leqslant a < b < T$,

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,b]} |\tilde{\vartheta}_t - \tilde{\vartheta}_a|^2 S_t^2 \mathrm{d}t\right] \leqslant c_{(5.4)}^2 \mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,b]} (b-t)\Upsilon(\cdot,\mathrm{d}t)\right] \quad \text{a.s.} \tag{5.4}$$

(iv) There is an a.s. non-decreasing process $\Theta \in \mathrm{CL}^+([0,T])$ such that
   (1) (*Growth condition*) One has

$$|\tilde{\vartheta}_a| \leqslant c_{(5.5)}(T-a)^{\frac{\theta-1}{2}}\Theta_a \quad \text{a.s., } \forall a \in [0,T). \tag{5.5}$$

   (2) (*Curvature condition*) One has

$$\mathbb{E}_{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta}\Upsilon(\cdot,\mathrm{d}t)\right] \leqslant c_{(5.6)}^2 \Phi_a^2 \quad \text{a.s., } \forall a \in [0,T), \tag{5.6}$$

   where

$$\Phi := \Theta S. \tag{5.7}$$

Here, $c_{(5.4)}$, $c_{(5.5)}$, $c_{(5.6)}$ are positive constants independent of a, b.

Condition (i) is equivalent to the square integrability of S. Condition (ii) means that the integrand $\tilde{\vartheta}$ has no fixed-time discontinuity, and this property is satisfied in various contexts. Conditions (iii)–(iv) are adapted from [37, Assumption 3.3], and the random measure Υ above describes some kind of curvature of the stochastic integral. Several specifications of Υ are provided in [15] (for the Brownian setting and the Lévy setting) and in [37] (for the exponential Lévy setting).

Theorem 5.7. *Let Assumption 5.6 hold for some $\tilde{\vartheta} \in \mathrm{CL}([0,T))$ and for some $\theta \in (0,1]$. For Φ given in (5.7), we define $\overline{\Phi} \in \mathrm{CL}^+([0,T])$ by setting*

$$\overline{\Phi}_t := \Phi_t + \sup_{u\in[0,t]} |\Delta\Phi_u|, \quad t \in [0,T].$$

Assume that $\Phi \in \mathcal{SM}_2(\mathbb{P})$. Then, the following assertions hold:

(1) *If $\int_{|x|\leqslant 1} |x|^r \nu(\mathrm{d}x) < \infty$ for some $r \in [1,2]$, then there is a constant $c_{(5.8)} > 0$ such that for all $\tau \in \mathcal{T}_{\det}$, $\varepsilon > 0$,*

$$\left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(5.8)} \max\left\{\varepsilon^{1-r}\sqrt{\|\tau\|_\theta}, \sqrt{\|\tau\|_\theta}, \varepsilon\right\}. \tag{5.8}$$

Consequently, choosing the adapted time-net τ_n^θ and $\varepsilon = n^{-\frac{1}{2r}}$ in (5.8) we obtain

$$\sup_{n\geqslant 1} n^{\frac{1}{2r}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}, \tau_n^\theta \middle| n^{-\frac{1}{2r}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty. \tag{5.9}$$

(2) *If $\sup_{r>0} \left| \int_{|\mathrm{e}^x-1|>r} (\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right| < \infty$, then there is a constant $c_{(5.10)} > 0$ such that for all $\tau \in \mathcal{T}_{\det}$, $\varepsilon > 0$,*

$$\left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}, \tau \middle| \varepsilon, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leqslant c_{(5.10)} \max\left\{\sqrt{\|\tau\|_\theta}, \varepsilon\right\}. \tag{5.10}$$

Consequently, choosing the adapted time-net τ_n^θ and $\varepsilon = n^{-\frac{1}{2}}$ in (5.10) we obtain

$$\sup_{n\geqslant 1} n^{\frac{1}{2}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}, \tau_n^\theta \middle| n^{-\frac{1}{2}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty. \tag{5.11}$$

(3) *If in addition $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in (2,\infty)$, then the conclusions of items (1), (2) hold for the $L_p(\mathbb{P})$-norm in place of the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$-norm.*

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

Proof. By Subsection 2.4, one has $\mathrm{d}S_t = S_{t-}\mathrm{d}Z_t$, where Z is a square integrable Lévy process with the Lévy measure $\nu_Z = \nu \circ h^{-1}$, where $h(x) := \mathrm{e}^x - 1$. Moreover, it is clear that $\int_{|x|\leqslant 1} |x|^r \nu(\mathrm{d}x) < \infty \Leftrightarrow \int_{|z|\leqslant 1} |z|^r \nu_Z(\mathrm{d}z) < \infty$. Then, we apply [37, Theorem 3.10] to obtain items (1) and (2). Items (3), (4) are due to Proposition 5.3 and Lemma A.2. □

Remark 5.8. The parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$-norm in (5.9) and (5.11) can be regarded as the $L_2(\mathbb{P})$-norm of the cardinality of the combined time-net $\tau_n^\theta \sqcup \rho(n^{-\frac{1}{2r}}, \frac{1-\theta}{2})$ and $\tau_n^\theta \sqcup \rho(n^{-\frac{1}{2}}, \frac{1-\theta}{2})$ respectively. This assertion is derived from [37, Proposition 3.13] (with $\mathbb{Q} = \mathbb{P}$, and $q = 2$, $r = \infty$).

5.4. **Hölder spaces and α-stable-like processes.** We first define some classes of Hölder continuous functions and bounded Borel functions, where the payoff functions are contained in.

Definition 5.9. Let $U \subseteq \mathbb{R}$ be a non-empty open interval.

(1) For $\eta \in [0,1]$, we let $C^{0,\eta}(U)$ denote the space of all Borel functions $f\colon U \to \mathbb{R}$ with $|f|_{C^{0,\eta}(U)} < \infty$, where

$$|f|_{C^{0,\eta}(U)} := \inf\{c \geqslant 0 : |f(x) - f(y)| \leqslant c|x-y|^\eta,\ \forall x, y \in U, x \neq y\}.$$

(2) For $q \in [1,\infty]$, we define

$$\mathring{W}^{1,q}(U) := \left\{ f\colon U \to \mathbb{R} : \exists k \in L_q(U),\ f(y) - f(x) = \int_x^y k(u)\mathrm{d}u, \forall x, y \in U, x < y\right\},$$

and let $|f|_{\mathring{W}^{1,q}(U)} := \|k\|_{L_q(U)}$.

It is obvious that $C^{0,\eta}(U)$ is the space of all η-Hölder continuous functions on U for $\eta \in (0,1]$, and $C^{0,0}(U)$ consists of all bounded and Borel functions on U. For $\eta \in [0,1]$, Hölder's inequality implies that

$$\mathring{W}^{1,\frac{1}{1-\eta}}(U) \subseteq C^{0,\eta}(U) \quad \text{with} \quad |f|_{C^{0,\eta}(U)} \leqslant |f|_{\mathring{W}^{1,\frac{1}{1-\eta}}(U)}, \quad \forall f \in \mathring{W}^{1,\frac{1}{1-\eta}}(U).$$

In particular, $\mathring{W}^{1,\infty}(U) = C^{0,1}(U)$, which is the space of Lipschitz functions on U.

We next introduce some classes of α-stable-like Lévy measures.

