Fine scale properties of ambit fields

– Limit theory and simulation –

PhD Thesis

Claudio Heinrich

Supervised by Mark Podolskij

Department of Mathematics
Aarhus University

April 2017

Fine Scale Properties of Ambit Fields
Limit Theory and Simulation

PhD thesis by
Claudio Heinrich
Department of Mathematics, Aarhus University
Ny Munkegade 118, 8000 Aarhus C, Denmark

Supervised by
Professor Mark Podolskij

Submitted April 30, 2017

Contents

Preface

The research presented in this thesis is the result of my three years of PhD studies at Aarhus University, Department of Mathematics. My work was supervised by Mark Podolskij. The thesis consists of an introductory chapter followed by three self-contained papers, two of which are at this time still work in progress but are expected to be ready for submission shortly.

Paper 1
A. Basse-O'Connor, C. Heinrich, M. Podolskij. On limit theory for Lévy semi-stationary processes. *Bernoulli*, 2016. Submitted.

Paper 2
A. Basse-O'Connor, C. Heinrich, M. Podolskij. On limit theory for functionals of stationary increments Lévy driven moving average processes. Work in progress.

Paper 3
C. Heinrich, M.S. Pakkanen, A.E.D. Veraart. A hybrid simulation scheme for volatility modulated moving average fields. Work in progress.

The introductory chapter provides mathematical background, mostly required for Papers 1 and 2. In these papers we investigate the limit theory for the power variation and related functionals of Lévy semi-stationary processes. Paper 3 is the result of my visit to Mikko Pakkanen and Almut Veraart in early 2016. We develop a simulation technique for a class of spatial ambit fields. The appendix contains technical details as well as the code for the simulation. I presented results contained in this thesis at the following conferences, seminars, workshops and visits.

- AHOI workshop on tempo-spatial stochastic processes, London, UK, February 2015
- DYNSTOCH conference 2015, Lund, Sweden, May 2015
- Annual DMV meeting 2015, Hamburg, Germany, September 2015
- Thiele seminar for applied mathematics, Aarhus, Denmark, December 2015
- Visit to Mikko Pakkanen and Almut Veraart at Imperial College London, London, UK, January to May 2016
- German Probability and Statistics Days 2016, Bochum, Germany, March 2016
- Workshop on spatio-temporal statistics, London, UK, April 2016
- DYNSTOCH conference 2016, Rennes, France, June 2016
- Workshop on rough path integration theory, Aarhus, Denmark, July 2016
- Conference on ambit fields, Aarhus, Denmark, August 2016
- ASC workshop on asymptotic statistic, Tokyo, Japan, January 2017

The past three years of my PhD studies have been a very rewarding and enjoyable time, and I am deeply grateful to the many people who contributed to that.

First of all, I would like to thank my supervisor Mark Podolskij for making it possible. When I was a student at Heidelberg University, his seminars and lectures sparked my interest in probability theory and statistics and I had the pleasure of writing my master thesis under his advice and guidance. To continue working with Mark was one of my main motivations for coming to Aarhus University. Here I met Andreas Basse-O'Connor who quickly became (not officially but in every other aspect) my cosupervisor. This thesis would not have been possible without Marks and Andreas' supervision and continuous motivation. In countless discussions and meetings they helped me to understand the wonders of stochastic processes and their limit theory. Working with them has been a truly inspiring experience and I am glad to have the opportunity to continue this work a little longer.

Studying is much easier in an environment where you feel at home, and for this I am deeply grateful to all my friends and colleagues in the statistics group at Aarhus University. In particular, my thanks goes to my current and former office mates Thorbjørn Grønbæk, Patrick Laub, Mikkel Nielsen, Victor Rhode, Christian Schmidt and Julie Thøgersen, who all helped me exploring the curiously enlightening effect of coffee & cake breaks.

My appreciation goes out to my coauthors Mikko Pakkanen and Almut Veraart, from Imperial College London, for making my visit possible and for many productive and pleasant meetings.

Last but not least I would like to thank my family and friends – for their wonderful never-ending support, advice on all matters of life and, occasionally, simply for distracting me from mathematics. All of this carried me through the last years and made me enjoy the time. This would not have been possible without your help and encouragement.

Claudio Heinrich,

Aarhus, April 2017

Summary

Ambit fields are a class of tempo-spatial stochastic processes that have been introduced for the purpose of modeling velocities in turbulent particle flows. The main contribution of this thesis is establishing limit theorems in the high frequency framework for a class of zero-spatial ambit processes called Lévy semi-stationary processes. These processes are of moving average type, driven by a pure jump Lévy process which is modulated by a stochastic volatility factor.

We establish the first order limit theory for power variations based on kth order increments of Lévy semi-stationary processes. The limiting behavior turns out to be heavily dependent on the interplay between the considered power, the order of increments k, the Blumenthal–Getoor index β of the driving Lévy process and the behaviour of the kernel function of the moving average at 0, which is specified by the power α. Our results can be used for statistical inference, in particular, they can be used to estimate the model parameters α and β. A natural generalisation of the power variation functional is obtained by applying an arbitrary continuous function f on kth order increments of the process. For this type of functionals the first order limit theory is investigated, when applied to stationary increments moving average processes, i.e. Lévy semi-stationary processes with constant volatility factor. In this framework we also prove a second order limit theorem, when the function f is bounded and the driving Lévy process is symmetric β-stable. Depending on the interplay of k, β and α, we obtain either a central limit theorem or convergence to a $(k-\alpha)\beta$-stable random variable. From a mathematical point of view, this part of the thesis extends the asymptotic theory investigated in the recent publication [20], where the first and partial second order limit theory for power variations of stationary increments Lévy driven moving averages have been studied.

In the last part of the thesis we develop and implement a simulation scheme for a certain class of spatial ambit fields often referred to as volatility modulated moving averages. Our technique of simulation is especially aimed at recovering the fine scale properties of the field correctly, and we demonstrate that it outperforms several other simulation schemes in that regard. The asymptotic behaviour of the mean square error of the simulation scheme is derived. The scheme relies on approximating the kernel function in the moving average representation partially by a step function and partially by a power function. For this type of approach the authors of [24], who considered a comparable model in one dimension, coined the expression hybrid simulation scheme.

Dansk sammenfatning

Ambit processer er en klasse af tids- og rumafhængige stokastiske processer som er blevet introduceret med formålet at modellere hastigheder af turbulente partikelstrømninger. Hovedbidraget fra denne afhandling er resultater om store tals love og stabile grænseværdisætninger i det højfrekvente tifælde, for en klasse af ambit processer kaldet Lévy semistationære processer. Disse processer er af typen glidende gennemsnit, som bliver drevet af en springfarlig Lévy process, hvor volatiliteten bliver påvirket af en stokastisk process.

Vi etablerer resultater om første ordens grænsesætninger for potens variation baseret på k-ordens tilvæksterne af Lévy semistationære processer. Opførslen af grænsevariablen viser sig at været dybt påvirket af sammenspillet mellem den betragtede potens, ordenen k af tilvæksterne, Blumenthal-Getoor indekset β for den drivende Lévy process og opførslen af integranden af det glidende gennemsnit tæt ved 0, som er specificeret ved potensen α. Vores resultater kan anvendes til statistisk inferens - mere præcist kan de bruges til at modellere parametrene α og β. En naturlig generalisering af potens variation funktionalet fås ved at anvendes en arbitrær funktion f på k'te ordens tilvæksterne af processen. For denne type af funktionaler viser vi resultater om første ordens grænsesætninger for glidende gennemsnit med stationære tilvækster, svarende til Lévy semistationære processer med konstant volatilitet. Indenfor denne ramme beviser vi også en anden-ordens grænseværdisætning, hvor funktionen f er begrænset og den drivende Lévy process har en symmetrisk β-stabil fordeling. Afhængig af sammenspillet mellem k, β og α, opnår vi enten en central grænseværdisætning eller en konvergens til en $(k-\alpha)\beta$-stabil stokastisk variabel. Fra et matematisk synspunkt udvider denne del af afhandlingen de nyere resultaterne fra artiklen [20], hvor første- og andens-ordens grænsesætninger for potens variation af glidende gennemsnit, med stationære tilvækster drevet af en Lévy process, er blevet studeret.

I den sidste del af afhandlingen udvikles og implementeres en simulationsalgoritme for en bestemt klasse af rumlige ambit processer - ofte refereret til som volatilitetsmodulerede glidende gennemsnit. Algoritmen approksimerer kernefunktionen i glidende gennemsnit's repræsentationen med en kombination af en trappefunktion og en potensfunktion. For denne type af approksimation har forfatterne i [24], som studerede en sammenlignelig model i en dimension, navngivet metoden hydrid simulationsalgoritme. Vores simulationsalgoritme sigter i særlig grad mod at reproducere opførslen af processen på mikroniveau, og vi demonstrerer at dette gør algoritmen bedre end flere andre simulationsalgortimer. Vi udleder den asymptotiske opførsel af den gennemsnitlige kvadratiske variation af simulationsalgoritmen.

Chapter 1

Introduction and preliminaries

A little over ten years ago, Ole E. Barndorff-Nielsen and Jürgen Schmiegel introduced the model of *ambit fields* in a series of papers [15, 16]. Their ambitious goal was to find a stochastic model that accurately captures characteristic features attributed to the velocity in turbulent flows, based on physical laws and measurements. Among these features are violent spontaneous changes in velocity and energy dissipation, i.e. the amount of kinetic turbulence energy transformed into heat by viscosity on small scales. Mathematically, an ambit field is a stochastic process, indexed by space and time, defined by the formula

$$X_t(x) = \int_{A_t(x)} g(t,s,x,\xi)\sigma_s(\xi)\, L(ds,d\xi) + \int_{D_t(x)} q(t,s,x,\xi)a_s(\xi)\, ds\, d\xi, \tag{1.1}$$

where L is a Lévy basis, to be defined in Section 1.3. The random value $X_t(x)$ models the turbulent velocity vector field at time $t \in \mathbb{R}_+$ and location in space $x \in \mathbb{R}^d$. The sets $A_t(x)$ and $D_t(x)$ resemble the area of space-time governing the velocity at (t,x), and are called *ambit sets*, giving the name to the model (the word 'ambit' comes from Latin and means 'sphere of influence'). The weight functions g and q are deterministic, whereas σ and a are stochastic processes representing aspects of the intermittency of the turbulence velocity field, which can be thought of as a measure for the local energy dissipation.

Ever since its introduction, ambit stochastics has been a rapidly expanding research field. The flexibility of the model quickly led to a range of applications beyond turbulence such as bioimaging, finance or metereology, examples being [6, 56, 67]. Recent reviews focusing on different aspects of ambit fields are [7, 13] and [69]. Due to the complexity of the model, much of the research so far focuses on one dimensional analoga of (1.1). Examples for such ambit processes can be obtained by observing an ambit field along a parametrised curve $(t, x(t))$ or by considering zero spatial ambit fields such as *Lévy semi-stationary* ($\mathcal{LSS}$) processes. A Lévy semi-stationary process is defined as

$$X_t = \int_{-\infty}^{t} \{g(t-s) - g_0(-s)\}\sigma_{s-}\, dL_s, \tag{1.2}$$

where g, g_0 are deterministic functions, σ is stochastic and L is a Lévy process. Papers I and II of this thesis investigate the limit theory for these processes in the high frequency setting, when the driving Lévy process is a pure jump process. In Paper I we consider the realised power variation of X based on kth order increments, defined as

$$V(p;k)_t^n := \sum_{i=k}^{[nt]} |\Delta_{i,k}^n X|^p, \qquad p > 0,$$

and derive its asymptotic behavior for $n \to \infty$. Here $[x]$ denotes the integer part of x, and the kth order increments $\Delta_{i,k}^n X$ are defined as

$$\Delta_{i,k}^n X := \sum_{j=0}^{k} (-1)^j \binom{k}{j} X_{(i-j)/n}, \qquad \text{for } i \geq k.$$

In particular, $\Delta_{i,1}^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ are the increments of the process, and $\Delta_{i,k}^n X = \Delta_{i,k-1}^n X - \Delta_{i-1,k-1}^n X$ for $k > 1$.

Over the last decade the limit theory of realised power variations has been an active field of research for a variety of stochastic processes. We refer to [18, 12] for the limit theory of Itô semimartingales, to [8, 10] for a class of Gaussian processes, including fractional Brownian motion, and to [32, 33] for the Rosenblatt process. In [9] the authors derive the limit theory for power variations of Brownian semi-stationary processes, which is the model (1.2) driven by a Brownian motion. In Paper I we present the first order limit theory for the power variation of $\mathcal{LSS}$ processes driven by a pure jump Lévy process. From a mathematical point of view this extends the asymptotic theory derived in [20], where the authors consider stationary increments Lévy driven moving averages, which is the model (1.2) with constant volatility σ. It turns out that the limiting behavior of $V(p;k)$ is divided into three different regimes, depending on the choice of p and k as well as on the Blumenthal-Getoor index β of the driving Lévy process and the behavior at 0 of the kernel function g, specified by the power α. We demonstrate that our results can be used to estimate α and β and the relative intermittency, which for $p = 2$ describes the relative amplitude of the velocity process on a fixed interval.

A natural generalisation of the realised power variation are functionals of the form

$$V(f;k)_t^n := a_n \sum_{i=k}^{[nt]} f(b_n \Delta_{i,k}^n X), \tag{1.3}$$

where f is a deterministic function, and $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are suitable normalising sequences. In Paper II we present the first order limit theory for such functionals, when X is a stationary increments Lévy driven moving average and the function f is continuous. Also in this framework the limiting behavior depends on the Blumenthal-Getoor index of the driving Lévy process L and the behavior of g at 0, as well as on certain properties of the function f. Moreover, we derive the second order asymptotics for one of the occurring cases, when the function f is bounded and has Appell rank greater than one.

Papers I and II of this thesis can be interpreted as a stepping stone on the road to understanding the limit theory for ambit fields. However, the model (1.2) is not only of high interest from an angle of ambit stochastics. The class of $\mathcal{LSS}$ processes contains as an important example linear fractional stable motions. This is the model (1.2) driven by a symmetric β-stable Lévy process, with $\sigma \equiv 1$ and $g(t) = g_0(t) = \max(t,0)^\alpha$ for some $\alpha \in (-1/\beta, 1-1/\beta)\setminus\{0\}$. These processes are self-similar of index $H = \alpha + 1/\beta$ and are a natural generalisation of the fractional Brownian motion, which is the linear fractional stable motion with $\beta = 2$. There is a wide range of literature on linear fractional stable motions and recent research addresses various topics such as, among others, semimartingale property [22], fine scale behavior [23, 40], simulation techniques [35] and statistical inference [4].

In Paper III we present and implement a simulation technique for purely spatial ambit fields of the form

$$X(x) = \int_{\mathbb{R}^2} g(x-\xi)\sigma(\xi)W(d\xi), \quad x \in \mathbb{R}^2,$$

where W is Gaussian white noise on $\mathbb{R}^2$. When the kernel g has a singularity at 0, the order of the singularity governs the roughness of the sample paths. The challenge in simulating X is to accurately recover the roughness while also capturing global properties of the model X. This can be achieved by using a hybrid simulation scheme that approximates the kernel g by a power function around 0, and by a step function away from 0. This idea is motivated by [24] where the authors propose a hybrid scheme for simulating Brownian semi-stationary processes, i.e. the model (1.2) driven by Brownian motion. We derive the asymptotic mean square error of the hybrid scheme and demonstrate in a simulation study that it outperforms other simulation methods in recovering the roughness of the field X.

In the remainder of this chapter we give some mathematical prerequisites that are essential for the results and proofs presented in this thesis. First we introduce our probabilistic setting and recall the main results of [20], which are an essential fundament for the theory and results presented in Papers I and II. Thereafter, in section 1.2, we give preliminaries for the proof of functional limit theorems and give some details on the Skorokhod M_1-topology. In section 1.3 we discuss the definition and important estimates for stochastic integrals with respect to Lévy processes. Section 1.4 summarises the key ideas and the intuition behind the proofs presented in Papers I and II.

1.1 Limit theory for Lévy driven moving average processes

In this section we introduce our basic assumptions and some notation. Thereafter we recall the limit theory for the power variation of stationary increments Lévy driven moving average processes presented in [20], which forms an essential prerequisite for Paper I and II of this thesis.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}}, \mathbb{P})$ be a filtered probability space. A Lévy process on the real line is an adapted process $(L_t)_{t\in\mathbb{R}}$ with stationary independent increments and càdlàg sample paths (the French acronym "càdlàg" stands for right continuous with left limits

– continue à droite, limite à gauche). We remark that the independence of increments is to be understood with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}}$ which might be larger than the filtration generated by L, i.e. $L_t - L_s$ is independent of $\mathcal{F}_s$ for all $s < t$. For simplicity we assume $L_0 = 0$.