Definition 5.10. Let ν be a Lévy measure and $\alpha \in (0,2)$.

(1) We let $\nu \in \mathcal{S}_1(\alpha)$ if one can decompose $\nu = \nu_1 + \nu_2$, where ν_1, ν_2 are Lévy measures and satisfy that

$$\limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu_2(\mathrm{d}x) < \infty, \tag{5.12}$$

$$\nu_1(\mathrm{d}x) = \frac{k(x)}{|x|^{\alpha+1}} \mathbb{1}_{\{x\neq 0\}}\mathrm{d}x, \tag{5.13}$$

where $0 < \liminf_{x\to 0} k(x) \leqslant \limsup_{x\to 0} k(x) < \infty$, and the function $x \mapsto \frac{k(x)}{|x|^\alpha}$ is non-decreasing on $(-\infty, 0)$ and non-increasing on $(0,\infty)$.

(2) We let $\nu \in \mathcal{S}_2(\alpha)$ if

$$0 < \liminf_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu(\mathrm{d}x) \leqslant \limsup_{|u|\to\infty} \frac{1}{|u|^\alpha} \int_{\mathbb{R}} (1-\cos(ux))\nu(\mathrm{d}x) < \infty. \tag{5.14}$$

In fact, $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$ for $\alpha \in (0,2)$, and moreover, the inclusion is strict. This assertion and some further properties of $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ are given in Lemma A.1.

Example 5.11. Let us provide some examples for those classes of Hölder functions and of α-stable-like processes used in financial modelling.

(1) The European call and put are Lipschitz, hence they belong to $\mathring{W}^{1,\infty}(\mathbb{R}_+)$.
The *power call* $g(y) := ((y-K)\vee 0)^\eta$ with $K > 0$ and $\eta \in (0,1)$ belongs to $C^{0,\eta}(\mathbb{R}_+)$, but $g \notin \mathring{W}^{1,q}(\mathbb{R}_+)$ for any $q \in (1,\infty)$. However, we can decompose $g = g_1 + g_2$, where $g_1 := ((y-K)\vee 0)^\eta \wedge 1$ and $g_2 := g - g_1$, so that $g_1 \in \cap_{1\leqslant q<\frac{1}{1-\eta}} \mathring{W}^{1,q}(\mathbb{R}_+)$ and g_2 is Lipschitz. By the linearity, the LRM strategy of g is the sum of the LRM strategies corresponding to g_1 and g_2.
The binary option $g(y) := \mathbb{1}_{[K,\infty)}(y)$ belongs to $C^{0,0}(\mathbb{R}_+)$ obviously.

(2) The CGMY process with parameters $C, G, M > 0$ and $Y \in (0,2)$ (see [32, Section 5.3.9]) 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 $\mathcal{S}_1(Y)$ due to Lemma A.1(3).
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

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

Hence, Lemma A.1(3) verifies that $\nu_{\mathrm{NIG}} \in \mathcal{S}_1(1)$.

5.5. **Discretisation of LRM strategies.** Let X be a Lévy process with $(X|\mathbb{P}) \sim (\gamma, \sigma, \nu)$ and $S = \mathrm{e}^X$. In this subsection, we apply results of Subsection 5.3, and the stochastic integral being approximated is the integral term in the FS decomposition of $g(S_T)$. Moreover, we choose the càdlàg version $\tilde{\vartheta}^g$ of the LRM strategy as mentioned in Theorem 1.1(2) so that the integral we are going to approximate is of the form

$$\int_0^T \tilde{\vartheta}^g_{t-} \mathrm{d}S_t.$$

Under the assumptions of Theorem 1.1, it follows from Remark 4.6 that, for $t \in [0, T)$,

$$\tilde{\vartheta}^g_t = \frac{1}{\|(\sigma,\nu)\|}\left(\sigma^2 \partial_y G^*(t, S_t) + \int_{\mathbb{R}} \frac{G^*(t, \mathrm{e}^x S_t) - G^*(t, S_t)}{S_t}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right) \quad \text{a.s.} \tag{5.15}$$

For $\eta \in [0,1]$ and $t \in [0,T]$, we define

$$\Theta(\eta)_t := \sup\nolimits_{u\in[0,t]}(S_u^{\eta-1}), \quad \Phi(\eta)_t := \Theta(\eta)_t S_t,$$
$$\overline{\Phi}(\eta)_t := \Phi(\eta)_t + \sup\nolimits_{u\in[0,t]} |\Delta\Phi(\eta)_u|.$$

The results about approximation are given in items (4)–(6) of Theorem 5.12 below. In fact, the LRM strategy $\tilde{\vartheta}^g_-$ is quite difficult to investigate directly under the original measure $\mathbb{P}$ but it fits well the main assumption Assumption 5.6 under the minimal martingale measure $\mathbb{P}^*$. Therefore, our idea is to switch between the original measure $\mathbb{P}$ and the minimal martingale measure $\mathbb{P}^*$ and use the fact that weighted BMO-norms allow a change of measure as given in Proposition 5.3(3). Moreover, regarding the drift coefficient γ_S given in (4.1), we now focus on the case $\gamma_S \neq 0$ since the case $\gamma_S = 0$, which corresponds to the martingale setting, was investigated in [37, Section 4].

Theorem 5.12. *Assume Setting 4.1, Assumption 4.5, $\gamma_S \neq 0$ and $\int_{|x|>1} \mathrm{e}^{3x}\nu(\mathrm{d}x) < \infty$. Let $g \in C^{0,\eta}(\mathbb{R}_+)$ with $\eta \in [0,1]$. Then, the following assertions hold:*

(1) *Both $\Phi(\eta)$ and $\overline{\Phi}(\eta)$ belong to $\mathcal{SM}_3(\mathbb{P}) \cap \mathcal{SM}_2(\mathbb{P}^*)$.*

(2) *$\mathbb{P}^* \in \mathcal{RH}_3(\mathbb{P})$ and $\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} \leqslant c\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})}$ for some constant $c > 0$.*

(3) *Set $M := \tilde{\vartheta}^g S$. Then, Assumption 5.6 is satisfied under $\mathbb{P}^*$ for the selection*

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

and for the parameter θ provided in Table 1.

(4) *With the adapted time-nets τ_n^θ given in (5.3), one has*

$$\sup_{n\geqslant 1} n^{\frac{1}{2r}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}^g, \tau_n^\theta \,\middle|\, n^{-\frac{1}{2r}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} < \infty, \tag{5.16}$$

where the parameters r and θ are provided in Table 1.

(5) *Let $s \in (1, \infty)$, and assume in addition when $\frac{\|(\sigma,\nu)\|}{\gamma_S} \in [-1, \infty)$ that $\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu(\mathrm{d}x) < \infty$. Then, $\mathbb{P} \in \mathcal{RH}_s(\mathbb{P}^*)$ and*

$$\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} \sim_c \|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})}$$

for some constant $c \geqslant 1$, and hence

$$\sup_{n\geqslant 1} n^{\frac{1}{2r}} \left\| E^{\mathrm{adj}}\left(\tilde{\vartheta}^g, \tau_n^\theta \,\middle|\, n^{-\frac{1}{2r}}, \tfrac{1-\theta}{2}\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})} < \infty, \tag{5.17}$$

where the parameters r and θ are provided in Table 1. Moreover, (5.17) holds true for the $L_3(\mathbb{P})$-norm in place of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm.

(6) *If in addition* $\int_{|x|>1} \mathrm{e}^{px}\nu(\mathrm{d}x) < \infty$ *for some* $p \in (3, \infty)$*, then* (5.16) (*resp.* (5.17)) *is satisfied for the* $L_{p-1}(\mathbb{P}^*)$*-norm* (*resp.* $L_p(\mathbb{P})$*-norm*) *in place of the* $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)$*-norm* (*resp.* $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$*-norm*)*.*

Table 1: Values of parameters r and θ

	σ and η	Small jump condition	Regularity of g	Conclusions for r and θ
C1	$\sigma > 0$ $\eta \in (0, 1]$	$\int_{\lvert x\rvert\leqslant 1} \lvert x\rvert^\alpha \nu(\mathrm{d}x) < \infty$ for some $\alpha \in [1, 2]$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\forall r \in [\alpha, 2]$ $\forall \theta \in (0, \eta)$ if $\eta \in (0, 1)$ $\theta = 1$ if $\eta = 1$
C2	$\sigma = 0$ $\eta \in [0, 1]$	$\int_{\lvert x\rvert\leqslant 1} \lvert x\rvert^\alpha \nu(\mathrm{d}x) < \infty$ for some $\alpha \in [1, \eta + 1]$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\forall r \in [\alpha, 2]$ $\theta = 1$
C3	$\sigma = 0$ $\eta \in [0, 1)$	$\nu \in \mathcal{S}_1(\alpha)$ for some $\alpha \in [1 + \eta, 2)$	$g \in C^{0,\eta}(\mathbb{R}_+)$	$\forall r \in (\alpha, 2]$ $\forall \theta \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$
C4	$\sigma = 0$ $\eta \in [0, 1)$	$\nu \in \mathcal{S}_2(\alpha)$ for some $\alpha \in [1 + \eta, 2)$	$g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$	$\forall r \in (\alpha, 2]$ $\forall \theta \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$

Remark 5.13. (1) Let us comment on the parameters r and θ in Table 1. First, since we use the adapted time-net τ_n^θ which leads to better estimates (see (5.9)), it follows that the parameter r only depends on the behavior of ν around zero. Moreover, the smaller r is, the better approximation accuracy one achieves. The parameter θ is the outcome of the interplay between the behavior of ν around zero and the Hölder regularity of the payoff function.

(2) Since X is a Lévy process under both measures $\mathbb{P}$ and $\mathbb{P}^*$, we apply [37, Proposition 5.3] (with $\alpha = 2$ and $\kappa = \frac{1-\theta}{2}$, $\varepsilon = n^{-\frac{1}{2r}}$) to conclude that the parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)$-norm in (5.16) can be regarded as the $L_2(\mathbb{P})$-norm and the $L_2(\mathbb{P}^*)$-norm of the cardinality of the combine time-net $\tau_n^\theta \sqcup \rho(n^{-\frac{1}{2r}}, \frac{1-\theta}{2})$. The parameter n in front of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm in (5.17) can be interpreted in a similar manner.

For the proof of Theorem 5.12, we need the following lemmas where we recall $\nu^*(\mathrm{d}x) = \big(1 - \frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x - 1)\big)\nu(\mathrm{d}x)$ from (4.6) and the classes $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ from Definition 5.10.

Lemma 5.14. *Under Assumption 4.5, the following assertions hold:*

(1) *For* $\beta \in [0, 2]$*, one has* $\int_{|x|\leqslant 1} |x|^\beta \nu(\mathrm{d}x) < \infty \Leftrightarrow \int_{|x|\leqslant 1} |x|^\beta \nu^*(\mathrm{d}x) < \infty$*.*

(2) *Assume* $\gamma_S \neq 0$*. Then, for* $r \in [1, \infty)$ *one has*

$$\begin{aligned}\mathbb{E}\mathrm{e}^{rX_t} < \infty, \forall t > 0 &\Leftrightarrow \int_{|x|>1} \mathrm{e}^{rx}\nu(\mathrm{d}x) < \infty \\ &\Leftrightarrow \int_{|x|>1} \mathrm{e}^{(r-1)x}\nu^*(\mathrm{d}x) < \infty \Leftrightarrow \mathbb{E}^*\mathrm{e}^{(r-1)X_t} < \infty, \forall t > 0.\end{aligned}$$

Proof. Item (1) is clear from the relation between ν and ν^*. A short computation and [31, Theorem 25.3] imply item (2). □

Lemma 5.15. *Assume Assumption 4.5 and* $\int_{|x|>1} \mathrm{e}^x\nu(\mathrm{d}x) < \infty$*. If* $\nu \in \mathcal{S}_i(\alpha)$ *for some* $\alpha \in (0, 2)$*, then* $\nu^* \in \mathcal{S}_i(\alpha)$ *for* $i = 1, 2$*.*

Proof. We first prove the assertion for $i=1$. Assume that $\mathcal{S}_1(\alpha)\ni\nu=\nu_1+\nu_2$, where ν_1, ν_2 are Lévy measures satisfying (5.13) and (5.12) respectively. Observe that $\operatorname{supp}\nu_i\subseteq\operatorname{supp}\nu$ for $i=1,2$. We define

$$\nu_1^*(\mathrm{d}x):=\begin{cases}\Big(\Big(1-\frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x-1)\Big)\mathbb{1}_{\{x<0\}}+\mathbb{1}_{\{x>0\}}\Big)\nu_1(\mathrm{d}x) & \text{if } \frac{\gamma_S}{\|(\sigma,\nu)\|}\leqslant 0\\ \Big(\Big(1-\frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x-1)\Big)\mathbb{1}_{\{x>0\}}+\mathbb{1}_{\{x<0\}}\Big)\nu_1(\mathrm{d}x) & \text{if } \frac{\gamma_S}{\|(\sigma,\nu)\|}>0,\end{cases}$$

and set

$$\begin{aligned}\nu_2^*(\mathrm{d}x)&:=\nu^*(\mathrm{d}x)-\nu_1^*(\mathrm{d}x)\\&=\begin{cases}-\frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x-1)\mathbb{1}_{\{x>0\}}\nu_1(\mathrm{d}x)+\Big(1-\frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x-1)\Big)\nu_2(\mathrm{d}x) & \text{if } \frac{\gamma_S}{\|(\sigma,\nu)\|}\leqslant 0\\ \frac{\gamma_S}{\|(\sigma,\nu)\|}(1-\mathrm{e}^x)\mathbb{1}_{\{x<0\}}\nu_1(\mathrm{d}x)+\Big(1-\frac{\gamma_S}{\|(\sigma,\nu)\|}(\mathrm{e}^x-1)\Big)\nu_2(\mathrm{d}x) & \text{if } \frac{\gamma_S}{\|(\sigma,\nu)\|}>0.\end{cases}\end{aligned}$$

It is clear that ν_1^* and ν_2^* are Lévy measures. Moreover, a short calculation shows that ν_1^* and ν_2^* satisfy (5.13) and (5.12) respectively, which verifies $\nu^*\in\mathcal{S}_1(\alpha)$.