The *Blumenthal–Getoor index* of L is defined as

$$\beta := \inf\Big\{r \geq 0 : \int_{-1}^{1} |x|^r \, \nu(dx) < \infty\Big\} \in [0,2],$$

where ν denotes the Lévy measure of L. Intuitively, this index measures the concentration of the small jumps of L. For example is $\beta = 0$ when L has only finitely many jumps on bounded intervals. It is well-known that $\sum_{s\in[0,1]} |\Delta L_s|^p$ is finite when $p > \beta$, while it is infinite for $p < \beta$. Here $\Delta L_s = L_s - L_{s-}$ where $L_{s-} = \lim_{u\uparrow s,\, u<s} L_u$. For a stable Lévy process with index of stability $\beta \in (0,2)$, the Blumenthal-Getoor index matches the index of stability and both will be denoted β.

Throughout this thesis, we will assume L to be a symmetric pure jump Lévy process, i.e. L has zero drift and no Gaussian part and its Lévy measure satisfies $\nu(-A) = \nu(A)$ for all $A \in \mathcal{B}(\mathbb{R})$. The functions g and g_0 in (1.2) are assumed to satisfy the following conditions, introduced in [20].

Assumption (A): *The function $g\colon \mathbb{R} \to \mathbb{R}$ satisfies*

$$g(t) \sim c_0 t^\alpha \qquad \text{as } t \downarrow 0 \quad \text{for some } \alpha > 0 \text{ and } c_0 \neq 0,$$

where $g(t) \sim f(t)$ as $t \downarrow 0$ means that $\lim_{t\downarrow 0} g(t)/f(t) = 1$. For some $\theta \in (0,2]$, $\limsup_{t\to\infty} \nu(x : |x| \geq t)t^\theta < \infty$ and $g - g_0$ is a bounded function in $L^\theta(\mathbb{R}_+)$. Furthermore, g is k-times continuously differentiable on $(0,\infty)$ and there exists a $\delta > 0$ such that $|g^{(k)}(t)| \leq Ct^{\alpha-k}$ for all $t \in (0,\delta)$, and such that both $|g'|$ and $|g^{(k)}|$ are in $L^\theta((\delta,\infty))$ and are decreasing on (δ,∞).

The volatility process σ (in the ambit framework usually called intermittency process) is assumed to be càdlàg and adapted, making the process $(\sigma_{t-})_{t\in\mathbb{R}}$ predictable. We recall that a stochastic process is called predictable if it is measurable with respect to the predictable σ-algebra on $\Omega \times \mathbb{R}$, which is generated by all left continuous adapted processes.

Occasionally, it is necessary to strengthen the condition $|g^{(k)}| \in L^\theta((\delta,\infty))$ slightly and assume the following.

Assumption (A-log): *In addition to (A) suppose that*

$$\int_\delta^\infty |g^{(k)}(s)|^\theta \log(1/|g^{(k)}(s)|)\, ds < \infty.$$

Assumption (A) ensures, in particular, that the process X with $\sigma = 1$ is well-defined, see Appendix A.3 for details. For θ as in the assumption, the Lévy process has moments of all orders smaller θ, cf. [78, Theorem 25.3]. When L is a β-stable Lévy process, we can and will always choose $\theta = \beta$. Even though the driving Lévy process is a pure jump process, it follows from the Kolmogorov moment criterion (see [59, Theorem 2.23]) that under the conditions above the process X admits a continuous version. Intuitively speaking, the kernel g smooths out the incoming shocks of the Lévy process, since it vanishes at 0. Indeed, the sample paths of X are smoother for larger α, and it is therefore not surprising that the parameter α has major influence on the limiting behaviour of the power variation. Visual evidence for this smoothing effect is given in Figure 1.1, where we show examples of Lévy driven moving average processes.

We now recall the first order limit theory for the power variation of stationary increments Lévy driven moving averages that was derived in [20]. To this end we introduce the following notation. Let $h_k : \mathbb{R} \to \mathbb{R}$ be given by

$$h_k(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (x-j)_+^{\alpha}, \qquad x \in \mathbb{R},$$

where $y_+ = \max\{y, 0\}$ for all $y \in \mathbb{R}$. Let $(T_m)_{m\geq 1}$ be a sequence of $\mathcal{F}$-stopping times that exhausts the jumps of $(L_t)_{t\geq 0}$. That is, $\{T_m(\omega) : m \geq 1\} \cap [0,\infty) = \{t \geq 0 : \Delta L_t(\omega) \neq 0\}$ and $T_m(\omega) \neq T_n(\omega)$ for all $m \neq n$ with $T_m(\omega) < \infty$. Let $(U_m)_{m\geq 1}$ be a sequence of independent and uniform $[0,1]$-distributed random variables, defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of the original probability space, which are independent of $\mathcal{F}$. For random variables $Z, Z_1, Z_2, \ldots$ defined on $(\Omega', \mathcal{F}', \mathbb{P}')$ we denote by $Z_n \xrightarrow{\mathcal{L}-s} Z$ the $\mathcal{F}$-stable convergence in law, see Section 1.2 for details.

Theorem 1.1.1 (Theorem 1.1, [20]). *Suppose that $X = (X_t)_{t\geq 0}$ is a stochastic process defined by* (1.2) *with $\sigma \equiv 1$, and that Assumption (A) is satisfied. Moreover, assume that the Blumenthal–Getoor index of L satisfies $\beta < 2$. Set $V(p;k)^n := V(p;k)_1^n$. We have the following three cases:*

(i) Suppose that (A-log) holds if $\theta = 1$. If $\alpha < k - 1/p$ and $p > \beta$ then the $\mathcal{F}$-stable convergence holds as $n \to \infty$

$$n^{\alpha p} V(p;k)^n \xrightarrow{\mathcal{L}-s} |c_0|^p \sum_{m:\, T_m \in [0,1]} |\Delta L_{T_m}|^p V_m, \qquad (1.4)$$

where $V_m = \sum_{l=0}^{\infty} |h_k(l + U_m)|^p$.

(ii) Suppose that L is a symmetric β-stable Lévy process with scale parameter $\gamma > 0$. If $\alpha < k - 1/\beta$ and $p < \beta$ then it holds

$$n^{-1+p(\alpha+1/\beta)} V(p;k)^n \xrightarrow{\mathbb{P}} m_p,$$

where $m_p = |c_0|^p \gamma^p (\int_{\mathbb{R}} |h_k(x)|^\beta \, dx)^{p/\beta} \mathbb{E}[|Z|^p]$ and Z is a symmetric β-stable random variable with scale parameter 1.

(iii) Suppose that $p \geq 1$. If $p = \theta$ suppose in addition that (A-log) holds. For all $\alpha > k - 1/(\beta \vee p)$ we have that

$$n^{-1+pk} V(p;k)^n \xrightarrow{\mathbb{P}} \int_0^1 |F_u|^p \, du,$$

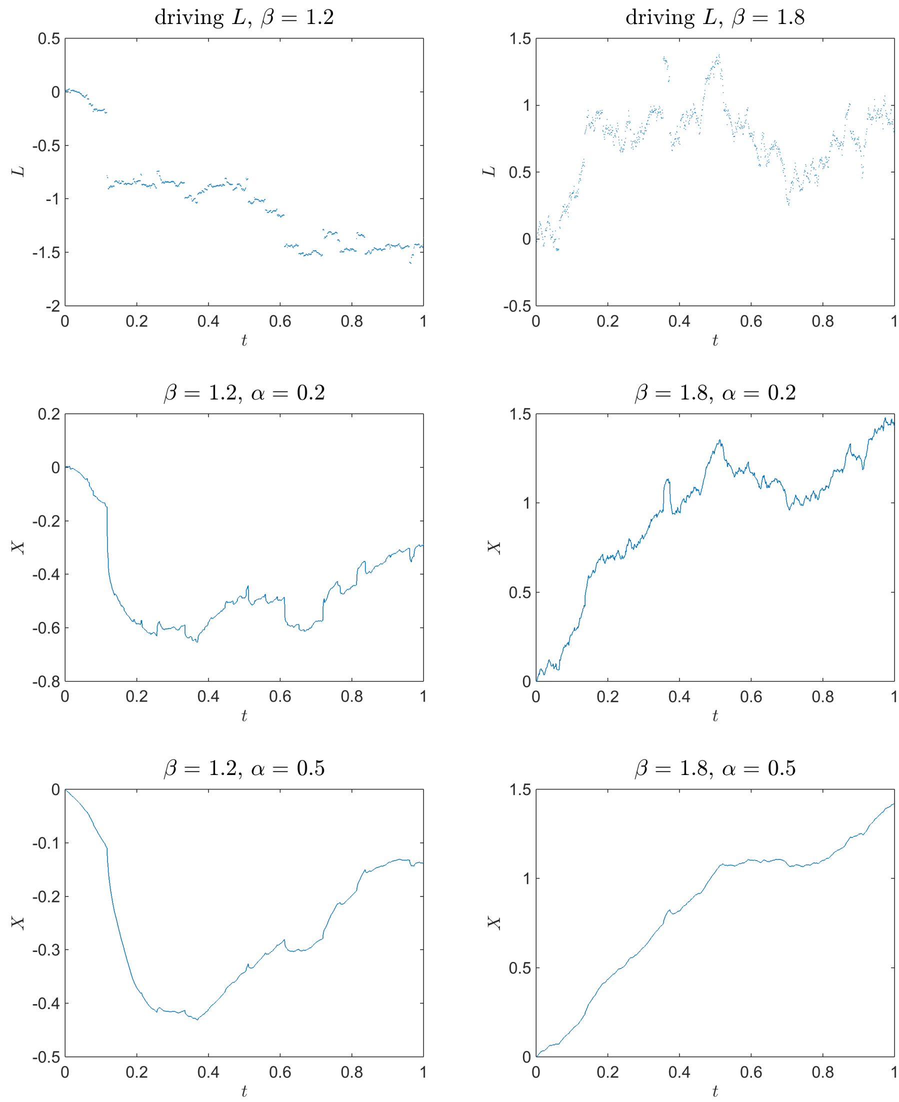

Figure 1.1: Realisations of the model (1.2) with constant volatility. The first row shows the driving Lévy process, row two and three show a Lévy driven moving average process X with $\alpha = 0.2$ and $\alpha = 0.5$, respectively. In the first column, the driving Lévy process is symmetric β-stable with $\beta = 1.2$ and in the second it is symmetric β-stable with $\beta = 1.8$. When the driving Lévy process has a jump that is much larger than the jumps surrounding it, the shape of the kernel function g at 0 becomes visible. The smoothing effect of the kernel, which becomes stronger as α increases, is apparent.

where $(F_u)_{u\in\mathbb{R}}$ *is a version with measurable sample paths of the process defined by*

$$F_u = \int_{-\infty}^{u} g^{(k)}(u-s)\, dL_s \quad \text{a.s. for all } u \in \mathbb{R},$$

which necessarily satisfies $\int_0^1 |F_u|^p\, du < \infty$, *almost surely.*

For a β-stable driving Lévy process and for $p \geq 1$, these three cases cover all possible configurations of α, β, p and k except the critical cases $p = \beta$ and $\alpha = k - 1/(\beta \vee p)$. The limit theory for the latter is discussed in [21].

In Paper I we extend this result to include a nontrivial volatility factor σ. We remark that, in contrast to the Brownian setting, the extension of Theorem 1.1.1 to Lévy semi-stationary processes is a more complex issue. This is due to the fact that it is harder to estimate various norms of X and related processes when the driving process L is a Lévy process. Our estimates on X rely heavily on decoupling techniques and isometries for stochastic integral mappings presented in the book of Kwapień and Woyczyński [61], which we will recall in Section 1.3. Moreover, we show functional convergence of the power variation – with respect to the Skorokhod M_1-topology in case (i) and uniform on compacts in probability in cases (ii) and (iii). See Section 1.2 for details.

In Paper II we consider more general variation functionals of the form (1.3) for continuous functions f. In this situation also three cases occur that are related to the three cases in Theorem 1.1.1. Which case applies depends not only on properties of the function f but also on the chosen normalising sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. In particular, for a fixed function f the variation functional (1.3) can converge to different limits for different normalising sequences. We also derive a second order limit theorem related to case (ii) when the function f is bounded. When $\alpha < k - 2/\beta$, a central limit theorem applies, and for $\alpha \in (k - 2/\beta, k - 1/\beta)$ we show convergence towards a $(k-\alpha)\beta$-stable random variable. This result relates to the second order asymptotic for the power variation for Lévy driven moving average processes derived in [20, Theorem 1.2].

1.2 Functional limit theorems and the Skorokhod M_1-topology

In this section we give preliminaries for the proof of functional convergence in Theorem 1.1.1 and its generalisations. In particular, we recall the notion of stable convergence and the definition and basic properties of the Skorokhod M_1-topology, which will be used in the functional version of Theorem 1.1.1 (i).

Theorem 1.1.1 shows the convergence of the sequence of real random variables $V(p;k)^n_t$, where $t \geq 0$ is fixed, under proper normalisation. More precisely, it only considers the case $t = 1$, but generalising it to arbitrary $t > 0$ is straightforward. However, the functionals $(V(p;k)^n_t)_{t\geq 0}$ and $(V(f;k)^n_t)_{t\geq 0}$ define stochastic processes with càdlàg sample paths, and it is natural to ask whether they converge as processes to a limiting process, i.e. whether the limit theorem holds functional. To this end we need to define notions of convergence for càdlàg processes, or equivalently, define metrics on the space $\mathbb{D} = \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàg functions from $\mathbb{R}_+$ into $\mathbb{R}$.

One mode of convergence on $\mathbb{D}$ is uniform convergence on compacts in probability, which will be denoted by $\xrightarrow{\text{u.c.p.}}$. For càdlàg stochastic processes $Z, Z^1, Z^2, \dots$ we have $Z^n \xrightarrow{\text{u.c.p.}} Z$ if for all $\varepsilon > 0$ and all $C > 0$ it holds that

$$\mathbb{P}\big(\|Z^n - Z\|_{C,\infty} > \varepsilon\big) \to 0, \quad \text{as } n \to \infty,$$

where $\|\cdot\|_{C,\infty}$ denotes the supremum norm on $[0, C]$. Equivalently, u.c.p.-convergence can be defined as convergence in probability of $\mathbb{D}$-valued random variables if $\mathbb{D}$ is equipped with a metric that metrises uniform convergence on compact sets, e.g. $d_{\text{uc}}(f, g) = \sum_{n=1}^{\infty} 2^{-n}(1 \wedge \|f - g\|_{[0,n],\infty})$. The following proposition is well known and the proof is straightforward, see [54, Equation (2.2.16)].

Proposition 1.2.1. *Let Z^n be a sequence of increasing processes in $\mathbb{D}(\mathbb{R}_+; \mathbb{R})$, such that $Z^n_t \xrightarrow{\mathbb{P}} Z_t$ for all t in a dense subset of $\mathbb{R}_+$. If the limiting process Z is continuous, it follows that $Z^n \xrightarrow{u.c.p.} Z$.*

This proposition indeed implies easily that the convergence in Theorem 1.1.1 (ii) and (iii) holds uniformly on compacts in probability, as was already remarked in [20]. See Theorem I.1.1 for details.

The situation is much more complicated in the framework of Theorem 1.1.1 (i), where the limit is not continuous and the convergence is stably in law. Let us briefly recall the definition of stable convergence, which was originally introduced in [72]. For a detailed treatment of the topic we refer to [49]. Consider a measurable space $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ and a Markov kernel $K : \Omega \times \mathcal{F} \to [0, 1]$, i.e. a mapping such that $K(\cdot, B) : \Omega \to [0, 1]$ is measurable for all $B \in \mathcal{F}$ and $K(\omega, \cdot)$ is a probability measure on $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ for all $\omega \in \Omega$. We obtain a probability measure $\mathbb{P}'$ on the measure space $(\Omega', \mathcal{F}') = (\Omega \times \widetilde{\Omega}, \mathcal{F} \otimes \widetilde{\mathcal{F}})$ by setting $\mathbb{P}'(d\omega, d\widetilde{\omega}) = K(\omega, d\widetilde{\omega})\mathbb{P}(d\omega)$. Random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ extend to $(\Omega', \mathcal{F}', \mathbb{P}')$ in the usual fashion, and we identify $\mathcal{F}$ with the sub σ-algebra $\mathcal{F} \otimes \{\emptyset, \widetilde{\Omega}\}$ of $\mathcal{F}'$. Let $(E, \mathcal{E})$ be a Polish space, i.e. a separable complete metric space, equipped with its Borel σ-algebra.