We now prove the statement for $i=2$. Assume that $\nu\in\mathcal{S}_2(\alpha)$. Let $\varepsilon\in(0,1)$ and $\delta>0$ be such that $|\frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}|<\varepsilon$ for all $|x|<\delta$. Then,

$$\begin{aligned}\int_{\mathbb{R}}(1-\cos(ux))\nu^*(\mathrm{d}x)&\geqslant\int_{|x|<\delta}(1-\cos(ux))\left(1-\frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\right)\nu(\mathrm{d}x)\\&\geqslant(1-\varepsilon)\int_{|x|<\delta}(1-\cos(ux))\nu(\mathrm{d}x)\\&=(1-\varepsilon)\left(\int_{\mathbb{R}}(1-\cos(ux))\nu(\mathrm{d}x)-\int_{|x|\geqslant\delta}(1-\cos(ux))\nu(\mathrm{d}x)\right).\end{aligned}$$

Since $\sup_{u\in\mathbb{R}}|\int_{|x|\geqslant\delta}(1-\cos(ux))\nu(\mathrm{d}x)|\leqslant 2\nu(\mathbb{R}\backslash(-\delta,\delta))<\infty$, it implies that

$$\liminf_{|u|\to\infty}\frac{1}{|u|^\alpha}\int_{\mathbb{R}}(1-\cos(ux))\nu^*(\mathrm{d}x)\geqslant(1-\varepsilon)\liminf_{|u|\to\infty}\frac{1}{|u|^\alpha}\int_{\mathbb{R}}(1-\cos(ux))\nu(\mathrm{d}x)>0.$$

For the upper limit, one has

$$\begin{aligned}&\int_{\mathbb{R}}(1-\cos(ux))\nu^*(\mathrm{d}x)\leqslant(1+\varepsilon)\int_{|x|<\delta}(1-\cos(ux))\nu(\mathrm{d}x)+\int_{|x|\geqslant\delta}(1-\cos(ux))\nu^*(\mathrm{d}x)\\&=(1+\varepsilon)\left(\int_{\mathbb{R}}-\int_{|x|\geqslant\delta}\right)(1-\cos(ux))\nu(\mathrm{d}x)+\int_{|x|\geqslant\delta}(1-\cos(ux))\nu^*(\mathrm{d}x),\end{aligned}$$

and hence,

$$\limsup_{|u|\to\infty}\frac{1}{|u|^\alpha}\int_{\mathbb{R}}(1-\cos(ux))\nu^*(\mathrm{d}x)\leqslant(1+\varepsilon)\limsup_{|u|\to\infty}\frac{1}{|u|^\alpha}\int_{\mathbb{R}}(1-\cos(ux))\nu(\mathrm{d}x)<\infty.$$

Combining those arguments, we get $\nu^*\in\mathcal{S}_2(\alpha)$. □

Proof of Theorem 5.12. Recall $(X|\mathbb{P}^*)\sim(\gamma^*,\sigma^*,\nu^*)$ from (4.6). Since the function g in Table 1 has at most linear growth and $\int_{|x|>1}\mathrm{e}^{3x}\nu(\mathrm{d}x)<\infty$, which is equivalent to $\int_{|x|>1}\mathrm{e}^{2x}\nu^*(\mathrm{d}x)<\infty$ by Lemma 5.14(2), the assumptions of Theorem 1.1 are satisfied so that (5.15) is applicable.

(1) Combining Lemma 5.14(2) with Lemma A.3, we obtain that $\Phi(\eta)\in\mathcal{SM}_3(\mathbb{P})\cap\mathcal{SM}_2(\mathbb{P}^*)$. Thanks to Lemma A.2, one has $\overline{\Phi}(\eta)\in\mathcal{SM}_3(\mathbb{P})\cap\mathcal{SM}_2(\mathbb{P}^*)$.

(2) We recall $\mathcal{E}(U)$ from Definition 4.4 and notice that $\mathcal{E}(U) > 0$ due to Assumption 4.5. According to Subsection 2.4, there is a Lévy process V with $(V|\mathbb{P}) \sim (\gamma_V, \sigma_V, \nu_V)$ such that $\mathcal{E}(U) = \mathrm{e}^V$. Due to (4.5), by letting $h(x) := \ln(1+x)$ for $x > -1$ one has

$$\nu_V = \nu_U \circ h^{-1} = (\nu \circ \alpha_U^{-1}) \circ h^{-1} = \nu \circ (h \circ \alpha_U)^{-1}. \tag{5.18}$$

Since $h(\alpha_U(x)) = \ln\big(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\big)$ for $x \in \operatorname{supp}\nu$, there exists an $\varepsilon_{(5.19)} > 0$ such that

$$\{x \in \operatorname{supp}\nu : |h(\alpha_U(x))| > 1\} \subseteq \mathbb{R}\backslash(-\varepsilon_{(5.19)}, \varepsilon_{(5.19)}). \tag{5.19}$$

Then, the assumption $\int_{|x|>1} \mathrm{e}^{3x}\nu(\mathrm{d}x) < \infty$ implies that

$$\int_{|x|>1} \mathrm{e}^{3x}\nu_V(\mathrm{d}x) = \int_{|h(\alpha_U(x))|>1} \mathrm{e}^{3(h(\alpha_U(x)))}\nu(\mathrm{d}x) \leqslant \int_{|x|\geqslant\varepsilon_{(5.19)}} \left(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\right)^3 \nu(\mathrm{d}x) < \infty,$$

Let $(V|\mathbb{P}) \sim \psi_V$. Since $(\mathrm{e}^{3V_t + t\psi_V(-3\mathrm{i})})_{t\in[0,T]}$ is a càdlàg martingale, it follows from the optional stopping theorem that for any stopping time $\rho\colon \Omega \to [0,T]$, a.s.,

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_\rho}\left[\mathrm{e}^{3V_T}\right] &= \mathrm{e}^{-T\psi_V(-3\mathrm{i})}\mathbb{E}_{\mathcal{F}_\rho}\left[\mathrm{e}^{3V_T+T\psi_V(-3\mathrm{i})}\right] = \mathrm{e}^{-T\psi_V(-3\mathrm{i})}\mathrm{e}^{3V_\rho+\rho\psi_V(-3\mathrm{i})}\\ &\leqslant \mathrm{e}^{T|\psi_V(-3\mathrm{i})|}\mathrm{e}^{3V_\rho} = \mathrm{e}^{T|\psi_V(-3\mathrm{i})|}\left|\mathbb{E}_{\mathcal{F}_\rho}\left[\mathrm{e}^{V_T}\right]\right|^3,\end{aligned}$$

where we use the martingale property of e^V for the last equality. According to Definition 5.2 and Proposition 5.3(3), $\mathrm{d}\mathbb{P}^* = \mathrm{e}^{V_T}\mathrm{d}\mathbb{P} \in \mathcal{RH}_3(\mathbb{P})$.

(3) In the notations of Assumption 5.6, let

$$\tilde{\vartheta} = \tilde{\vartheta}^g, \quad \Upsilon(\cdot,\mathrm{d}t) = \mathrm{d}\langle M, M\rangle_t^{\mathbb{P}^*} + M_t^2\mathrm{d}t, \quad \Theta = \Theta(\eta), \quad \Phi = \Phi(\eta).$$

We now verify the requirements of Assumption 5.6 *under the measure* $\mathbb{P}^*$.