Definition 1.2.2. A sequence Z_n of E-valued random variables defined on $(\Omega, \mathcal{F})$ *converges $\mathcal{F}$-stably in law* to Z defined on the extension $(\Omega', \mathcal{F}')$, denoted $Z_n \xrightarrow{\mathcal{L}-s} Z$, if it satisfies one of the following two equivalent conditions.

(S1) For all real valued $\mathcal{F}$-measurable random variables $Y \in L^1(\Omega)$ and all bounded continuous functions $g : E \to \mathbb{R}$ it holds that

$$\mathbb{E}[g(Z_n)Y] \to \mathbb{E}'[g(Z)Y],$$

where $\mathbb{E}'$ denotes the expectation on the probability space $(\Omega', \mathcal{F}', \mathbb{P}')$.

(S2) For all $\mathcal{F}$-measurable random variables Y the joint convergence in law $(Z_n, Y) \xrightarrow{\mathcal{L}} (Z, Y)$ holds.

Clearly, stable convergence in law implies convergence in law. Conversely, it is implied by convergence in probability, i.e. $Z^n \xrightarrow{\mathbb{P}'} Z$ implies $Z^n \xrightarrow{\mathcal{L}-s} Z$, which follows easily from (S2). When both Z_n and Z are $\mathcal{F}$-measurable $Z_n \xrightarrow{\mathcal{L}-s} Z$, is equivalent to $Z_n \xrightarrow{\mathbb{P}} Z$. The main advantage of stable convergence over convergence in law is the following desirable property, see [49, Theorems 3.17, 3.18]. For sequences $(Z_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ of $\mathcal{F}$-measurable random variables with $Z_n \xrightarrow{\mathcal{L}-s} Z$ and $Y_n \xrightarrow{\mathbb{P}} Y$ it holds that $(Z_n, Y_n) \xrightarrow{\mathcal{L}} (Z, Y)$. This property is often useful for statistical applications, since in many frameworks it holds that $Z_n \xrightarrow{\mathcal{L}} Z$ and the limiting distribution depends on an unknown random quantity Y. The stable convergence $Z_n \xrightarrow{\mathcal{L}-s} Z$ then allows, roughly speaking, to replace Y by a consistent estimator.

In order to show functional stable convergence in Theorem 1.1.1 (i), we now need to choose a metric on $\mathbb{D}$. The metric d_{UC} of uniform convergence on compacts introduced above is not a good candidate, since the limiting process Z is not continuous and the approximating sequences $V(p;k)^n$ and $V(f;k)^n$ do not jump at the same times as Z. For characterising convergence of càdlàg functions to a discontinuous limit, Skorokhod [79] introduced 4 different topologies on the linear space $\mathbb{D}(\mathbb{R}_+;\mathbb{R})$, which are typically called the J_1-, J_2-, M_1- and M_2-topology, all of which can be given by a metric. The by far most popular one is the J_1-topology, which is also the strongest, i.e. convergence with respect to J_1 implies convergence with respect to the three other topologies. However, it can be shown that the convergence in Theorem 1.1.1 (i) does not hold functional with respect to the J_1-topology, see Appendix A.1. We prove that it holds with respect to the M_1-topology, which we introduce next. Some details to the other topologies are given in Appendix A.1.

In order to define the M_1-metric, we first consider a finite time horizon $0 < t_\infty < \infty$ and consider for a function $f \in \mathbb{D}([0,t_\infty];\mathbb{R})$ the completed graph, which is the subset of $\mathbb{R}^2$ obtained by 'filling in the jumps of f', i.e.

$$\Gamma_f = \{(t,x) \in [0,t_\infty] \times \mathbb{R} : x = \alpha f(t-) + (1-\alpha) f(t), \text{ for some } \alpha \in [0,1]\}.$$

For a visualisation of the functioning of the M_1-metric consider two functions $f, g \in \mathbb{D}([0,t_\infty];\mathbb{R})$, and imagine two ants positioned at the starting points of the completed graphs Γ_f and Γ_g, i.e. at the points $(0, f(0))$ and $(0, g(0))$ in $\mathbb{R}^2$. We now let the ants walk on the graphs, but forbid them to change directions, so they are only allowed to walk forward. If the two ants can find a way to walk the graphs to the end, i.e. to $(t_\infty, f(t_\infty))$ and $(t_\infty, g(t_\infty))$ respectively, without ever being further apart than ε (in $\mathbb{R}^2$), then it holds that the M_1-distance of f and g is smaller or equal ε (in $\mathbb{D}$). See Figure 1.2 for an example. More formally, a parametric representation of f is a continuous bijection $\phi : [0,1] \to \Gamma_f$ with $\phi(0) = (0, f(0))$. Denoting by $\Pi(f)$ the set of parametric representations of f, the M_1-metric is defined as

$$d_{M_1}(f_1, f_2) = \inf_{\substack{\phi_i \in \Pi(f_i)\\ i=1,2}} \{\|\phi_1 - \phi_2\|_\infty\},$$

where for a function $\phi : [0,1] \to \mathbb{R}^2$, $\phi(t) = (u(t), r(t))$ we denote $\|\phi\|_\infty := \sup_{t\in[0,1]}\{|u(t)| \vee |r(t)|\}$ (most ants prefer to measure distances in the maximum metric on $\mathbb{R}^2$). It is not

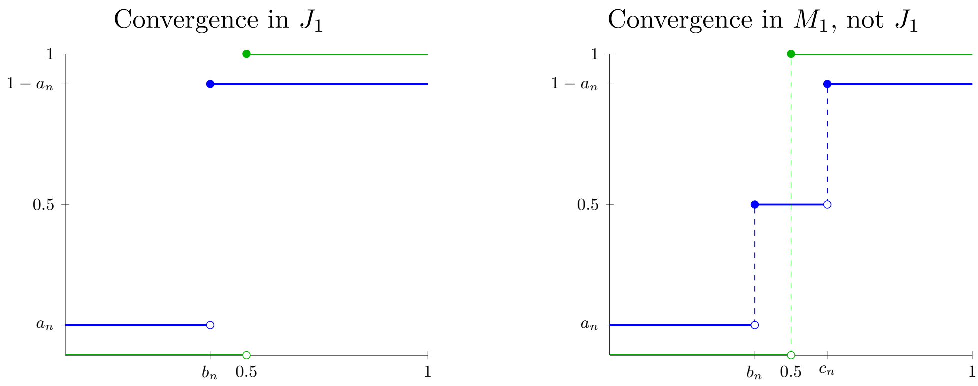

Figure 1.2: Examples for convergence in J_1 and M_1. The functions plotted in blue converge to the function $\mathbb{1}_{[1/2,1]}$, plotted in green if the sequences $(a_n), (b_n)$ and (c_n) are chosen such that $a_n \to 0$ and $b_n, c_n \to 1/2$. The first plot shows J_1-convergence (which implies M_1-convergence), in the second plot we have only M_1-convergence. The dashed lines show the completed graphs. See Appendix A.1 for definition and examples of J_2- and M_2-convergence.

difficult to show that d_{M_1} indeed defines a metric. The M_1-topology is weaker than J_1, i.e. every sequence that converges in J_1, converges in M_1 as well. A typical example for convergence in M_1 but not in J_1 is a monotonic staircase converging to a single jump, see Figure 1.2. Convergence with respect to M_1 can be generalised to $\mathbb{D}(\mathbb{R}_+;\mathbb{R})$ by defining $f_n \to f$ in $(\mathbb{D}(\mathbb{R}_+;\mathbb{R}), M_1)$ if and only if $f_n \to f$ in $(\mathbb{D}([0,t];\mathbb{R}), M_1)$ for all $t \geq 0$ such that f is continuous at t.

Since it is given by a metric, the M_1-topology can alternatively be defined by characterising convergence of sequences. This characterisation is often more convenient and will be used throughout our proofs. A sequence f_n of functions in $\mathbb{D}(\mathbb{R}_+;\mathbb{R})$ converges to $f \in \mathbb{D}(\mathbb{R}_+;\mathbb{R})$ with respect to the Skorokhod M_1-topology if and only if $f_n(t) \to f(t)$ for all t in a dense subset of $[0,\infty)$, and for all $t_\infty \in [0,\infty)$ it holds that

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{0\leq t\leq t_\infty} w(f_n, t, \delta) = 0.$$

Here, the oscillation function w is defined as

$$w(f,t,\delta) = \sup_{0\vee(t-\delta)\leq t_1<t_2<t_3\leq(t+\delta)\wedge t_\infty} \{|f(t_2) - [f(t_1), f(t_3)]|\},$$

where for $b < a$ the interval $[a,b]$ is defined to be $[b,a]$, and $|a-[b,c]| := \inf_{d\in[b,c]} |a-d|$. We remark that stochastic process convergence with respect to M_1, but not with respect to J_1, is a rare phenomenon in the literature, examples being [3, 60] and [86].

We conclude this subsection by sketching the typical approach to proving stable convergence of a sequence of processes Z^n in the M_1-topology, which will be denoted $Z^n \xrightarrow{\mathcal{L}_{M_1}-s} Z$. This technique is almost identically used to show convergence with respect to the J_1-topology and is discussed in detail in [27], see also [54, 87]. The key idea is that $Z^n \xrightarrow{\mathcal{L}_{M_1}-s} Z$ is equivalent to $(Z^n)_{n\in\mathbb{N}}$ satisfying the following two conditions.

(i) The sequence $(Z^n)_{n\in\mathbb{N}}$ is tight in $(\mathbb{D}(\mathbb{R}_+;\mathbb{R}), d_{M_1})$.

(ii) The finite dimensional distributions converge stably in law, i.e. for all $t_1, ..., t_d \geq 0$ we have the joint stable convergence in distribution of $\mathbb{R}^d$ valued random variables

$$(Z^n_{t_1}, \dots, Z^n_{t_d}) \xrightarrow{\mathcal{L}-s} (Z_{t_1}, \dots, Z_{t_d}).$$

Recall that a sequence of random variables $(Z^n)_{n\in\mathbb{N}}$ with values in a metric space $(E, \mathcal{E})$ is called tight if for all $\varepsilon > 0$ there is a compact set $K \subset E$ such that $\mathbb{P}(Z^n \in K) > 1 - \varepsilon$ for all n. The justification that it is sufficient to show (i) and (ii) above is the following corollary to Prokhorov's theorem.

Corollary 1.2.3. *([27, Theorem 5.1]) Let $(E, \mathcal{E})$ be a Polish space, and $(Z^n)_{n\in\mathbb{N}}$ be a sequence of $(E, \mathcal{E})$-valued random variables. Then $(Z^n)_{n\in\mathbb{N}}$ is tight if and only if every subsequence of $(Z^n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence. If moreover the limit of every weakly convergent subsequence of $(Z^n)_{n\in\mathbb{N}}$ must be Z, it follows already that Z^n converges in law to Z.*

We remark that the space $\mathbb{D}$ equipped with the M_1-topology is indeed Polish, see [87, Section 12.8]. The convergence of the finite dimensional distributions (ii) implies that the limit of every weakly convergent subsequence of $(Z^n)_{n\in\mathbb{N}}$ must be Z, see [27, Theorem 13.1] and [87, Theorem 11.6.6]. This argument is easily generalised to stable convergence in law by using (S2) of Definition 1.2.2.

1.3 Integration with respect to Lévy processes and Musielak-Orlicz spaces

In this section we give an overview of stochastic integration with respect to Lévy processes and infinitely divisible random measures, and present several estimates for Lévy integrals. When proving limit theorems for a Lévy driven process $Y_t = \int_{-\infty}^t F_{t,s} dL_s$, it is essential to have sharp control on the order of magnitude of increments

$$Y_{t+\Delta} - Y_t = \int_{-\infty}^{t+\Delta} (F_{t+\Delta,s} - F_{t,s}\mathbb{1}_{\{s\leq t\}}) dL_s, \qquad \text{as} \qquad \Delta \to 0.$$

Typically, it is much easier to control the order of magnitude of the integrand $F_{t+\Delta,s} - F_{t,s}\mathbb{1}_{\{s\leq t\}}$. Therefore, a crucial ingredient to our proofs are several isometries for the integral mapping $F \mapsto \int_{-\infty}^t F_{t,s} dL_s$ that we present below. These estimates were derived by Rajput and Rosiński [71] for deterministic integrands, and by Kwapién, Rosiński and Woyczyński [61, 75] for predictable integrands. They play a similar role for our proofs as Burkholder's inequality plays for proofs of limit theorems for processes driven by Brownian motion, e.g. continuous Itô semimartingales. In our framework, however, an application of Burkholder's inequality is not possible as the Lévy process does not necessarily have sufficiently high moments.

Consider a σ-finite measure space $(A, \mathcal{A}, \mu)$, for our purposes mostly $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, and let $\mathcal{A}_b$ denote the sets in $\mathcal{A}$ of finite measure. An independently scattered infinitely divisible random measure on A is a collection of real valued random variables $\{\Lambda(A) : A \in \mathcal{A}_b\}$ satisfying the following properties

1. For $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}_b$ with $\bigcup_n A_n \in \mathcal{A}_b$ it holds that $\Lambda(\bigcup_n A_n) = \sum_{n=1}^{\infty} \Lambda(A_n)$, almost surely.

2. For disjoint sets $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}_b$ the random variables $\{\Lambda(A_n)\}_{n\in\mathbb{N}}$ are independent.

3. For all $A \in \mathcal{A}_b$, the law of $\Lambda(A)$ is infinitely divisible.

A Lévy basis on $\mathbb{R}^d$, as used in the definition of ambit fields (1.1), is an independently scattered infinitely divisible random measure that is stationary in the sense that $\Lambda(A) \overset{\mathrm{d}}{=} \Lambda(A + x)$ for all $x \in \mathbb{R}^d$. A popular example of a Lévy basis is Gaussian white noise on $\mathbb{R}^d$, see Paper III. For simple functions $f : A \to \mathbb{R}$ of the form $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ where $\alpha_i \in \mathbb{R}$ and $A_i \in \mathcal{A}_b$, the stochastic integral is then defined as $\int_{A_0} f d\Lambda := \sum_{i=1}^{n} \alpha_i \Lambda(A_i \cap A_0)$ for any $A_0 \in \mathcal{A}$.

Taking limits in probability, the integral can be extended to the class $\mathbf{L}_{\mathrm{nr}}(d\Lambda)$ of all (nonrandom) functions $f : A \to \mathbb{R}$ such that there exists a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ with

(i) $f_n \to f$ μ-almost everywhere and

(ii) $\int_{A_0} f_n d\Lambda$ converges in probability for all $A_0 \in \mathcal{A}$.

For $f \in \mathbf{L}_{\mathrm{nr}}(d\Lambda)$, the integral $\int_{A_0} f d\Lambda$ is then defined as $\mathbb{P}\text{-}\lim_{n\to\infty} \int_{A_0} f_n d\Lambda$, which does not depend on the choice of the approximating sequence (f_n), as was demonstrated in [84]. In [71], the authors derived a more explicit equivalent definition for the class of integrands $\mathbf{L}_{\mathrm{nr}}(d\Lambda)$, see Theorem 1.3.2 below.

We are mostly interested in the case $(A, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ where λ denotes the Lebesgue measure and the independently scattered infinitely divisible random measure is generated by a pure jump symmetric Lévy process. More precisely, given a Lévy process L on the real line and letting $\Lambda((a, b]) = L_b - L_a$ for $a < b$, Λ extends uniquely to a random measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ by a standard argument, cf. [58, Theorem 3.4]. In this framework, the discussed integration theory can be extended to include predictable integrands. This extension relies on the use of decoupling inequalities and a complete account can be found in the monograph [61].

In order to derive continuity and isometry properties of the integral mapping, the space of integrands $\mathbf{L}_{\mathrm{nr}}(dL)$ (and certain subspaces) need to be equipped with topologies induced by normlike functionals called modulars. We recall now the definition and some basic properties of modulars. A detailed account can be found in [64]. Let us remark that in the literature there exist several slightly different definitions of modulars. We follow mostly [61] and [64].

Definition 1.3.1. Let E be a linear space over $\mathbb{R}$. A function $\Phi : E \to [0, \infty]$ is called a *modular* on E if it satisfies the following conditions

(i) $\Phi(e) = 0$ if and only if $e = 0$.

(ii) For any $e \in E$ the function $\mathbb{R} \to [0, \infty], t \mapsto \Phi(te)$ is continuous, even and nondecreasing on $\mathbb{R}_+$.

A modular is *of moderate growth* if it additionally satisfies

(iii) There is a finite constant C such that $\Phi(e+f) \le C(\Phi(e)+\Phi(f))$ for all $e, f \in E$.

It is 0-*convex* if it satisfies

(iv) For any $e, f \in E$ and $\alpha, \beta \ge 0$ with $\alpha + \beta = 1$ it holds that $\Phi(\alpha e + \beta f) \le \Phi(e) + \Phi(f)$.

A 0-convex modular of moderate growth defines a topology on E, which is determined by the condition that e_n converges to e if $\Phi(e_n - e) \to 0$. It is often convenient to work instead with either of the two following regularized modulars, both of which induce the same topology as Φ,

$$\|e\|_\Phi := \inf\{t > 0 \ : \ \Phi(e/t) \le 1\}, \quad \text{or} \quad |||e|||_\Phi := \inf\{t > 0 \ : \ \Phi(e/t) \le t\}.$$

The modular $\|\cdot\|_\Phi$ is of moderate growth and is homogeneous, i.e. it satisfies $\|te\|_\Phi = |t|\|e\|_\Phi$ for all $t \in \mathbb{R}$ and $e \in E$. It is not necessarily 0-convex and does not necessarily obey the triangle inequality. However, if Φ is convex, then $\|\cdot\|_\Phi$ is a norm, and is called the Luxemburg norm, see [64, Theorem 1.5]. This is used in Proposition 1.3.4 below. The modular $|||\cdot|||_\Phi$, on the other hand, is an F-norm, i.e. it obeys the triangle inequality and satisfies $|||-e|||_\Phi = |||e|||_\Phi$, but is not homogeneous. In particular, $d(e, f) = |||e - f|||_\Phi$ defines a metric on E. It can be shown that for a 0-convex modular Φ of moderate growth the conditions $d(e_n, e) \to 0$ and $\Phi(e_n - e) \to 0$ are equivalent, cf. [64, Theorem 1.6]. Since in metric spaces the topology is completely determined by characterising convergent sequences, this justifies our definition of the topology induced by Φ. For further details about the modulars $\|\cdot\|_\Phi$ and $|||\cdot|||_\Phi$ we refer to [63] and [61, Chapter 0.7].

Now let L be a pure jump symmetric Lévy process with Lévy measure ν. For $p \in [0, \infty)$ and measurable $f : \mathbb{R} \to \mathbb{R}$ define

$$\Phi_{p,L}(f) := \int_{\mathbb{R}^2} \phi_p(f(s)u)\, ds\, \nu(du), \quad \text{where} \quad \phi_p(x) := |x|^p \mathbb{1}_{\{|x|>1\}} + x^2 \mathbb{1}_{\{|x|\le 1\}} \tag{1.5}$$

Then, the functional $\Phi_{p,L}$ defines a modular on the space

$$\mathbf{L}^p_{\mathrm{nr}}(dL) := \{f : \mathbb{R} \to \mathbb{R} \text{ measurable} : \Phi_{p,L}(f) < \infty\}.$$

We show in Appendix A.2 that $\Phi_{p,L}$ is 0-convex and of moderate growth. The latter implies in particular that $\mathbf{L}^p_{\mathrm{nr}}(dL)$ is a vector space. For $p > 0$, this type of modular space is called Musielak-Orlicz space. It is complete with respect to the F-norm $|||\cdot|||_{p,L} := |||\cdot|||_{\Phi_{p,L}}$ and simple functions are dense in it, cf. [64]. We remark that for $p > 0$ the Lévy process L needs to admit pth moment in order for $\mathbf{L}^p_{\mathrm{nr}}(dL)$ to be nontrivial. The following theorem is a corollary to several results from [71, Section 2 & 3], the proof can be found in Appendix A.2.

Theorem 1.3.2. *(i) A function f is integrable with respect to the Lévy process L if and only if $f \in \mathbf{L}^0_{\mathrm{nr}}(dL)$, i.e. $\mathbf{L}^0_{\mathrm{nr}}(dL) = \mathbf{L}_{\mathrm{nr}}(dL)$. Moreover, for any $p > 0$ the integral $\int_{\mathbb{R}} f(s)dL_s$ is in $L^p(\Omega)$ if and only if $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$.*

(ii) Let $p > 0$ and equip $\mathbf{L}^p_{\mathrm{nr}}(dL)$ with the homogeneous modular $\|\cdot\|_{p,L} := \|\cdot\|_{\Phi_{p,L}}$ introduced above. Then, the integral mapping $\mathbf{L}^p_{\mathrm{nr}}(dL) \to L^p(\Omega)$, $f \mapsto \int_{\mathbb{R}} f(s)dL_s$ is a quasi-isometry, i.e. there are constants c, C, depending only on p, such that for all $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$

$$c\left\| \int_{\mathbb{R}} f(s)dL_s \right\|_p \leq \|f\|_{p,L} \leq C\left\| \int_{\mathbb{R}} f(s)dL_s \right\|_p.$$

Here and in the following we use for $p > 0$ and random variables Z the notation $\|Z\|_p = \mathbb{E}[|Z|^p]^{\frac{1}{p}}$, which defines a norm when $p \geq 1$ and a homogeneous modular when $p < 1$.

The following generalisation to predictable integrands is discussed in detail in [61]. A modular Φ on a linear metric space E defines by composition a mapping $\Phi : L^0(E) \to L^0([0,\infty])$, where $L^0(E)$ and $L^0([0,\infty])$ denote the spaces of E- and $[0,\infty]$-valued random variables, respectively. For $p \geq 0$, we define the random Musielak-Orlicz space

$$\mathbf{L}^p(dL) := \{F = (F_t)_{t\in\mathbb{R}} \in \mathcal{P} \,:\, \Phi_{p,L}(F) < \infty, \text{ almost surely}\},$$

where $\mathcal{P}$ denotes the class of predictable processes. The following result from [61] generalises Theorem 1.3.2 and will play a key role for our proofs.

Theorem 1.3.3. *A predictable process F is integrable with respect to L if and only if $F \in \mathbf{L}^0(dL)$. For all $p \geq 1$ there are constants c, C, depending only on p, such that for all $F \in \mathbf{L}^p(dL)$ it holds that*

$$c\mathbb{E}\big[\|F\|^p_{p,L}\big] \leq \mathbb{E}\left[\left| \int_{\mathbb{R}} F_s \, dL_s \right|^p\right] \leq C\mathbb{E}\big[\|F\|^p_{p,L}\big].$$

This result follows from [61, Theorem 9.1.1], [61, Equation (9.5.3)] and the comments following it. The restriction $p \geq 1$ is inherent to the decoupling inequality used in [61]. In general the modulars $\|\cdot\|_{p,L}$ are much better behaved for $p \geq 1$, as they are equivalent to a norm in this case. This fact will also be essential for some of our proofs.

Proposition 1.3.4. *Assume $p \geq 1$. There is a norm $\|\cdot\|'_{p,L}$ on $\mathbf{L}^p_{\mathrm{nr}}(dL)$, called the Luxemburg norm, and constants $c, C > 0$ such that*

$$c\|f\|'_{p,L} \leq \|f\|_{p,L} \leq C\|f\|'_{p,L}$$

for all $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$. The modular $\|\cdot\|_{p,L}$ has the following properties

(i) Homogeneity: For all $\lambda \in \mathbb{R}$, $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$, $\|\lambda f\|_{p,L} = |\lambda|\|f\|_{p,L}$.

(ii) Triangle inequality (up to a constant): There exists a constant $C > 0$ such that, for all $m \geq 1$ and $f_1, ..., f_m \in \mathbf{L}^p_{\mathrm{nr}}(dL)$ we have

$$\|f_1 + \cdots + f_m\|_{p,L} \leq C\big(\|f_1\|_{p,L} + \cdots + \|f_m\|_{p,L}\big).$$

(iii) Upper bound: For all $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$ it holds that

$$\|f\|_{p,L} \leq \Phi^{1/2}_{p,L}(f) \vee \Phi^{1/p}_{p,L}(f).$$

The proof is given in Appendix A.2. Properties ((i))-((iii)) obviously continue to hold, ω by ω, for processes in $\mathbf{L}^p(dL)$. Fortunately, the restriction $p \geq 1$ becomes unnecessary when the driving Lévy process L is symmetric β-stable, as we can rely on an isometry derived in [75]. We use the notation $\|Z\|^\beta_{\beta,\infty} = \sup_{\lambda>0} \lambda^\beta \mathbb{P}[|Z| > \lambda]$ for an arbitrary random variable Z. In the literature, $\|\cdot\|_{\beta,\infty}$ is often referred to as the weak L^β-norm, even though it satisfies the triangle inequality only up to a constant. For $p < \beta$ it holds that $\|Z\|_p \leq \|Z\|_{\beta,\infty} \leq (\frac{\beta}{\beta-p})^{1/p}\|Z\|_\beta$. In particular, we can have $\|Z\|_{\beta,\infty} < \infty$ even though $\|Z\|_\beta = \infty$, which is for example the case when Z is β-stable.

Theorem 1.3.5 ([75], Theorem 2.1)**.** *Let $(L_t)_{t\in\mathbb{R}}$ be a symmetric β-stable Lévy process. Then there are positive constants $c, C > 0$ such that for all F in $\mathbf{L}^0(dL)$ it holds that*

$$c\mathbb{E}\bigg[\int_{\mathbb{R}} |F_s|^\beta \, ds\bigg] \leq \bigg\| \int_{\mathbb{R}} F_s \, dL_s \bigg\|^\beta_{\beta,\infty} \leq C\mathbb{E}\bigg[\int_{\mathbb{R}} |F_s|^\beta \, ds\bigg].$$

We remark that Theorem 1.3.3 and 1.3.5 consider in the original references only integrals over a finite time interval, say $\int_0^t F_s \, dL_s$. However, the definition of the stochastic integral and the estimates of the integral extend to the case of $\int_{\mathbb{R}} F_s \, dL_s$ in a natural way.

The theory of Lévy integration developed in [61] and [71] is not restricted to symmetric Lévy processes. For non-symmetric Lévy processes, however, the corresponding modulars become more involved and are much harder to control. As an example, Assumption (A) is no longer sufficient to guarantee the existence of the integral (1.2) with $\sigma = 1$ when the Lévy process is non-symmetric, which is easily seen by considering a pure drift process L. In Section I.3 of Paper I we present an estimate for integrals with respect to non-symmetric Lévy processes that we use in the proof of Theorem I.1.1.

Finally, let us remark that a general approach to define stochastic space-time integrals with random integrand as in (1.1) dates back to Bichteler [26] and constructs the stochastic integral by the Daniell procedure. In the recent publication [31] the authors derive an explicit characterisation of the class of possible integrands for this integration theory, which coincides with the class $\mathbf{L}(dL)$ when applied to the framework discussed above. This general integration theory can in particular be used to show the existence of general tempo-spatial ambit fields with stochastic integrand.

1.4 Methodology of the proofs

The proofs of the generalisations of Theorem 1.1.1 in the articles below contain many technical details, sometimes making it difficult to grasp the general idea behind them. It adds to this effect that some of the essential steps of the proof of Theorem 1.1.1 given in [20] can be transferred to the generalisations presented in Paper I and II in a straightforward manner, and are then referenced rather than repeated. In this section we explain therefore the intuition and methodology of the proof of Theorem 1.1.1 and discuss some aspects of the generalisation to include nontrivial volatility and to general variation functionals $V(f;k)_t^n$. We motivate how the limits and convergence rates emerge, prioritising simplicity over mathematical preciseness. Throughout this section we denote by X_t the model (1.2), and by Y_t the same model with $\sigma \equiv 1$. For simplicity of exposition we only consider the case $k = 1$ and we set $\Delta_i^n X := \Delta_{i,1}^n X$ and $h := h_1$. By $V(p;X)_t^n$ and $V(p;Y)_t^n$ we denote the realised power variation of the processes X and Y, respectively, and similarly $V(f;Y)_t^n$ denotes the general variation functional introduced in (1.3).

Theorem 1.1.1 (i)

Let us first remark that the limit in Theorem 1.1.1 (i) is indeed finite almost surely by the following argument. By mean value theorem there is a constant $C > 0$ such that $|h(x)| \le C|x|^{\alpha-1}$ for all $x \in \mathbb{R}$, implying that $|V_m| \le C\big(|U_m|^{\alpha p} + \sum_{l=1}^{\infty} |l+U_m|^{(\alpha-1)p}\big)$. Since $(\alpha-1)p < -1$ by assumption, the random variables V_m are uniformly bounded. It follows that $V(p;Y) \le C\sum_{m\,:\,T_m\in[0,1]} |\Delta L_{T_m}|^p$, which is finite almost surely by the assumption $p > \beta$.

Now, let us recall the basic intuition behind the proof of Theorem 1.1.1 (i). We first discuss the asymptotic distribution of the increments

$$\Delta_i^n Y = \int_{-\infty}^{i/n} g\Big(\frac{i}{n} - s\Big) - g\Big(\frac{i-1}{n} - s\Big) dL_s.$$

In the situation of Theorem 1.1.1 (i) it holds that $\alpha < 1 - 1/p$, implying that the derivative g' explodes at 0. This explosive behaviour dominates the asymptotics of the increments, and justifies the approximation

$$\Delta_i^n Y \approx \int_{\frac{i}{n}-1}^{i/n} g\Big(\frac{i}{n} - s\Big) - g\Big(\frac{i-1}{n} - s\Big) dL_s.$$

Although the process L typically has infinitely many jumps on finite intervals, we assume for simplicity of exposition that $T \in [(j-1)/n, j/n)$ is the only jump time of L within the interval $[-1, t]$. Recalling the assumption $g(t) \sim c_0 t^\alpha$ for $t \to 0$, we consider the approximation

$$\begin{aligned}\Delta_i^n Y &\approx A_i^n + B_i^n \\ &:= c_0 \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \Big(\frac{i}{n} - s\Big)^\alpha dL_s + \int_{\frac{i}{n}-1}^{\frac{i-1}{n}} \Big\{ \Big(\frac{i}{n} - s\Big)^\alpha - \Big(\frac{i-1}{n} - s\Big)^\alpha \Big\} dL_s \right)\end{aligned}$$

Since $T \in [(j-1)/n, j/n)$ is the only jump time of L, we observe that $A_i^n = 0$ for all $i \neq j$ and $B_i^n = 0$ for all $i < j$. More precisely, we deduce that

$$\Delta_{j+l}^n Y \approx \begin{cases} c_0 \Delta L_T \left(\frac{j}{n} - T\right)^\alpha & l = 0 \\ c_0 \Delta L_T \left(\left(\frac{j+l}{n} - T\right)^\alpha - \left(\frac{j+l-1}{n} - T\right)^\alpha\right) & l \geq 1 \end{cases} \tag{1.6}$$

Now, we use the following result, which is essentially due to Tukey [83] (see also [38] and [20, Lemma 4.1]): Let Z be a random variable with an absolutely continuous distribution and let $\{x\} := x - [x] \in [0, 1)$ denote the fractional part of $x \in \mathbb{R}$. Then it holds that

$$\{nZ\} \xrightarrow{\mathcal{L}-s} U \sim \mathcal{U}([0,1]),$$

where U is defined on an extension of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and is independent of $\mathcal{F}$. Using $j - nT = 1 - \{nT\}$, the approximation (1.6) now implies the stable convergence of scaled increments

$$n^\alpha \Delta_{j+l}^n Y \xrightarrow{\mathcal{L}-s} c_0 \Delta L_T \left((l+U)_+^\alpha - (l-1+U)_+^\alpha\right), \qquad l \geq 0. \tag{1.7}$$

Thus, we obtain the result of (1.4) for one jump time:

$$\sum_{i=j}^{[nt]} |n^\alpha \Delta_i^n Y|^p \xrightarrow{\mathcal{L}-s} c_0^p |\Delta L_T|^p \sum_{l=0}^{\infty} \left|(l+U)_+^\alpha - (l-1+U)_+^\alpha\right|^p. \tag{1.8}$$

In Paper I, where we extend the model to contain a nontrivial volatility factor σ, the formal proof becomes more complicated, but the intuition remains largely the same. We can follow essentially the same argument as above, replacing dL_s by $\sigma_{s-} dL_s$ and ΔL_T by $\sigma_{T-} \Delta L_T$. This leads us to presume that the limit of the power variation in the 1 jump scenario above is

$$V(p; X)_t^n \xrightarrow{\mathcal{L}-s} c_0^p |\Delta L_T \sigma_{T-}|^p \sum_{l=0}^{\infty} \left|(l+U)_+^\alpha - (l-1+U)_+^\alpha\right|^p.$$

This intuition proves to be correct, as we will show in Theorem I.1.1.