Item (i) is clear. For item (ii), Theorem 1.1(2) verifies that $M = \tilde{\vartheta}^g S$ is a $\mathbb{P}^*$-martingale adapted to the augmented natural filtration of X, which is a quasi-left continuous filtration (see [29, p.150]). This implies that $\tilde{\vartheta}^g_{t-}S_{t-} = \tilde{\vartheta}^g_t S_t$ a.s. for each $t \in [0,T)$ (see [29, p.191]), and hence $\tilde{\vartheta}^g_{t-} = \tilde{\vartheta}^g_t$ a.s. due to $S_{t-} = S_t$ a.s.

For item (iii), we can prove (5.4) as in [37, Example 3.2] (with $\sigma(x) = x$), where the square $\mathbb{P}^*$-integrability of M can be inferred from (5.20).

For item (iv), it follows from the proof of [37, Theorem 4.6(3)] that for any $a \in [0,T)$, a.s.,

$$\begin{aligned}&\mathbb{E}^*_{\mathcal{F}_a}\left[\int_{(a,T)} (T-t)^{1-\theta}\Upsilon(\cdot,\mathrm{d}t)\right]\\ &\leqslant \begin{cases}\mathbb{E}^*_{\mathcal{F}_a}\Big[\lim_{t\uparrow T} M_t^2 + \int_{(a,T)} M_t^2\mathrm{d}t\Big] & \text{if } \theta = 1\\ \mathbb{E}^*_{\mathcal{F}_a}\Big[\int_{(a,T)} \left((1-\theta)(T-t)^{-\theta} + (T-t)^{1-\theta}\right) M_t^2\mathrm{d}t\Big] & \text{if } \theta \in (0,1).\end{cases}\end{aligned}$$

Hence, in order to verify (iv), thanks to $\Phi(\eta) \in \mathcal{SM}_2(\mathbb{P}^*)$, it suffices to show that there is a constant $c_{(5.20)} > 0$ which might depend on $\hat{\theta}$ but is independent of t such that

$$|\tilde{\vartheta}^g_t| \leqslant c_{(5.20)}(T-t)^{\frac{\hat{\theta}-1}{2}}\Theta(\eta)_t \quad \text{a.s., } \forall t \in [0,T), \tag{5.20}$$

where $\hat{\theta}$ is given according to the cases C1 and C2 of Table 1 as follows

$$\hat{\theta} = \begin{cases}\eta & \text{in the case C1}\\ 1 & \text{in the case C2.}\end{cases}$$

Regarding C3 and C4, it is sufficient to prove that (5.20) holds for any $\hat{\theta} \in \big(0, \frac{2(1+\eta)}{\alpha} - 1\big)$.

Indeed, we first let $\mathbb{Q} := \mathbb{P}^*$ and $\ell := \nu$ in (A.1) and then derive from (5.15) that

$$\tilde{\vartheta}^g_t = \frac{\varGamma^{\mathbb{P}^*}_\nu(T-t, S_t)}{\|(\sigma,\nu)\|} \quad \text{a.s., } \forall t \in [0,T).$$

Case C1: Since $\sigma^* = \sigma > 0$ and $\int_{|x|>1} \mathrm{e}^{2x}\nu^*(\mathrm{d}x) < \infty$, Proposition A.4(1) implies (5.20) with $\hat{\theta} = \eta$.
Case C2: Since $\int_{|x|\leqslant 1} |x|^{\eta+1}\nu(\mathrm{d}x) < \infty$, combining Lemma 5.14(1) with Proposition A.4(2) we obtain (5.20) with $\hat{\theta} = 1$.
Case C3: Due to Lemma 5.15, we have $\nu^* \in \mathcal{S}_1(\alpha)$. Let $\varepsilon \in (0, 2-\alpha]$ be arbitrary. Then it follows from Lemma A.1(2) that $\int_{|x|\leqslant 1} |x|^{\alpha+\varepsilon}\nu(\mathrm{d}x) < \infty$. We apply Proposition A.4(3) and Remark A.5 with $\beta = \alpha+\varepsilon$ to obtain that, for any $t \in [0,T)$, a.s.,

$$|\tilde{\vartheta}^g_t| = \frac{|\varGamma^{\mathbb{P}^*}_\nu(T-t, S_t)|}{\|(\sigma,\nu)\|} \leqslant c_\varepsilon (T-t)^{\frac{\eta+1}{\alpha}-1-\frac{\varepsilon}{\alpha}} S_t^{\eta-1} \leqslant c_\varepsilon (T-t)^{\frac{1}{2}\left(\left(\frac{2(\eta+1)}{\alpha}-1-\frac{2\varepsilon}{\alpha}\right)-1\right)} \Theta(\eta)_t,$$

where $c_\varepsilon > 0$ is some constant which might depend on ε. Since $\varepsilon > 0$ can be arbitrarily small, the assertion (5.20) holds for any $\hat{\theta} \in \left(0, \frac{2(1+\eta)}{\alpha} - 1\right)$.
Case C4: Again, one has $\nu^* \in \mathcal{S}_2(\alpha)$ due to Lemma 5.15, and Lemma A.1(2) verifies $\int_{|x|\leqslant 1} |x|^{\alpha+\varepsilon}\nu(\mathrm{d}x) < \infty$ for all $\varepsilon \in (0, 2-\alpha]$. By Proposition A.4(4) and Remark A.5 with $\beta = \alpha+\varepsilon$ and by the same reason as in the case C3 above, we get (5.20).

(4) By the relation between the behavior of ν and of ν^* around zero given in Lemma 5.14(1), we use item (3) and apply (5.9) to obtain (5.16).

(5) *Step 1.* For ν_V given in (5.18), we first show that $\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu_V(\mathrm{d}x) < \infty$. Indeed,

$$\begin{aligned}\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu_V(\mathrm{d}x) &= \int_{|h(\alpha_U(x))|>1} \left(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\right)^{1-s} \nu(\mathrm{d}x)\\ &\leqslant \int_{|x|\geqslant \varepsilon_{(5.19)}} \left(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\right)^{1-s} \nu(\mathrm{d}x) =: \mathrm{I}_{(5.21)}.\end{aligned} \tag{5.21}$$

We consider three cases regarding $\frac{\|(\sigma,\nu)\|}{\gamma_S}$ as follows:
Case 1: $\frac{\|(\sigma,\nu)\|}{\gamma_S} > -1$. We denote $x_0 := \ln\left(1 + \frac{\|(\sigma,\nu)\|}{\gamma_S}\right)$. Then, Assumption 4.5 verifies that $x_0 \notin \operatorname{supp}\nu$, which means $\nu((x_0-\varepsilon_0, x_0+\varepsilon_0)) = 0$ for some $\varepsilon_0 > 0$. Moreover, using the mean value theorem we infer that $1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|} \geqslant |x-x_0|\frac{|\gamma_S|}{\|(\sigma,\nu)\|}\mathrm{e}^{x\wedge x_0}$ for all $x \in \operatorname{supp}\nu$. Hence,

$$\begin{aligned}\mathrm{I}_{(5.21)} &= \int_{|x|\geqslant\varepsilon_{(5.19)}, |x-x_0|\geqslant\varepsilon_0} \left(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\right)^{1-s} \nu(\mathrm{d}x),\\ &\leqslant \varepsilon_0^{1-s}\frac{|\gamma_S|^{1-s}}{\|(\sigma,\nu)\|^{1-s}} \int_{|x|\geqslant\varepsilon_{(5.19)}, |x-x_0|\geqslant\varepsilon_0} \mathrm{e}^{(1-s)(x\wedge x_0)}\nu(\mathrm{d}x)\\ &\leqslant \varepsilon_0^{1-s}\frac{|\gamma_S|^{1-s}}{\|(\sigma,\nu)\|^{1-s}} \int_{|x|\geqslant\varepsilon_{(5.19)}} \mathrm{e}^{(1-s)(x\wedge x_0)}\nu(\mathrm{d}x) < \infty,\end{aligned}$$

where the finiteness is due to the assumption $\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu(\mathrm{d}x) < \infty$.
Case 2: $\frac{\|(\sigma,\nu)\|}{\gamma_S} = -1$. We have $\mathrm{I}_{(5.21)} = \int_{|x|\geqslant\varepsilon_{(5.19)}} \mathrm{e}^{(1-s)x}\nu(\mathrm{d}x) < \infty$.

Case 3: $\frac{\|(\sigma,\nu)\|}{\gamma_S} < -1$. In this case, one has $\gamma_S < 0$, which implies that $\inf_{x\in\mathbb{R}}\big(1 - \frac{\gamma_S(\mathrm{e}^x-1)}{\|(\sigma,\nu)\|}\big) = 1 + \frac{\gamma_S}{\|(\sigma,\nu)\|} > 0$. Hence,

$$\mathrm{I}_{(5.21)} \leqslant \left(1 + \frac{\gamma_S}{\|(\sigma,\nu)\|}\right)^{1-s} \int_{|x|\geqslant\varepsilon_{(5.19)}} \nu(\mathrm{d}x) < \infty.$$

We conclude from three cases above that $\int_{|x|>1} \mathrm{e}^{(1-s)x}\nu_V(\mathrm{d}x) < \infty$, or equivalently

$$\mathrm{e}^{-t\psi_V((s-1)\mathrm{i})} = \mathbb{E}\mathrm{e}^{(1-s)V_t} < \infty, \quad t > 0.$$

Step 2. We show $\mathbb{P} \in \mathcal{RH}_s(\mathbb{P}^*)$. By writing $\mathrm{d}\mathbb{P} = \mathrm{e}^{-V_T}\mathrm{d}\mathbb{P}^*$ and since $\mathrm{e}^V = \mathcal{E}(U)$ is a $\mathbb{P}$-martingale, it implies that e^{-V} is a $\mathbb{P}^*$-martingale. We have for any $t \in [0,T]$ that, a.s.,

$$\mathbb{E}^*_{\mathcal{F}_t}\Big[\mathrm{e}^{s(-V_T)}\Big] = \mathrm{e}^{-V_t}\mathbb{E}_{\mathcal{F}_t}\big[\mathrm{e}^{-sV_T}\mathrm{e}^{V_T}\big] = \mathrm{e}^{-V_t}\mathbb{E}_{\mathcal{F}_t}\Big[\mathrm{e}^{(1-s)V_T}\Big] \leqslant \mathrm{e}^{T|\psi_V((s-1)\mathrm{i})|}\mathrm{e}^{-sV_t}.$$

By a similar argument as in the proof of [15, Proposition A.1], we infer that

$$\mathbb{E}^*_{\mathcal{F}_\rho}\Big[\mathrm{e}^{s(-V_T)}\Big] \leqslant \mathrm{e}^{T|\psi_V((s-1)\mathrm{i})|}\mathrm{e}^{-sV_\rho} = \mathrm{e}^{T|\psi_V((s-1)\mathrm{i})|}\big|\mathbb{E}^*_{\mathcal{F}_\rho}\big[\mathrm{e}^{-V_T}\big]\big|^s \quad \text{a.s.}$$

for any stopping times $\rho\colon \Omega \to [0,T]$, which implies $\mathbb{P} \in \mathcal{RH}_s(\mathbb{P}^*)$.

Step 3. Thanks to *Step 2* and items (1), (2), we apply Proposition 5.3(3) to obtain

$$\|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P}^*)} \sim_c \|\cdot\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})}.$$

Then, assertion (5.17) is clear due to (5.16). The "Moreover" part holds because of $\overline{\Phi}(\eta) \in \mathcal{SM}_3(\mathbb{P})$ and (5.2).

(6) A similar argument as in the proof of item (1) shows that both $\Phi(\eta)$ and $\overline{\Phi}(\eta)$ belong to $\mathcal{SM}_p(\mathbb{P}) \cap \mathcal{SM}_{p-1}(\mathbb{P}^*)$. We now apply (5.2) to derive the assertion. □

Appendix A. Some technical results

A.1. **Some properties of classes $\mathcal{S}_1(\alpha)$ and $\mathcal{S}_2(\alpha)$.** We recall $\mathcal{S}_1(\alpha)$ and $\mathcal{S}_2(\alpha)$ from Definition 5.10.

Lemma A.1 (See also [37], Remark 4.5)**.** *For $\alpha \in (0,2)$, the following assertions hold:*

(1) $\mathcal{S}_1(\alpha) \subsetneq \mathcal{S}_2(\alpha)$.

(2) *If $\nu \in \mathcal{S}_2(\alpha)$, then $\alpha = \inf\{r \in [0,2] : \int_{|x|\leqslant 1} |x|^r\nu(\mathrm{d}x) < \infty\}$.*

(3) *If a Lévy measure ν has a density $p(x) := \frac{\nu(\mathrm{d}x)}{\mathrm{d}x}$ which satisfies*

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

then $\nu \in \mathcal{S}_1(\alpha)$.

Proof. (1) Let $\mathcal{S}_1(\alpha) \ni \nu = \nu_1 + \nu_2$. A short calculation shows that (5.14) holds for ν_1 in place of ν. Combining this with (5.12) yields that (5.14) holds for ν, and hence $\nu \in \mathcal{S}_2(\alpha)$. Since $\nu(\mathrm{d}x) := x^{-1-\alpha}\mathbb{1}_{(0,1)}(x)\mathrm{d}x \in \mathcal{S}_2(\alpha)\backslash\mathcal{S}_1(\alpha)$, the inclusion $\mathcal{S}_1(\alpha) \subseteq \mathcal{S}_2(\alpha)$ is strict.

(2) follows from [4, Theorem 3.2].