In Paper II we consider the more general variation functional introduced in (1.3) assuming that σ is constant. The intuitive approximations (1.7) and (1.8) above show that the appropriate choice for the normalising sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ is $a_n = 1$ and $b_n = n^\alpha$. Arguing as above, we then expect the stable convergence in law

$$V(f; Y)_t^n \xrightarrow{\mathcal{L}-s} \sum_{l=0}^{\infty} f\left\{c_0 \Delta L_T \left((l+U)_+^\alpha - (l-1+U)_+^\alpha\right)\right\}.$$

The function f needs to satisfy a certain growth condition to ensure that the limit is finite. See Theorem II.1.1 for details.

Theorem 1.1.1 (ii)

Here we present the intuition behind the proof of Theorem 1.1.1 (ii). We first turn our attention to the small scale behaviour of the stationary increments Lévy driven moving averages Y. Under the conditions of Theorem 1.1.1 (ii), $\alpha < 1-1/\beta$ and thus g' explodes at 0. Hence, we intuitively deduce the following approximation for the increments of Y for a small $\Delta > 0$:

$$\begin{aligned}
Y_{t+\Delta} - Y_t &= \int_{-\infty}^{t+\Delta} \{g(t+\Delta-s) - g(t-s)\}\, dL_s \\
&\approx \int_{t+\Delta-\varepsilon}^{t+\Delta} \{g(t+\Delta-s) - g(t-s)\}\, dL_s \\
&\approx c_0 \int_{t+\Delta-\varepsilon}^{t+\Delta} \{(t+\Delta-s)_+^{\alpha} - (t-s)_+^{\alpha}\}\, dL_s \\
&\approx c_0 \int_{-\infty}^{t+\Delta} \{(t+\Delta-s)_+^{\alpha} - (t-s)_+^{\alpha}\}\, dL_s = \widetilde{Y}_{t+\Delta} - \widetilde{Y}_t,
\end{aligned}$$

where

$$\widetilde{Y}_t := c_0 \int_{\mathbb{R}} \{(t-s)_+^{\alpha} - (-s)_+^{\alpha}\}\, dL_s,$$

and $\varepsilon > 0$ is an arbitrary small real number with $\varepsilon \gg \Delta$. In the classical terminology $\widetilde{Y}$ is called the *tangent process of* Y. In the framework of Theorem 1.1.1 (ii), the process $\widetilde{Y}$ is a symmetric fractional β-stable motion. We recall that $(\widetilde{Y}_t)_{t\geq 0}$ has stationary increments and is self-similar with index $H = \alpha + 1/\beta \in (1/2, 1)$, i.e.

$$(\widetilde{Y}_{at})_{t\geq 0} \overset{d}{=} a^H (\widetilde{Y}_t)_{t\geq 0}.$$

Furthermore, the symmetric fractional β-stable noise $(\widetilde{Y}_t - \widetilde{Y}_{t-1})_{t\geq 1}$ is mixing; see e.g. [30]. Thus, using Birkhoff's ergodic theorem we conclude that

$$\begin{aligned}
V(p;Y)_t^n &= \frac{1}{n} \sum_{i=1}^{[nt]} \left| n^H \Delta_i^n Y \right|^p \\
&\approx \frac{1}{n} \sum_{i=1}^{[nt]} \left| n^H \Delta_i^n \widetilde{Y} \right|^p \\
&\overset{d}{=} \frac{1}{n} \sum_{i=1}^{[nt]} \left| \widetilde{Y}_i - \widetilde{Y}_{i-1} \right| \xrightarrow{\mathbb{P}} t\mathbb{E}[|\widetilde{Y}_1 - \widetilde{Y}_0|^p] = t m_p,
\end{aligned}$$

where m_p is the constant defined in 1.1.1 (ii).

For the generalised variation functional (1.3) we expect by the same arguments the convergence in probability

$$V(f;Y)_t^n \xrightarrow{\mathbb{P}} t\mathbb{E}[f(\widetilde{Y}_1 - \widetilde{Y}_0)],$$

with the scaling sequences $a_n = n^{-1}$ and $b_n = n^H$, provided f is such that the expectation exists.

In our first paper we derive the convergence of the realised power variation of the process X with nontrivial volatility by the following blocking technique. In the first step of the proof we freeze σ over blocks of length $1/n$ and replace the power variation by the functional

$$\widetilde{V}(p;X)_t^n = \sum_{i=1}^{[nt]} |\sigma_{\frac{i-1}{n}} \Delta_i^n Y|^p.$$

This replacement is justified by the asymptotic equivalence

$$\left| n^{p(\alpha+1/\beta)-1} (\widetilde{V}(p;X)_t^n - V(p;X)_t^n) \right| \xrightarrow{\mathbb{P}} 0, \tag{1.9}$$

which we derive in the proof. Thereafter, we introduce a new step size $1/l$ satisfying $1/n \ll 1/l \ll 1$, and freeze the volatility at the beginning of each blocks of length $1/l$. More precisely, we consider the functional

$$\widetilde{V}(p;X)_t^{n,l} = \sum_{j=1}^{[tl]} |\sigma_{\frac{j-1}{l}}|^p \Bigg(\sum_{i \,:\, \frac{i}{n} \in \left[\frac{j-1}{l}, \frac{j}{l}\right)} |\Delta_i^n Y|^p \Bigg).$$

Thereafter , we establish asymptotic equivalence of $\widetilde{V}(p;X)_t^{n,l}$ and $\widetilde{V}(p;X)_t^n$ by showing that

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}(|n^{p(\alpha+1/\beta)-1}(\widetilde{V}(p;X)_t^{l,n} - \widetilde{V}(p;X)_t^n)| > \varepsilon) = 0,$$

for all $\varepsilon > 0$. Then, applying the limit theorem for the process Y on each block of size $1/l$, we obtain

$$n^{p(\alpha+1/\beta)-1}\widetilde{V}(p;X)_t^{l,n} \quad \xrightarrow[n\to\infty]{\mathbb{P}} \quad \sum_{j=1}^{[lt]} |\sigma_{\frac{j-1}{l}}|^p \frac{m_p}{l} \quad \xrightarrow[l\to\infty]{\text{a.s.}} \quad m_p \int_0^t |\sigma_t|^p dt,$$

where the second step is convergence of Riemann sums. The integral on the right hand side is indeed the limit in Theorem I.1.1 (ii).

Remark 1.4.1. The approach of freezing σ over blocks of different sizes is quite popular for extending limit theorems to volatility modulated processes, and has for example been used for Itô semimartingales [12] and Brownian semi-stationary processes [9], i.e. the process X driven by a Brownian motion. It is therefore remarkable that this technique is not applicable in the proofs of Theorem I.1.1 (i) and (iii) for the following reason. The fundamental idea behind the blocking technique is the approximation

$$\begin{aligned} X_{t+\Delta} - X_t &= \int_{-\infty}^{t+\Delta} \{g(t+\Delta-s) - g(t-s)\}\sigma_{s-} dL_s \\ &\approx \sigma_{t-} \int_{-\infty}^{t+\Delta} g(t+\Delta-s) - g(t-s) dL_s, \end{aligned}$$

for $\Delta > 0$ small. This approximation is justified if the integrand gains asymptotically most weight around t, which is the case when α is small and g' explodes at 0. Consequently, the blocking technique must fail in the framework of Theorem 1.1.1 (iii), where we consider large α.

In the framework of Theorem I.1.1 (i) we assume $\alpha < 1 - 1/p$ and it is therefore somewhat surprising that the blocking technique is not applicable either. Considering the one jump scenario and the notation of the last subsection, (1.6) yields the approximation

$$\Delta_{j+l}^n X \approx c_0 \sigma_{T-} \Delta L_T \left(\left(\frac{j+l}{n} - T \right)^{\alpha} - \left(\frac{j+l-1}{n} - T \right)^{\alpha} \right) \approx \sigma_{T-} \Delta_i^n Y,$$

for $l \geq 1$. The first step of the blocking technique, however, approximates the increment $\Delta_{j+l}^n X$ by $\sigma_{\frac{j+l-1}{n}} \Delta_{j+l}^n Y \approx \sigma_T \Delta_{j+l}^n Y$, leading to a different result if σ and L jump at the same time. Consequently, the asymptotic equivalence (1.9), properly scaled, does not hold.

Theorem 1.1.1 (iii)

In order to uncover the path properties of the process Y we perform a formal differentiation with respect to time. Since $g(0) = 0$ we obtain a formal representation

$$dY_t = g(0) dL_t + \left(\int_{-\infty}^{t} g'(t-s)\, dL_s \right) dt = F_t\, dt. \tag{1.10}$$

We remark that the random variable F_t is not necessarily finite under assumption (A). However, under conditions of Theorem 1.1.1 (iii), the process Y is differentiable almost everywhere and $Y' = F \in L^p([0,1])$, although the process F explodes at jump times of L when $\alpha < 1$. Thus, under the conditions of Theorem 1.1.1 (iii), an application of the mean value theorem gives an intuitive proof of (iii):

$$\mathbb{P}\text{-}\lim_{n\to\infty} V(p;Y)_t^n = \mathbb{P}\text{-}\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{[nt]} |F_{\xi_i^n}|^p = \int_0^t |F_u|^p\, du,$$

where $\xi_i^n \in ((i-1)/n, i/n)$. This gives a sketch of the proof of the asymptotic result in Theorem 1.1.1 (iii).

For extending the result to the process X the intuition remains largely the same, with the process F_t replaced by $U_t = \int_{-\infty}^t g'(t-s)\sigma_s dL_s$. We do not show that the sample paths of X are differentiable with derivative U but derive a stochastic Fubini result for Lévy integrals to formalize the idea behind (1.10).

For the variation functional $V(f;Y)_t^n$ the arguments above show that with the normalizing sequences $a_n = n^{-1}$ and $b_n = n$ we can expect

$$V(f;Y)_t^n \xrightarrow{\mathbb{P}} \int_0^t f(F_u) du,$$

when the function f is such that the integral exists.

Bibliography

[1] Abramowitz, M. and I. Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. U.S. Government Printing Office, Washington, D.C.

[2] Adler, R. J. (1981). *The geometry of random fields*. John Wiley & Sons, Ltd., Chichester. Wiley Series in Probability and Mathematical Statistics.

[3] Avram, F. and M. Taqqu (1992). Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab. 20*(1), 483–503.

[4] Ayache, A. and J. Hamonier (2012). Linear fractional stable motion: A wavelet estimator of the parameter. *Statistics & Probability Letters 82*(8), 1569 – 1575.

[5] Barndorff-Nielsen, O. and A. Basse-O'Connor (2011). Quasi Ornstein-Uhlenbeck processes. *Bernoulli 17*(3), 916–941.

[6] Barndorff-Nielsen, O., F. Benth, and A. Veraart (2011). Modelling electricity forward markets by ambit fields. *Available at https://ssrn.com/abstract=1938704*.

[7] Barndorff-Nielsen, O., F. Benth, and A. Veraart (2015). Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency. In *Advances in mathematics of finance*, pp. 25–60. Polish Acad. Sci. Inst. Math., Warsaw.

[8] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2009). Power variation for Gaussian processes with stationary increments. *Stochastic Process. Appl. 119*(6), 1845–1865.

[9] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2011). Multipower variation for Brownian semistationary processes. *Bernoulli 17*(4), 1159–1194.

[10] Barndorff-Nielsen, O., J. Corcuera, M. Podolskij, and J. Woerner (2009). Bipower variation for Gaussian processes with stationary increments. *J. Appl. Probab. 46*(1), 132–150.

[11] Barndorff-Nielsen, O., J. M. Corcuera, and M. Podolskij (2013). Limit theorems for functionals of higher order differences of Brownian semi-stationary processes. In *Prokhorov and contemporary probability theory*, Volume 33 of *Springer Proc. Math. Stat.*, pp. 69–96. Springer, Heidelberg.

[12] Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From stochastic calculus to mathematical finance*, pp. 33–68. Springer, Berlin.

[13] Barndorff-Nielsen, O., E. Hedevang, J. Schmiegel, and B. Szozda (2016). Some recent developments in ambit stochastics. In *Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015*, Volume 138, pp. 3–25. Springer, Cham.

[14] Barndorff-Nielsen, O., M. Pakkanen, and J. Schmiegel (2014). Assessing relative volatility/intermittency/energy dissipation. *Electron. J. Stat. 8*(2), 1996–2021.

[15] Barndorff-Nielsen, O. and J. Schmiegel (2007). Ambit processes; with applications to turbulence and tumour growth. In *Stochastic analysis and applications*, pp. 93–124. Springer.

[16] Barndorff-Nielsen, O. and J. Schmiegel (2008). Time change, volatility, and turbulence. In *Mathematical control theory and finance*, pp. 29–53. Springer, Berlin.

[17] Barndorff-Nielsen, O. and J. Schmiegel (2009). Brownian semistationary processes and volatility/intermittency. In *Advanced financial modelling*, Volume 8 of *Radon Ser. Comput. Appl. Math.*, pp. 1–25. Walter de Gruyter, Berlin.

[18] Barndorff-Nielsen, O. and N. Shephard (2003). Realized power variation and stochastic volatility models. *Bernoulli 9*(2), 243–265.

[19] Basse-O'Connor, A., C. Heinrich, and M. Podolskij (2017). On limit theory for Lévy semi-stationary processes. available at arXiv:1604.02307.

[20] Basse-O'Connor, A., R. Lachiéze-Rey, and M. Podolskij (2016). Power variation for a class of stationary increments levy driven moving averages. *Annals of Probability*. To appear.

[21] Basse-O'Connor, A. and M. Podolskij (2017). On critical cases in limit theory for stationary increments Lévy driven moving averages. *Stochastics 89*(1), 360–383.

[22] Basse-O'Connor, A. and J. Rosiński (2016). On infinitely divisible semimartingales. *Probab. Theory Related Fields 164*(1-2), 133–163.

[23] Benassi, A., S. Cohen, and J. Istas (2004). On roughness indices for fractional fields. *Bernoulli 10*(2), 357–373.

[24] Bennedsen, M., A. Lunde, and M. Pakkanen (2016). Hybrid scheme for brownian semistationary processes. available at arXiv:1507.03004.

[25] Berk, K. (1973). A central limit theorem for m-dependent random variables with unbounded m. *Ann. Probability 1*, 352–354.

[26] Bichteler, K. (2002). *Stochastic integration with jumps.* Cambridge University Press, Cambridge.

[27] Billingsley, P. (1999). *Convergence of probability measures* (Second ed.). John Wiley & Sons, Inc., New York.

[28] Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989). *Regular variation.* Cambridge University Press, Cambridge.

[29] Braverman, M. and G. Samorodnitsky (1998). Symmetric infinitely divisible processes with sample paths in Orlicz spaces and absolute continuity of infinitely divisible processes. *Stochastic Process. Appl. 78*(1), 1–26.

[30] Cambanis, S., C. Hardin, Jr., and A. Weron (1987). Ergodic properties of stationary stable processes. *Stochastic Process. Appl. 24*(1), 1–18.

[31] Chong, C. and C. Klüppelberg (2015). Integrability conditions for space-time stochastic integrals: theory and applications. *Bernoulli 21*(4), 2190–2216.

[32] Chronopoulou, A., C. Tudor, and F. Viens (2011). Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes. *Commun. Stoch. Anal. 5*(1), 161–185.

[33] Chronopoulou, A., F. Viens, and C. Tudor (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electron. J. Stat. 3*, 1393–1435.

[34] Coeurjolly, J. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process. 4*(2), 199–227.

[35] Cohen, S., C. Lacaux, and M. Ledoux (2008). A general framework for simulation of fractional fields. *Stochastic Process. Appl. 118*(9), 1489–1517.

[36] Dang, T. and J. Istas (2015). Estimation of the hurst and the stability indices of a h-self-similar stable process. Working paper. Available at arXiv:1506.05593.

[37] Davies, S. and P. Hall (1999). Fractal analysis of surface roughness by using spatial data. *J. R. Stat. Soc. Ser. B Stat. Methodol. 61*(1), 3–37.

[38] Delattre, S. and J. Jacod (1997). A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli 3*(1), 1–28.

[39] Gärtner, K. and M. Podolskij (2015). On non-standard limits of Brownian semi-stationary processes. *Stochastic Process. Appl. 125*(2), 653–677.