(3) By assumption, there exist constants $0 < c \leqslant C < \infty$ and $\varepsilon > 0$ such that

$$c|x|^{-1-\alpha} \leqslant p(x) \leqslant C|x|^{-1-\alpha}, \quad \forall |x| \leqslant \varepsilon.$$

We let

$$\nu_1(\mathrm{d}x) := c\mathbb{1}_{\{0<|x|\leqslant\varepsilon\}}|x|^{-1-\alpha}\mathrm{d}x \quad \text{and} \quad \nu_2(\mathrm{d}x) := \nu(\mathrm{d}x) - \nu_1(\mathrm{d}x).$$

Then, ν_1 satisfies (5.13). For ν_2, we have

$$\int_{\mathbb{R}}(1-\cos(ux))\nu_2(\mathrm{d}x) \leqslant (C-c)\int_{|x|\leqslant\varepsilon}\frac{1-\cos(ux)}{|x|^{1+\alpha}}\mathrm{d}x + 2\int_{|x|>\varepsilon}\nu(\mathrm{d}x),$$

which implies that (5.12) holds for ν_2. Hence, $\nu \in \mathcal{S}_1(\alpha)$. □

A.2. **Regularity of weight processes.** Let $T \in (0,\infty)$. We assume that $\mathbb{Q}$ is a probability measure and $X = (X_t)_{t\in[0,T]}$ is a Lévy process with $(X|\mathbb{Q}) \sim (\gamma^{\mathbb{Q}}, \sigma^{\mathbb{Q}}, \nu^{\mathbb{Q}})$.

The regularity of the weight $\overline{\Phi}$ used in Theorem 5.7 is verified by Lemma A.2 below. For $\Phi \in \mathrm{CL}^+([0,T])$, we let $\overline{\Phi} \in \mathrm{CL}^+([0,T])$ by setting

$$\overline{\Phi}_t := \Phi_t + \sup\nolimits_{u\in[0,t]}|\Delta\Phi_u|, \quad t \in [0,T].$$

It is clear that $\Phi \vee \Phi_- \leqslant \overline{\Phi}$, and $\Phi \equiv \overline{\Phi}$ if and only if Φ is continuous.

Lemma A.2 ([37], Proposition 7.1). *If $\Phi \in \mathcal{SM}_q(\mathbb{Q})$ for some $q \in (0,\infty)$, then $\overline{\Phi} \in \mathcal{SM}_q(\mathbb{Q})$.*

We next recall the process $\Phi(\eta) \in \mathrm{CL}^+([0,T])$ used in Theorem 5.12, that is

$$\Phi(\eta)_t := \mathrm{e}^{X_t}\sup\nolimits_{u\in[0,t]}\mathrm{e}^{(\eta-1)X_u}, \quad t\in[0,T], \eta\in[0,1].$$

Lemma A.3 ([37], Proposition 7.2). *If $\int_{|x|>1}\mathrm{e}^{qx}\nu^{\mathbb{Q}}(\mathrm{d}x) < \infty$ for some $q\in(1,\infty)$, then $\Phi(\eta) \in \mathcal{SM}_q(\mathbb{Q})$ for all $\eta\in[0,1]$.*

A.3. **Gradient type estimates for a Lévy semigroup on Hölder spaces.** Assume that $X = (X_t)_{t\geqslant 0}$ is a Lévy process with respect to a probability measure $\mathbb{Q}$ with $(X|\mathbb{Q}) \sim (\gamma^{\mathbb{Q}}, \sigma^{\mathbb{Q}}, \nu^{\mathbb{Q}})$. We let

$$\mathcal{D}_{\exp}(X|\mathbb{Q}) := \left\{g\colon \mathbb{R}_+ \to \mathbb{R} \text{ Borel} : \mathbb{E}^{\mathbb{Q}}|g(y\mathrm{e}^{X_t})| < \infty, \forall y>0, t\geqslant 0\right\},$$

where $\mathbb{E}^{\mathbb{Q}}$ is the expectation computed under $\mathbb{Q}$. For $t \geqslant 0$, we define $Q_t\colon \mathcal{D}_{\exp}(X|\mathbb{Q}) \to \mathcal{D}_{\exp}(X|\mathbb{Q})$ by setting

$$Q_t g(y) := \mathbb{E}^{\mathbb{Q}} g(y\mathrm{e}^{X_t}).$$

It is clear that $Q_{t+s} = Q_t \circ Q_s$ for all $s,t \geqslant 0$ which means that $(Q_t)_{t\geqslant 0}$ is a semigroup.

For a Lévy measure ℓ on $\mathcal{B}(\mathbb{R})$ and a Borel function g, we write symbolically

$$\mathit{\Gamma}^{\mathbb{Q}}_{\ell}(t,y) := |\sigma^{\mathbb{Q}}|^2\partial_y Q_t g(y) + \int_{\mathbb{R}}\frac{Q_t g(\mathrm{e}^x y) - Q_t g(y)}{y}(\mathrm{e}^x - 1)\ell(\mathrm{d}x) \tag{A.1}$$

for $(t,y) \in \mathbb{R}^2_+$, where $\partial_y Q_t g := 0$ if $\sigma^{\mathbb{Q}} = 0$. We recall $C^{0,\eta}(\mathbb{R}_+)$, $\mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$ from Definition 5.9 and $\mathcal{S}_1(\alpha)$, $\mathcal{S}_2(\alpha)$ from Definition 5.10.

Proposition A.4 ([37], Proposition 8.6). *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 exists a constant $c_{(\mathrm{A.2})} > 0$ such that*

$$|\mathit{\Gamma}^{\mathbb{Q}}_{\ell}(t,y)| \leqslant c_{(\mathrm{A.2})} R_t y^{\eta-1}, \quad \forall (t,y)\in(0,T]\times\mathbb{R}_+, \tag{A.2}$$

where the cases for R_t are provided in the following cases:

(1) *If $\sigma^{\mathbb{Q}} > 0$ and $\int_{|x|>1}\mathrm{e}^{2x}\nu^{\mathbb{Q}}(\mathrm{d}x) < \infty$, then $R_t = t^{\frac{\eta-1}{2}}$.*

(2) *If $\sigma^{\mathbb{Q}} = 0$, $\int_{|x|>1}\mathrm{e}^{\eta x}\nu^{\mathbb{Q}}(\mathrm{d}x) < \infty$ and $\int_{|x|\leqslant 1}|x|^{\eta+1}\ell(\mathrm{d}x) < \infty$, then $R_t = 1$.*

(3) *If $\sigma^{\mathbb{Q}} = 0$ and if the following two conditions hold:*
   (a) $\nu^{\mathbb{Q}} \in \mathcal{S}_1(\alpha)$ *for some* $\alpha \in (0,2)$ *and* $\int_{|x|>1} \mathrm{e}^x \nu^{\mathbb{Q}}(\mathrm{d}x) < \infty$,
   (b) *there is a* $\beta \in (1+\eta, 2]$ *such that*

$$0 < \sup_{r \in (0,1]} r^\beta \int_{|x| \leqslant 1} \left(\left| \frac{x}{r} \right|^2 \wedge \left| \frac{x}{r} \right|^{\eta+1} \right) \ell(\mathrm{d}x) < \infty, \tag{A.3}$$

   then one has $R_t = t^{\frac{\eta+1-\beta}{\alpha}}$.

(4) *If* $\sigma^{\mathbb{Q}} = 0$, $g \in \mathring{W}^{1,\frac{1}{1-\eta}}(\mathbb{R}_+)$, *and if the following two conditions hold:*
   (a) $\nu^{\mathbb{Q}} \in \mathcal{S}_2(\alpha)$ *for some* $\alpha \in (0,2)$ *and* $\int_{|x|>1} \mathrm{e}^x \nu^{\mathbb{Q}}(\mathrm{d}x) < \infty$,
   (b) *there is a* $\beta \in (1+\eta, 2]$ *such that* (A.3) *is satisfied,*
   then one has $R_t = t^{\frac{\eta+1-\beta}{\alpha}}$.