[40] Glaser, S. (2015). A law of large numbers for the power variation of fractional Lévy processes. *Stoch. Anal. Appl. 33*(1), 1–20.

[41] Gneiting, T., H. Ševčíková, and D. Percival (2012). Estimators of fractal dimension: assessing the roughness of time series and spatial data. *Statist. Sci. 27*(2), 247–277.

[42] Gneiting, T., H. Ševčíková, D. Percival, M. Schlather, and Y. Jiang (2006). Fast and exact simulation of large Gaussian lattice systems in $\mathbf{R}^2$: exploring the limits. *J. Comput. Graph. Statist. 15*(3), 483–501.

[43] Goff, J. and T. Jordan (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research: Solid Earth 93*(B11), 13589–13608.

[44] Grahovac, D., N. Leonenko, and M. Taqqu (2015). Scaling properties of the empirical structure function of linear fractional stable motion and estimation of its parameters. *J. Stat. Phys. 158*(1), 105–119.

[45] Guttorp, P. and T. Gneiting (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika 93*(4), 989–995.

[46] Guyon, X. and J. León (1989). Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbf{R}$. *Ann. Inst. H. Poincaré Probab. Statist. 25*(3), 265–282.

[47] Hansen, L. and T. Thorarinsdottir (2013). A note on moving average models for Gaussian random fields. *Statist. Probab. Lett. 83*(3), 850–855.

[48] Hansen, L., T. Thorarinsdottir, E. Ovcharov, T. Gneiting, and D. Richards (2015). Gaussian random particles with flexible Hausdorff dimension. *Adv. in Appl. Probab. 47*(2), 307–327.

[49] Häusler, E. and H. Luschgy (2015). *Stable convergence and stable limit theorems.* Springer, Cham.

[50] Hellmund, G., M. Prokešová, and E. Jensen (2008). Lévy-based Cox point processes. *Adv. in Appl. Probab. 40*(3), 603–629.

[51] Ho, H. and T. Hsing (1997). Limit theorems for functionals of moving averages. *Ann. Probab. 25*(4), 1636–1669.

[52] Huang, W., K. Wang, F. Breidt, and R. Davis (2011). A class of stochastic volatility models for environmental applications. *J. Time Series Anal. 32*(4), 364–377.

[53] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic processes and their applications 118*(4), 517–559.

[54] Jacod, J. and P. Protter (2012). *Discretization of processes.* Springer, Heidelberg.

[55] Jacod, J. and A. Shiryaev (2003). *Limit theorems for stochastic processes* (Second ed.). Springer-Verlag, Berlin.

[56] Jensen, E., K. Jonsdottir, J. Schmiegel, and O. Barndorff-Nielsen (2006). Spatio-temporal modelling-with a view to biological growth. *Monographs on statistics and applied probability 107*, 47.

[57] Jónsdóttir, K., A. Rønn-Nielsen, K. Mouridsen, and E. Jensen (2013). Lévy-based modelling in brain imaging. *Scand. J. Stat. 40*(3), 511–529.

[58] Kallenberg, O. (1983). *Random measures* (Third ed.). Akademie-Verlag, Berlin; Academic Press, Inc., London.

[59] Kallenberg, O. (2002). *Foundations of modern probability* (Second ed.). Springer-Verlag, New York Berlin Heidelberg.

[60] Kella, O. and W. Whitt (1990). Diffusion approximations for queues with server vacations. *Adv. in Appl. Probab. 22*(3), 706–729.

[61] Kwapień, S. and W. Woyczyński (1992). *Random series and stochastic integrals: single and multiple*. Birkhäuser Boston, Inc., Boston, MA.

[62] Matérn, B. (1986). *Spatial variation* (Second ed.). Springer-Verlag, Berlin.

[63] Mazur, S. and W. Orlicz (1958). On some classes of linear spaces. *Studia Math. 17*, 97–119.

[64] Musielak, J. (1983). *Orlicz spaces and modular spaces*. Springer-Verlag, Berlin.

[65] Nguyen, M. and A. Veraart (2017). Modelling spatial heteroskedasticity by volatility modulated moving averages. *Spatial Statistics*.

[66] Nourdin, I. and A. Réveillac (2009). Asymptotic behavior of weighted quadratic variations of fractional Brownian motion: the critical case $H = 1/4$. *Ann. Probab. 37*(6), 2200–2230.

[67] Noven, R., A. D. Veraart, and A. Gandy (2015). A levy-driven rainfall model with applications to futures pricing. *available at arXiv: 1511.08190*.

[68] Pipiras, V. and M. Taqqu (2003). Central limit theorems for partial sums of bounded functionals of infinite-variance moving averages. *Bernoulli 9*(5), 833–855.

[69] Podolskij, M. (2015). Ambit fields: survey and new challenges. In *XI Symposium on Probability and Stochastic Processes*, Volume 69 of *Progr. Probab.*, pp. 241–279. Birkhäuser/Springer.

[70] Podolskij, M. and M. Vetter (2010). Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl. 64*(3), 329–351.

[71] Rajput, B. and J. Rosiński (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields 82*(3), 451–487.

[72] Rényi, A. (1963). On stable sequences of events. *Sankhyā Ser. A 25*, 293 302.

[73] Rolewicz, S. (1985). *Metric linear spaces* (Second ed.). D. Reidel Publishing Co., Dordrecht; PWN—Polish Scientific Publishers, Warsaw.

[74] Rosiński, J. (1990). On series representations of infinitely divisible random vectors. *Ann. Probab. 18*(1), 405–430.

[75] Rosiński, J. and W. Woyczyński (1986). On Itô stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab. 14*(1), 271–286.

[76] Samorodnitsky, G. and M. Taqqu (1994). *Stable non-Gaussian random processes.* Chapman & Hall, New York.

[77] Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge.

[78] Sato, K. (2013). *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge.

[79] Skorohod, A. (1956). Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen. 1*, 289–319.

[80] Surgailis, D. (2004). Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli 10*(2), 327–355.

[81] Taqqu, M. and R. Wolpert (1983). Infinite variance self-similar processes subordinate to a Poisson measure. *Z. Wahrsch. Verw. Gebiete 62*(1), 53–72.

[82] Tudor, C. A. and F. G. Viens (2009). Variations and estimators for self-similarity parameters via Malliavin calculus. *Ann. Probab. 37*(6), 2093–2134.

[83] Tukey, J. (1938). On the distribution of the fractional part of a statistical variable. *Mat. Sbornik 4*(3), 561–562.

[84] Urbanik, K. and W. Woyczyński (1967). A random integral and Orlicz spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 15*, 161–169.

[85] von Bahr, B. and C. Esseen (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist 36*, 299–303.

[86] Whitt, W. (1971). Weak convergence of first passage time processes. *J. Appl. Probability 8*, 417–422.

[87] Whitt, W. (2002). *Stochastic-process limits.* Springer-Verlag, New York.

[88] Yan, J. (2007). Spatial stochastic volatility for lattice data. *J. Agric. Biol. Environ. Stat. 12*(1), 25–40.

Paper I

On limit theory for Lévy semi-stationary processes

Andreas Basse-O'Connor[1], *Claudio Heinrich*[1] *and Mark Podolskij*[1]
[1] *Department of Mathematics, Aarhus University, Denmark*

Abstract: In this paper we present some limit theorems for power variation of Lévy semi-stationary processes in the setting of infill asymptotics. Lévy semi-stationary processes, which are a one-dimensional analogue of ambit fields, are moving average type processes with a multiplicative random component, which is usually referred to as volatility or intermittency. From the mathematical point of view this work extends the asymptotic theory investigated in [12], where the authors derived the limit theory for kth order increments of stationary increments Lévy driven moving averages. The asymptotic results turn out to heavily depend on the interplay between the given order of the increments, the considered power $p > 0$, the Blumenthal–Getoor index $\beta \in (0, 2)$ of the driving pure jump Lévy process L and the behaviour of the kernel function g at 0 determined by the power α. In this paper we will study the first order asymptotic theory for Lévy semi-stationary processes with a random volatility/intermittency component and present some statistical applications of the probabilistic results.

I.1 Introduction and main results

Over the last ten years there has been a growing interest in the theory of ambit fields. Ambit fields is a class of spatio-temporal stochastic processes that has been originally introduced by Barndorff-Nielsen and Schmiegel in a series of papers [9, 10, 11] in the context of turbulence modelling, but which has found manifold applications in mathematical finance and biology among other sciences; see e.g. [2, 25].

Ambit processes describe the dynamics in a stochastically developing field, for instance a turbulent wind field, along curves embedded in such a field. A key characteristic of the modelling framework is that beyond the most basic kind of random noise it also specifically incorporates additional, often drastically changing, inputs referred to as *volatility* or *intermittency*. In terms of mathematical formulae an ambit field is specified via

$$X_t(x) = \mu + \int_{A_t(x)} g(t,s,x,\xi)\sigma_s(\xi)\, L(ds,d\xi) + \int_{D_t(x)} q(t,s,x,\xi)a_s(\xi)\, ds\, d\xi, \quad \text{(I.1.1)}$$

where t denotes time while x gives the position in space. Further, $A_t(x)$ and $D_t(x)$ are Borel measurable subsets of $\mathbb{R}\times\mathbb{R}^d$, g and q are deterministic weight functions, σ represents the intermittency field, a is a drift field and L denotes an independently scattered infinitely divisible random measure on $\mathbb{R}\times\mathbb{R}^d$ (see e.g. [30] for details). In the literature, the sets $A_t(x)$ and $D_t(x)$ are usually referred to as *ambit sets*. In the framework of turbulence modelling the stochastic field $(X_t(x))_{t\geq 0,\, x\in\mathbb{R}^3}$ describes the velocity of a turbulent flow at time t and position x, while the ambit sets $A_t(x), D_t(x)$ are typically bounded.

In this paper we consider a purely temporal analogue of ambit fields (without drift) $(X_t)_{t\in\mathbb{R}}$, defined on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathbb{R}},\mathbb{P})$, which is given as

$$X_t = \int_{-\infty}^{t} \big\{g(t-s) - g_0(-s)\big\}\sigma_{s-}\, dL_s, \quad \text{(I.1.2)}$$

and is usually referred to as a *Lévy semi-stationary (LSS) process*. Here $L=(L_t)_{t\in\mathbb{R}}$ is a symmetric Lévy process on $\mathbb{R}$ with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}}$ with $L_0=0$ and without a Gaussian component. That is, for all $u\in\mathbb{R}$, the process $(L_{t+u}-L_u)_{t\geq 0}$ is a symmetric Lévy process on $\mathbb{R}_+$ with respect to $(\mathcal{F}_{t+u})_{t\geq 0}$. The process $(\sigma_t)_{t\in\mathbb{R}}$ is assumed to be càdlàg and adapted to $(\mathcal{F}_t)_{t\in\mathbb{R}}$, and g and g_0 are deterministic continuous functions vanishing on $(-\infty,0)$. The name Lévy semi-stationary process refers to the fact that the process $(X_t)_{t\in\mathbb{R}}$ is stationary whenever $g_0=0$ and $(\sigma_t)_{t\in\mathbb{R}}$ is stationary and independent of $(L_t)_{t\in\mathbb{R}}$. It is assumed throughout this paper that g, g_0, σ and L are such that the process (X_t) is well-defined, which is in particular satisfied under the conditions stated in Remark I.3.3 below. We are interested in the asymptotic behaviour of the power variation of the process X. More precisely, let us consider the kth order increments $\Delta^n_{i,k}X$ of X, $k\in\mathbb{N}$, that are defined by

$$\Delta^n_{i,k}X := \sum_{j=0}^{k}(-1)^j\binom{k}{j}X_{(i-j)/n}, \qquad \text{where } i\geq k.$$

For instance, we have that $\Delta_{i,1}^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ and $\Delta_{i,2}^n X = X_{\frac{i}{n}} - 2X_{\frac{i-1}{n}} + X_{\frac{i-2}{n}}$. The main functional of interest is the power variation computed on the basis of kth order increments:

$$V(p;k)_t^n := \sum_{i=k}^{[nt]} |\Delta_{i,k}^n X|^p, \qquad p > 0. \tag{I.1.3}$$

At this stage we remark that power variation of stochastic processes has been a very active research area in the last decade. We refer e.g. to [7, 22, 23, 29] for limit theory for power variations of Itô semimartingales, to [3, 5, 17, 21, 28] for the asymptotic results in the framework of fractional Brownian motion and related processes, and to [16, 34] for investigations of power variation of the Rosenblatt process. The power variation of Brownian semi-stationary processes, which is the model (I.1.2) driven by a Brownian motion, has been studied in [4, 6, 19]. Under proper normalisation the authors have shown convergence in probability for the statistic $V(p;k)_t^n$ and proved its asymptotic mixed normality.

However, when the driving motion in (I.1.2) is a pure jump Lévy process, the asymptotic theory is very different from the Brownian case. In the recent work [12] the power variation of the model (I.1.2) with constant intermittency σ has been studied. The authors showed that the asymptotic behavior of $V(p;k)_t^n$ is greatly affected by the Blumenthal–Getoor index β of the driving Lévy motion as well as the behavior of the kernel function g at 0. The goal of this work is to extend the result of [12] to LSS-processes with nontrivial intermittency process σ. Such extensions are important in applications, say in the framework of turbulence, since the intermittency is often the main object of interest. Moreover, we show that the convergence holds functional with respect to the Skorokhod M_1-topology in the setting of Theorem I.1.1 (i), and with respect to the uniform norm in the settings of Theorem I.1.1 (ii) and (iii).

Throughout this article, β denotes the Blumenthal–Getoor index of the driving Lévy process, which is defined as

$$\beta := \inf\left\{r \geq 0 : \int_{-1}^{1} |x|^r \,\nu(dx) < \infty\right\} \in [0,2],$$

where ν denotes the Lévy measure of L. It is well-known that $\sum_{s\in[0,1]} |\Delta L_s|^p$ is finite when $p > \beta$, while it is infinite for $p < \beta$. Here $\Delta L_s = L_s - L_{s-}$ where $L_{s-} = \lim_{u\uparrow s,\, u<s} L_u$. We recall that for a stable Lévy processes the Blumenthal–Getoor index matches the index of stability. The authors of [12] impose the following set of assumptions on g, g_0 and ν, which we assume to hold throughout this paper.

Assumption (A): *The function $g : \mathbb{R} \to \mathbb{R}$ satisfies $\lim_{t\downarrow 0} g(t)t^{-\alpha} = c_0$ for some $\alpha > 0$ and $c_0 \neq 0$. There is a $\theta \in (0,2]$, such that $\limsup_{t\to\infty} \nu(x : |x| \geq t)t^\theta < \infty$ and $g - g_0$ is a bounded function in $L^\theta(\mathbb{R}_+)$. Furthermore, g is k-times continuously differentiable on $(0,\infty)$ and there exists a $\delta > 0$ such that $|g^{(k)}(t)| \leq Ct^{\alpha-k}$ for all $t \in (0,\delta)$, and such that both $|g'|$ and $|g^{(k)}|$ are in $L^\theta((\delta,\infty))$ and are decreasing on (δ,∞).*

Assumption (A-log): *In addition to (A) suppose that*

$$\int_\delta^\infty |g^{(k)}(s)|^\theta \log(1/|g^{(k)}(s)|) \, ds < \infty.$$

Assumption (A) ensures, in particular, that the process X with $\sigma = 1$ is well-defined, cf. [12]. When L is a β-stable Lévy process, we can and will always choose $\theta = \beta$ in assumption (A). In addition to these assumptions we use in our main result the following integrability conditions on the stochastic process $H_s := g^{(k)}(-s)\sigma_s \mathbb{1}_{(-\infty,-\delta]}(s)$, $s \in \mathbb{R}$, where δ is defined as in assumption (A).

Assumption (B1): There exists $\rho > 0$ with $\rho \leq 1 \wedge \theta$ and $\beta' > \beta$ with $\beta' \geq p$ such that

$$\mathbb{E}\Big[\Big(\int_{\mathbb{R}} \big(|H_s|^\rho \vee |H_s|^{\beta'}\big)\, ds\Big)^{1\vee \frac{p}{2}}\Big] < \infty. \tag{I.1.4}$$

For $\theta = 1$ suppose in addition that we may choose $\rho < 1$ in (I.1.4).

Assumption (B2): It holds that

$$\mathbb{E}\Big[\int_{\mathbb{R}} |H_s|^\beta\, ds\Big] < \infty.$$

For $p \leq 2$ it is not difficult to show that (B1) is at least satisfied when we can choose $\theta < 1$ in (A), and the intermittency satisfies $\sup_{s\in(-\infty,-\delta]} \mathbb{E}[|\sigma_s|^{1\vee\beta'}] < \infty$. Assumption (B2) will only be used in the case where L is a β-stable Lévy motion (see Theorem I.1.1 (ii) below), and is e.g. satisfied when $\sup_{s\in(-\infty,-\delta]} \mathbb{E}[|\sigma_s|^\beta] < \infty$. These stronger assumptions are satisfied in many applications, as σ is often assumed to be stationary.

Before we state our main theorem we introduce some more notation. Let $h_k : \mathbb{R} \to \mathbb{R}$ be given by

$$h_k(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (x-j)_+^\alpha, \qquad x \in \mathbb{R}, \tag{I.1.5}$$

where $y_+ = \max\{y, 0\}$ for all $y \in \mathbb{R}$. Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ and $(T_m)_{m\geq 1}$ be a sequence of $\mathbb{F}$-stopping times that exhausts the jumps of $(L_t)_{t\geq 0}$. That is, $\{T_m(\omega) : m \geq 1\} \cap [0,\infty) = \{t \geq 0 : \Delta L_t(\omega) \neq 0\}$ and $T_m(\omega) \neq T_n(\omega)$ for all $m \neq n$ with $T_m(\omega) < \infty$. Let $(U_m)_{m\geq 1}$ be independent and uniform $[0,1]$-distributed random variables, defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of the original probability space, which are independent of $\mathcal{F}$. By $\big(\mathbb{D}(\mathbb{R}_+;\mathbb{R}), M_1\big)$ we denote the Skorokhod space of càdlàg functions from $\mathbb{R}_+$ into $\mathbb{R}$, equipped with the Skorokhod M_1-topology, making it a Polish space. The M_1-topology was originally introduced in [33]. We give a definition in Section I.4, a detailed account and many properties can be found in [35]. For stochastic processes Z^n, Z with càdlàg sample paths that are defined on $(\Omega', \mathcal{F}')$, we denote by $Z^n \xrightarrow{\mathcal{L}_{M_1}-s} Z$ the functional $\mathcal{F}$-stable convergence in law with respect to the M_1-topology. That is, $Z^n \xrightarrow{\mathcal{L}_{M_1}-s} Z$ means that $\mathbb{E}'[\phi(Z^n)Y] \to \mathbb{E}'[\phi(Z)Y]$ for all bounded continuous functions $\phi : (\mathbb{D}(\mathbb{R}_+;\mathbb{R}), M_1) \to \mathbb{R}$, and all bounded $\mathcal{F}$-measurable Y, where $\mathbb{E}'$ denotes the expectation on the extended space $(\Omega', \mathcal{F}', \mathbb{P}')$. By $\xrightarrow{\text{u.c.p.}}$ we denote uniform convergence on compact sets in probability. That is, $(Z_t^n)_{t\geq 0} \xrightarrow{\text{u.c.p.}} (Z_t)_{t\geq 0}$ as $n \to \infty$ means that $\mathbb{P}(\sup_{t\in[0,N]} |Z_t^n - Z_t| > \varepsilon) \to 0$ for all $N \in \mathbb{N}$ and all $\varepsilon > 0$.

The following extension of [12, Theorem 1.1], to include a non-trivial σ process and functional convergence, is the main result of this paper.

Theorem I.1.1. *Let $X = (X_t)_{t\geq 0}$ be a stochastic process defined by* (I.1.2)*. Let (A) be satisfied and assume that the Blumenthal–Getoor index satisfies $\beta < 2$.*

(i) Suppose that (B1) holds and if $\theta = 1$ assume additionally that (A-log) is satisfied. Let $\alpha < k - 1/p$, $p > \beta$ and $p \geq 1$. Then, as $n \to \infty$, the functional $\mathcal{F}$-stable convergence holds

$$n^{\alpha p} V(p;k)^n_t \xrightarrow{\mathcal{L}_{M_1}-s} |c_0|^p \sum_{m\colon T_m \in [0,t]} |\Delta L_{T_m} \sigma_{T_m -}|^p V_m$$

where $V_m = \sum_{l=0}^{\infty} |h_k(l + U_m)|^p$.

(ii) Suppose that L is a symmetric β-stable Lévy process with $\beta \in (0,2)$ and scale parameter $\gamma > 0$. Suppose that (B2) holds and that $\alpha < k - 1/\beta$ and $p < \beta$. Then as $n \to \infty$

$$n^{-1+p(\alpha+1/\beta)} V(p;k)^n_t \xrightarrow{u.c.p.} m_p \int_0^t |\sigma_s|^p ds,$$

where $m_p = |c_0|^p \gamma^p (\int_{\mathbb{R}} |h_k(x)|^\beta \, dx)^{p/\beta} \mathbb{E}[|Z|^p]$, where Z is a symmetric β-stable random variable with scale parameter 1.

(iii) Suppose that (B1) holds, $\theta > 1$, $\alpha > k - 1/(\beta \vee p)$ and $p \geq 1$. If $p = \theta$ assume additionally that (A-log) is satisfied. Then, as $n \to \infty$,

$$n^{-1+pk} V(p;k)^n_t \xrightarrow{u.c.p.} \int_0^t |F_u|^p \, du,$$

where $(F_u)_{u \in \mathbb{R}}$ is a version with measurable sample paths of the process defined by

$$F_u = \int_{-\infty}^u g^{(k)}(u-s)\sigma_{s-} \, dL_s \quad \text{a.s. for all } u \in \mathbb{R},$$

which necessarily satisfies $\int_0^t |F_u|^p \, du < \infty$, almost surely.

Under the integrability assumptions (B1) and (B2), Theorem I.1.1 covers all possible choices of $\alpha > 0, \beta \in [0,2)$ and $p \geq 1$ except the critical cases where $p = \beta$, $\alpha = k - 1/p$ or $\alpha = k - 1/\beta$. The two critical cases $\alpha = k - 1/p$, $p > \beta$ and $\alpha = k - 1/\beta$, $p < \beta$ have been studied in [13] in the case $\sigma \equiv 1$. We conjecture that analogous results hold for LSS processes with non-trivial intermittency component, but will not pursue this theory in the paper.

First order asymptotic theory for Lévy semi-stationary processes can be used to draw inference on the parameters α, β and on certain intermittency functionals in the context of high frequency observations, see Section I.2. Furthermore, this type of limit theory is an intermediate step towards asymptotic results for general ambit fields of the form (I.1.1). We remark that, in contrast to the Brownian setting, extending the first order limit theory presented in [12] to Lévy semi-stationary processes with non-trivial σ is a more complex issue. This is due to the fact that it is harder to estimate various norms of X and related processes when the driving process L is a

Lévy process. To this end, we rely heavily on decoupling techniques and isometries for stochastic integral mappings presented in the monograph [26] and [31], see Section I.3 for more details.

This paper is structured as follows. Section I.2 is devoted to various statistical applications of our limit theory. In Section I.3 we discuss properties of Lévy integrals of predictable processes and recall essential estimates from [26] for those integrals. All proofs are demonstrated in Section I.4.

I.2 Some statistical applications

We start this section by giving an interpretation to the parameters $\alpha > 0$ and $\beta \in (0,2)$. Let us consider the linear fractional stable motion defined by

$$Y_t := c_0 \int_{\mathbb{R}} \{(t-s)_+^{\alpha} - (-s)_+^{\alpha}\} \, dL_s,$$

where L is symmetric β-stable, and the constant c_0 has been introduced in assumption (A). It is well known that the process $(Y_t)_{t\geq 0}$ is well defined whenever $H = \alpha + 1/\beta < 1$. Furthermore, the process $(Y_t)_{t\geq 0}$ has stationary symmetric β-stable increments, Hölder continuous paths of all orders smaller than α and self-similarity index H, i.e.

$$(Y_{at})_{t\geq 0} \stackrel{\mathrm{d}}{=} \left(a^H Y_t\right)_{t\geq 0} \qquad \text{for any } a \in \mathbb{R}_+.$$

We refer to e.g. [14] for more details. As it has been discussed in [12, 13] in the setting $\sigma = 1$, the small scale behaviour of the process X is well approximated by the corresponding behaviour of the linear fractional stable motion Y. In other words, when the intermittency process σ is smooth, we have that

$$X_{t+\Delta} - X_t \approx \sigma_t (Y_{t+\Delta} - Y_t)$$

for small $\Delta > 0$. Thus, intuitively speaking, the properties of Y (Hölder smoothness, self-similarity) transfer to the process X on small scales.

Having understood the role of the parameters $\alpha > 0$ and $H = \alpha + 1/\beta \in (1/2, 1)$ from the modelling perspective, it is obviously important to investigate estimation methods for these parameters. We note that the conditions $\alpha > 0$ and $H \in (1/2, 1)$ imply the restrictions $\beta \in (1,2)$ and $\alpha < 1 - 1/\max\{p, \beta\}$. Hence, the regime of Theorem I.1.1 (iii) is never applicable.

We start with a direct estimation procedure, which identifies the convergence rates in Theorem I.1.1 (i)-(ii). We apply these convergence results only for $t = 1$ and $k = 1$. For $p \in [\underline{p}, \overline{p}]$ with $\underline{p} \in (0,1)$ and $\overline{p} > 2$, we introduce the statistic

$$S(n,p) := -\frac{\log V(p)^n}{\log n} \qquad \text{with} \qquad V(p)^n = V(p;1)_1^n.$$

When the underlying Lévy motion L is symmetric β-stable and the assumptions of Theorems I.1.1 (i)-(ii) are satisfied, we obtain that

$$S(n,p) \xrightarrow{\mathbb{P}} S_{\alpha,\beta}(p) := \begin{cases} \alpha p: & \alpha < 1 - 1/p \text{ and } p > \beta \\ pH - 1: & \alpha < 1 - 1/\beta \text{ and } p < \beta \end{cases}, \tag{I.2.1}$$

if the parameter is (α, β). Indeed, the result of Theorem I.1.1 (i) shows that

$$\frac{\alpha p \log n + \log V(p)^n}{\log n} \xrightarrow{\mathcal{L}-s} 0 \qquad \Rightarrow \qquad \frac{\alpha p \log n + \log V(p)^n}{\log n} \xrightarrow{\mathbb{P}} 0.$$

This explains the first line in (I.2.1), and the second line follows similarly from Theorem I.1.1 (ii). At this stage we remark that the limit $S_{\alpha,\beta} : [\underline{p}, \overline{p}] \setminus \{\beta\} \to \mathbb{R}$ is a piecewise linear function with two different slopes. It can be continuously extended to the function $S_{\alpha,\beta} : [\underline{p}, \overline{p}] \to \mathbb{R}$, whose definition can be further extended to include all values

$$(\alpha, \beta) \in J := \left\{(\alpha, \beta) \in \mathbb{R}^2 : \ \beta \in [1, 2], \ \alpha \in [0, 1 - 1/\beta]\right\}.$$

For estimation of (α, β), it is natural to minimise the L^2-distance between the observed scale function $S(n, p)$ and the theoretical limit $S_{\alpha,\beta}(p)$:

$$(\hat{\alpha}_n, \hat{\beta}_n) \in \operatorname{argmin}_{(\alpha,\beta)\in J} \|S(n) - S_{\alpha,\beta}\|_{L^2([\underline{p},\overline{p}])} \tag{I.2.2}$$

with $S(n) := S(n, \cdot)$. This approach is somewhat similar to the estimation method proposed in [20]. For finite n, the minimum of the $L^2([\underline{p}, \overline{p}])$-distance at (I.2.2) is not necessarily obtained at a unique point, and we take an arbitrary measurable minimiser $(\hat{\alpha}_n, \hat{\beta}_n)$. Our next result shows consistency of the estimator $(\hat{\alpha}_n, \hat{\beta}_n)$.

Corollary I.2.1. *Let $(\alpha_0, \beta_0) \in J^\circ$, where J° is the set of all inner points of J, denote the true parameter of the model* (I.1.2)*, and let L be a symmetric β_0-stable Lévy motion. Assume that the conditions of Theorem I.1.1 (i) (resp. Theorem I.1.1 (ii)) hold when $\alpha_0 \in (0, 1 - 1/p)$ and $p > \beta_0$ (resp. $\alpha_0 \in (0, 1 - 1/\beta_0)$ and $p < \beta_0$). Then we obtain convergence in probability*

$$(\hat{\alpha}_n, \hat{\beta}_n) \xrightarrow{\mathbb{P}} (\alpha_0, \beta_0).$$

Proof. Set $r_0 = (\alpha_0, \beta_0)$ and $\widehat{r}_n = (\hat{\alpha}_n, \hat{\beta}_n)$. We first show the convergence

$$\|S(n) - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} \xrightarrow{\mathbb{P}} 0. \tag{I.2.3}$$

From (I.2.1) we deduce that $S(n, p) \xrightarrow{\mathbb{P}} S_{r_0}(p)$ for all $p \in [\underline{p}, \overline{p}] \setminus \{\beta_0\}$. Furthermore, for any $p \in [\underline{p}, \overline{p}]$, it holds that

$$\left(V(\overline{p})^n\right)^{1/\overline{p}} \le \left(V(p)^n\right)^{1/p} \le \left(V(\underline{p})^n\right)^{1/\underline{p}}.$$

Hence, we deduce the inequality

$$\left|\frac{\log V(p)^n}{\log n}\right| \le \max\left\{\frac{p}{\overline{p}} \cdot \left|\frac{\log V(\overline{p})^n}{\log n}\right|, \frac{p}{\underline{p}} \cdot \left|\frac{\log V(\underline{p})^n}{\log n}\right|\right\}.$$

Since $|\log V(\underline{p})^n / \log n| \xrightarrow{\mathbb{P}} \underline{p}(\alpha_0 + 1/\beta_0) - 1$ and $|\log V(\overline{p})^n / \log n| \xrightarrow{\mathbb{P}} \alpha_0 \overline{p}$, because $\underline{p} < 1 < \beta_0$ and $\overline{p} > 2 > \beta_0$, we readily deduce the convergence at (I.2.3) by dominated convergence theorem.

Now, we note that the mapping $G : J \to G(J) \subset L^2([\underline{p}, \overline{p}])$, $r \mapsto S_r$, is a homeomorphism. Thus, it suffices to prove that $\|S_{\widehat{r}_n} - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} \xrightarrow{\mathbb{P}} 0$ to conclude $\widehat{r}_n \xrightarrow{\mathbb{P}} r_0$. To show the former we observe that

$$\begin{aligned}
\|S_{\widehat{r}_n} - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} &\leq \|S(n) - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} + \|S(n) - S_{\widehat{r}_n}\|_{L^2([\underline{p},\overline{p}])} \\
&= \|S(n) - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} + \min_{r \in J} \|S(n) - S_r\|_{L^2([\underline{p},\overline{p}])} \\
&\leq 2\|S(n) - S_{r_0}\|_{L^2([\underline{p},\overline{p}])} \xrightarrow{\mathbb{P}} 0.
\end{aligned}$$

This completes the proof of Corollary I.2.1. □

In practice the integral in (I.2.2) needs to be discretised. We further remark that the estimator $S(n,p)$ has the rate of convergence $\log n$ due to the bias $V(p)/\log n$, where $V(p)$ denotes the limit of $V(p)^n$.

As for the estimation of the self-similarity parameter $H = \alpha + 1/\beta \in (1/2, 1)$, there is an alternative estimator based on a ratio statistic. Recalling that $\beta \in (1,2)$, we deduce for any $p \in (0,1]$

$$R(n,p) := \frac{\sum_{i=2}^n |X_{\frac{i}{n}} - X_{\frac{i-2}{n}}|^p}{\sum_{i=1}^n |X_{\frac{i}{n}} - X_{\frac{i-1}{n}}|^p} \xrightarrow{\mathbb{P}} 2^{pH}$$

by a direct application of Theorem I.1.1 (ii). Thus, we immediately conclude that

$$\hat{H}_n := \frac{\log R(n,p)}{p \log 2} \xrightarrow{\mathbb{P}} H.$$

This type of idea is rather standard in the framework of a fractional Brownian motion with Hurst parameter H. It has been also applied to Brownian semi-stationary processes in [4, 6]. Theorem 1.2 (i) in [12], which has been shown in the setting $\sigma = 1$, suggests that the statistic $\hat{H}_n$ has convergence rate $n^{1-1/(1-\alpha)\beta}$ whenever $p \in (0, 1/2]$. Furthermore, the rate of convergence can be improved to $\sqrt{n}$ via using kth order increments with $k \geq 2$ (cf. [12, Theorem 1.2 (ii)]). However, we dispense with the precise proof of these statements for non-constant intermittency process σ. In a recent work [18] it was shown that for linear fractional stable motions the convergence $\hat{H}_n \xrightarrow{\mathbb{P}} H$ continues to hold for powers $p \in (-1, 0)$. This is particularly useful, since choosing p negative ensures that the condition $p < \beta$ of Theorem I.1.1 (ii) is always satisfied. However, proving this result for a general Lévy semi-stationary process is a much more delicate issue.