Here, the constant $c_{(\mathrm{A.2})}$ *might depend on* β *in items* (3) *and* (4).

Remark A.5. Since $|\frac{x}{r}|^2 \wedge |\frac{x}{r}|^{\eta+1} \leqslant |\frac{x}{r}|^\beta$ for $\beta \in (1+\eta, 2]$, a sufficient condition for (A.3) is that $0 < \int_{|x|\leqslant 1} |x|^\beta \ell(\mathrm{d}x) < \infty$.

Acknowledgment. The author would like to sincerely thank his Ph.D. supervisors, Christel Geiss and Stefan Geiss, for helpful discussions.

References

1. D. Applebaum, *Lévy processes and stochastic calculus (2nd ed.)*, Cambridge University Press, Cambridge, 2009.
2. T. Arai, R. Suzuki, Local risk-minimization for Lévy markets, *Int. J. Financ. Eng.* 2(2), 2015, 1550015.
3. F. Benth, G. Di Nunno, A. Løkka, B. Øksendal, F. Proske, Explicit representation of the minimal variance portfolio in markets driven by Lévy processes, *Math. Finance* 13(1), 2003, 55–72.
4. R. Blumenthal, R. Getoor, Sample functions of stochastic processes with stationary independent increments, *J. Math. Mech.* 10, 1961, 493–516.
5. M. Brodén, P. Tankov, Tracking errors from discrete hedging in exponential Lévy model, *Int. J. Theor. Appl. Finance* 14(6), 2011, 803–837.
6. T. Choulli, N. Vandaele, M. Vanmaele, The Föllmer–Schweizer decomposition. Comparison and description, *Stochastic Process. Appl.* 120, 2010, 853–872.
7. T. Choulli, L. Krawczyk, C. Stricker, $\mathcal{E}$-martingales and their applications in mathematical finance, *Ann. Probab.* 26(2), 1998, 853–876.
8. R. Cont, P. Tankov, E. Voltchkova, *Hedging with options in models with jumps*, In: Stochastic Analysis and Applications – The Abel Symposium 2005, Springer, Berlin, 2007.
9. F. Esche, M. Schweizer, Minimal entropy preserves the Lévy property: how and why, *Stochastic Process. Appl.* 115, 2005, 299–327.
10. C. Geiss, E. Laukkarinen, Denseness of certain smooth Lévy functionals in $\mathbb{D}_{1,2}$, *Probab. Math. Statist.* 31, 2011, 1–15.
11. C. Geiss, S. Geiss, E. Laukkarinen, A note on Malliavin fractional smoothness for Lévy processes and approximation, *Potential Anal.* 39, 2013, 203–230.
12. C. Geiss, A. Steinicke, Malliavin derivative of random functions and applications to Lévy driven BSDEs, *Electron. J. Probab.* 21(10), 2016, 28 pp.
13. S. Geiss, Quantitative approximation of certain stochastic integrals, *Stochastics Stochastics Rep.* 73(3–4), 2002, 241–270.
14. S. Geiss, Weighted BMO and discrete time hedging within the Black-Scholes model, *Probab. Theory Related Fields* 132, 2005, 13–38.
15. S. Geiss, N. T. Thuan, On Riemann–Liouville operators, BMO, gradient estimates in the Lévy–Itô space, and approximation, 2020. arXiv:2009.00899.
16. S. Goutte, N. Oudjane, F. Russo, Variance optimal hedging for continuous time additive processes and applications, *Stochastics* 86(1), 2014, 147–185.
17. F. Hubalek, J. Kallsen, L. Krawczyk, Variance-optimal hedging for processes with stationary independent increments, *Ann. Appl. Probab.* 16(2), 2006, 853–885.

and Lévy-type processes, *J. Theor. Probab.* 25, 2012, 144–170.

18. K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, *Trans. Amer. Math. Soc.* 81, 1956, 253–263.
19. J. Jacob, S. Méléard, P. Protter, Explicit form and robustness of martingale representations, *Ann. Probab.* 28(4), 2000, 1747–1780.
20. J. Kallsen, A. Pauwels, Variance-optimal hedging in general affine stochastic volatility models, *Adv. Appl. Prob.* 42, 2010, 83–105.
21. E. Laukkarinen, *On Malliavin calculus and approximation of stochastic integrals for Lévy process*, PhD thesis, University of Jyväskylä, 2013.
22. E. Laukkarinen, Malliavin smoothness on the Lévy space with Hölder continuous or BV functionals, *Stochastic Process. Appl.* 130(8), 2020, 4766–4792.
23. A. Løkka, Martingale representation of functionals of Lévy processes, *Stoch. Anal. Appl.* 22(4), 2004, 867–892.
24. J. Ma, P. Protter, J. Zhang, Explicit form and path regularity of martingale representations, In: *Lévy processes: Theory and applications*, Springer, New York, 2001.
25. P. Malliavin, H. Airault, L. Kay, G. Letac, *Integration and probability*, Springer, New York, 1995.
26. P. Monat, C. Stricker, Föllmer–Schweizer decomposition and mean-variance hedging for general claims, *Ann. Probab.* 23(2), 1995, 605–628.
27. D. Nualart, E. Nualart, *Introduction to Malliavin calculus*, Cambridge University Press, Cambridge, 2018.
28. G. Di Nunno, B. Øksendal, F. Proske, *Malliavin calculus for Lévy processes with applications to finance*, Springer-Verlag Berlin Heidelberg, 2009.
29. P. Protter, *Stochastic integration and differential equations (2nd ed., ver. 2.1)*, Springer-Verlag, Berlin, 2005.
30. M. Rosenbaum, P. Tankov, Asymptotically optimal discretization of hedging strategies with jumps, *Ann. Appl. Probab.* 24(3), 2014, 1002–1048.
31. K. Sato, *Lévy processes and infinitely divisible distributions (2nd ed.)*, Cambridge University Press, Cambridge, 2013.
32. W. Schoutens, *Lévy processes in finance: Pricing financial derivatives*, Wiley, 2003.
33. M. Schweizer, On the minimal martingale measure and the Föllmer–Schweizer decomposition, *Stoch. Anal. Appl.*, 13(5), 1995, 573–599.
34. M. Schweizer, A guided tour through quadratic hedging approaches, In: *Option pricing, interest rates and risk management*, Cambridge Univ. Press, Cambridge, 2001.
35. J. Solé, F. Utzet, J. Vives, *Chaos expansion and Malliavin calculus for Lévy processes*, In: Stochastic Analysis and Applications – The Abel Symposium 2005, Springer, Berlin, 2007.
36. P. Tankov, *Pricing and hedging in exponential Lévy models: review of recent results*, Paris-Princeton Lecture Notes in Mathematical Finance, Springer, 2010.
37. N. T. Thuan, Approximation of stochastic integrals with jumps in weighted bounded mean oscillation spaces, 2020. arXiv:2009.02116.

[1]Department of Mathematics and Statistics, P.O.Box 35, FI-40014 University of Jyväskylä, Finland

Email address: `thuan.t.nguyen@jyu.fi`

[2]Department of Mathematics, Vinh University, 182 Le Duan, Vinh, Nghe An, Vietnam