Another important object for applications in turbulence modelling is the intermittency process σ. First of all, we remark that the process σ in the general model (I.1.2) is statistically not identifiable. This is easily seen, because multiplication of σ by a constant can not be distinguished from the multiplication of, say, Lévy process L by the same constant. However, it is very well possible to estimate the *relative intermittency*, which is defined as

$$RI(p) := \frac{\int_0^t |\sigma_s|^p ds}{\int_0^1 |\sigma_s|^p ds}, \qquad t \in (0,1),$$

for $p \in (0,1]$. The relative intermittency, which has been introduced in [8] for $p = 2$ in the context of Brownian semi-stationary processes, describes the relative amplitude of the velocity process on an interval $[0,1]$. Applying the convergence result of Theorem I.1.1 (ii) for $p \in (0,1]$, the relative intermittency can be consistently estimated via

$$RI(n,p) := \frac{V(p)^n_t}{V(p)^n_1} \xrightarrow{\mathbb{P}} RI(p).$$

Again we suspect that the associated convergence rate is $n^{1-1/(1-\alpha)\beta}$ whenever $p \in (0,1/2]$ as suggested by [12, Theorem 1.2 (i)].

I.3 Preliminaries: Estimates on Lévy integrals

To prove the various limit theorems we need very sharp estimates of the pth moments of the increments of process X defined in (I.1.2). In fact, we need such estimates for several different processes related to X obtained by different truncations. When $F : \mathbb{R}_+ \to \mathbb{R}$ is a deterministic function, the estimates for integrals $\int_0^t F_s \, dL_s$ go back to Rajput and Rosiński [30, Theorem 3.3]. Their results imply the existence of a constant $C > 0$ such that

$$\mathbb{E}\bigg[\bigg|\int_0^t F_s \, dL_s\bigg|^q\bigg] \le C\|F\|^q_{L,q},$$

where $\|\cdot\|_{L,q}$ is a certain functional to be defined below (when L is symmetric and without Gaussian component). The decoupling approach used in Kwapién and Woyczyński [26] provides an extension of the results to general predictable F, see Lemmas I.3.1 and I.3.2 below. Before stating the results precisely, we need the following notation.

Let $L = (L_t)_{t\in\mathbb{R}}$ be a symmetric Lévy process on the real line with $L_0 = 0$, Lévy measure ν and without a Gaussian component. For a predictable process $(F_t)_{t\in\mathbb{R}}$ and for $q = 0$ or $q \ge 1$ we define

$$\Phi_{q,L}(F) := \int_{\mathbb{R}^2} \phi_q(F_s u) \, ds \, \nu(du), \quad \text{where} \quad \phi_q(x) := |x|^q \mathbb{1}_{\{|x|>1\}} + x^2 \mathbb{1}_{\{|x|\le 1\}}.$$

A predictable process $F = (F_t)_{t\in\mathbb{R}}$ is integrable with respect to $(L_t)_{t\in\mathbb{R}}$ in the sense of [26] if and only if $\Phi_{0,L}(F) < \infty$ almost surely (cf. [26, Theorem 9.1.1]). The linear space of predictable processes satisfying $\Phi_{q,L}(F) < \infty$ will be denoted by $\mathbf{L}^q(dL)$L. In order to estimate the qth moments of stochastic integrals we introduce for all $q \ge 1$

$$\|F\|_{q,L} := \inf\{\lambda > 0 \, : \, \Phi_{q,L}(F/\lambda) \le 1\}, \qquad F \in \mathbf{L}^q(dL)L. \tag{I.3.1}$$

The following two results from [26] and [31] will play a key role for our proofs.

Lemma I.3.1 ([26], Equation (9.5.3))**.** *For all $q \ge 1$ there is a constant C, depending only on q, such that we obtain for all $F \in \mathbf{L}^q(dL)L$*

$$\mathbb{E}\bigg[\bigg|\int_{\mathbb{R}} F_s \, dL_s\bigg|^q\bigg] \le C\mathbb{E}\big[\|F\|^q_{q,L}\big]. \tag{I.3.2}$$

The above lemma follows by [26, Equation (9.5.3)] and the comments following it. Actually, [26, Equation (9.5.3)] only treats the case where the stochastic integral in (I.3.2) is over a finite time interval, say $\int_0^t F_s \, dL_s$. However, the definition of the stochastic integral and the estimates of the integral in [26, Chapters 8–9] extend to the case of $\int_{\mathbb{R}} F_s \, dL_s$ in a natural way.

For the next result, which is an immediate consequence of [31, Theorem 2.1], we use the notation $\|Z\|_{\beta,\infty}^{\beta} = \sup_{\lambda>0} \lambda^{\beta}\mathbb{P}[|Z| > \lambda]$ for an arbitrary random variable Z. For $q < \beta$ it holds that $\mathbb{E}[|Z|^q]^{1/q} \leq \|Z\|_{\beta,\infty} \leq (\frac{\beta}{\beta-q})^{1/q}\mathbb{E}[|Z|^{\beta}]^{1/\beta}$. In the literature, $\|\cdot\|_{\beta,\infty}$ is often referred to as the weak L^{β}-norm. However, $\|\cdot\|_{\beta,\infty}$ satisfies the triangle inequality only up to a constant.

Lemma I.3.2 ([31], Theorem 2.1)**.** *Let $(L_t)_{t\in\mathbb{R}}$ be a symmetric β-stable Lévy process. Then there is a positive constant $C > 0$ such that for all $(F_t)_{t\in\mathbb{R}}$ in $\mathbf{L}^0(dL)L$ it holds that*

$$\Big\| \int_{\mathbb{R}} F_s \; dL_s \Big\|_{\beta,\infty}^{\beta} \leq C\mathbb{E}\Big[\int_{\mathbb{R}} |F_s|^{\beta} \; ds \Big].$$

The next remark gives sufficient conditions for the process X introduced at (I.1.2) to be well-defined.

Remark I.3.3. Suppose that (A) is satisfied and define the two processes $F^{(1)}$ and $F^{(2)}$ by $F_s^{(1)} = (g(-s) - g_0(-s))\sigma_s$ and $F_s^{(2)} = g'(-s)\sigma_s$ for $s < 0$. Then the process X given by (I.1.2) is well-defined if there exists a $\beta' > \beta$ such that

$$\int_{-\infty}^{-\delta} \Big(|F_s^{(i)}|^{\theta} \mathbb{1}_{\{|F_s^{(i)}| \leq 1\}} + |F_s^{(i)}|^{\beta'} \mathbb{1}_{\{|F_s^{(i)}| > 1\}} \Big) \, ds < \infty \tag{I.3.3}$$

almost surely for $i = 1, 2$. To show the above we argue as follows: For any $\beta' \in (\beta, 2]$ we deduce from (A) and simple calculations the estimate

$$\int_{\mathbb{R}} \big(|ux|^2 \wedge 1\big)\nu(dx) \leq C(|u|^{\theta}\mathbb{1}_{\{|u|\leq 1\}} + |u|^{\beta'}\mathbb{1}_{\{|u|>1\}}), \qquad u \in \mathbb{R}. \tag{I.3.4}$$

Then, an application of the mean value theorem combined with assumption (I.3.3) yields that $\Phi_{0,L}(H^{(t)}) < \infty$ almost surely for all $t > 0$, where $H_s^{(t)} = (g(t-s) - g_0(-s))\sigma_s$. This guarantees the existence of the process X due to [26, Theorem 9.1.1].

In our proofs we will need the following properties of the functional $\|\cdot\|_{L,q}$ defined in (I.3.1).

i Homogeneity: For all $\lambda \in \mathbb{R}$, $F \in \mathbf{L}^q(dL)L$, $\|\lambda F\|_{q,L} = |\lambda| \|F\|_{q,L}$.

ii Triangle inequality (up to a constant): There exists a constant $C > 0$ such that for all $F^1, ..., F^m \in \mathbf{L}^q(dL)L$ we have

$$\|F^1 + \cdots + F^m\|_{q,L} \leq C\big(\|F^1\|_{q,L} + \cdots + \|F^m\|_{q,L}\big), \tag{I.3.5}$$

and the constant C does not depend on m or L.

iii Upper bound: For all $F \in \mathbf{L}^q(dL)L$ we have

$$\|F\|_{q,L} \le \Phi_{q,L}^{1/2}(F) \vee \Phi_{q,L}^{1/q}(F). \tag{I.3.6}$$

Property (i) follows directly from the definition of $\|\cdot\|_{L,q}$ in (I.3.1). To show property (ii) it is sufficient to derive (I.3.5) for $F^1, \dots, F^m \in \mathbf{L}^q_{\mathrm{nr}}(dL)L$, where $\mathbf{L}^q_{\mathrm{nr}}(dL)L$ denotes the subspace of nonrandom processes in $\mathbf{L}^q(dL)$L. We will show that there is a norm $\|\cdot\|'_{q,L}$ on $\mathbf{L}^q_{\mathrm{nr}}(dL)L$ and $c > 0$ and $C > 0$ such that $c\|F\|'_{q,L} \le \|F\|_{q,L} \le C\|F\|'_{q,L}$, for all $F \in \mathbf{L}^q_{\mathrm{nr}}(dL)L$, which then implies (I.3.5). To this end, let

$$\widetilde{\phi}_q(x) := (2/q|x|^q + 1 - 2/q)\mathbb{1}_{\{|x|>1\}} + x^2\mathbb{1}_{\{|x|\le 1\}}.$$

Clearly, there exist $c, C > 0$ such that $c\widetilde{\phi}_q(x) \le \phi_q(x) \le C\widetilde{\phi}_q(x)$ for all $x \in \mathbb{R}$. Since the function $\widetilde{\phi}_q$ is convex, the functional

$$\|F\|'_{q,L} = \inf\left\{\lambda \ge 0 \,:\, \int_{\mathbb{R}^2} \widetilde{\phi}_q(F_s u/\lambda)\, ds\, \nu(du) \le 1\right\}$$

is a norm on $\mathbf{L}^q_{\mathrm{nr}}(dL)L$, called the Luxemburg norm (cf. [27, Chapter 1]). Using convexity of $\widetilde{\phi}_q$ it follows by straightforward calculations that $c\|F\|'_{q,L} \le \|F\|_{q,L} \le C\|F\|'_{q,L}$ for all $F \in \mathbf{L}^q_{\mathrm{nr}}(dL)L$. This implies (I.3.5). Finally, property (iii) follows by the fact that $\phi_q(\lambda x) \le (\lambda^2 \vee \lambda^q)\phi_q(x)$ for all $\lambda \ge 0$.

We conclude this subsection with a remark on the situation when the integrator is a non-symmetric Lévy process $(\widetilde{L}_t)_{t\in\mathbb{R}}$ with $\widetilde{L}_0 = 0$, Lévy measure $\widetilde{\nu}$, shift parameter η, without a Gaussian part, and the truncation function $\tau\colon x \mapsto 1_{\{|x|<1\}} + \mathrm{sign}(x)\mathbb{1}_{\{|x|\ge 1\}}$. That is, for all $\theta \in \mathbb{R}$,

$$\mathbb{E}[e^{i\theta \tilde{L}_1}] = \exp\Big(i\theta\eta + \int_{\mathbb{R}} \big(e^{i\theta x} - 1 - i\theta\tau(x)\big)\, \widetilde{\nu}(dx)\Big).$$

In this situation the modulars and norms defined above become much more involved and harder to control, which is the main reason why we consider only symmetric Lévy motions as driving processes. Moreover, assumptions (A), (B1) and (B2) are not sufficient to guarantee the existence of the integral (I.1.2) if we consider non-symmetric Lévy processes, e.g. if $L_t = \eta t$ with $\eta \ne 0$. For more details we refer to [26, Chapter 9.1]. For our purposes, the following integrability criterion with respect to non-symmetric Lévy processes will suffice. For a predictable process $(F_t)_{t\in\mathbb{R}}$ define

$$\Psi_{0,\widetilde{L}}(F) = \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \tau(uF_s) - \tau(u)F_s\widetilde{\nu}(du) + \eta F_s \Big|\, ds.$$

Then, the condition

$$\Phi_{0,\widetilde{L}}(F) + \Psi_{0,\widetilde{L}}(F) < \infty \quad \text{almost surely} \tag{I.3.7}$$

is sufficient for the integral $\int_{\mathbb{R}} F_s\, d\widetilde{L}_s$ to exist, and we write $F \in \mathbf{L}^0(dL)\widetilde{L}$. Indeed, this is a consequence of [26, Theorem 9.1.1 and pp. 217–218] combined with the estimate [30, Lemma 2.8].

I.4 Proofs

In this section we present the proofs of our main results. The proof of (i) is divided into two parts and is similar to the proof of the corresponding result in [12]. First we show the theorem under the assumption that L is a compound Poisson process with jumps bounded away from zero in absolute value by some $a > 0$. Thereafter, we argue that the contribution of the jumps of L with absolute value $\leq a$ to the power variation becomes negligible as $a \to 0$. The proof of Theorem I.1.1 (ii) relies on freezing the intermittency σ over small blocks and then deducing the result from [12, Theorem 1.1]. A key step in the proof of Theorem I.1.1 (iii) is an application of a suitable stochastic Fubini result that we introduce in Subsection I.4.

Throughout the proofs we denote all positive constants that do not depend on n or ω by C, even though they may change from line to line. Similarly, we will denote by K any positive random variable that does not depend on n, but may change from line to line. For a random variable Y and $q > 0$ we denote $\|Y\|_q = \mathbb{E}[|Y|^q]^{1/q}$. We frequently use the notation

$$g_{i,n}(s) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} g((i-j)/n - s),$$

which allows us to express the kth order increments of X as

$$\Delta_{i,k}^n X = \int_{-\infty}^{i/n} g_{i,n}(s) \sigma_{s-} \, dL_s.$$

Recalling that $|g^{(k)}(s)| \leq Ct^{\alpha-k}$ for all $s \in (0, \delta)$ and $|g^{(k)}|$ is decreasing on (δ, ∞) by assumption (A), Taylor expansion leads to the following important estimates.

Lemma I.4.1. *Suppose that assumption (A) is satisfied. It holds that*

$$\begin{aligned}
&|g_{i,n}(s)| \leq C(i/n - s)^{\alpha} \qquad \textit{for } s \in [(i-k)/n, i/n], \\
&|g_{i,n}(s)| \leq Cn^{-k}((i-k)/n - s)^{\alpha-k} \qquad \textit{for } s \in (i/n - \delta, (i-k)/n), \textit{ and} \\
&|g_{i,n}(s)| \leq Cn^{-k}\big(\mathbb{1}_{[(i-k)/n-\delta, i/n-\delta]}(s) + g^{(k)}((i-k)/n - s)\mathbb{1}_{(-\infty,(i-k)/n-\delta)}(s)\big), \\
&\qquad\qquad \textit{for } s \in (-\infty, i/n - \delta].
\end{aligned}$$

Applying a standard localisation argument (cf. [7, Section 3]) we can and will assume throughout the proofs that the process σ is uniformly bounded by a constant on $[-\delta, \infty)$.

We conclude this subsection with a definition and some brief remarks on the Skorokhod M_1-topology. It was originally introduced by Skorokhod [33] by defining a metric on the completed graphs of càdlàg functions, where the completed graph of f is defined as

$$\Gamma_f = \{(x,t) \in \mathbb{R} \times \mathbb{R}_+ \, : \, x = \alpha f(t-) + (1-\alpha) f(t), \text{ for some } \alpha \in [0,1]\}.$$

The M_1-topology is weaker as the more commonly used J_1-topology but still strong enough to make many important functionals, such as sup and inf, continuous. It can be shown that the stable convergence in Theorem I.1.1 does not hold with respect to the J_1-topology. As M_1 is metrisable, it is entirely defined by characerising convergence of sequences, as we do in the following. A sequence f_n of functions in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ converges to $f \in \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ with respect to the Skorokhod M_1-topology if and only if $f_n(t) \to f(t)$ for all t in a dense subset of $[0, \infty)$, and for all $t_\infty \in [0, \infty)$ it holds that

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{0\le t\le t_\infty} w(f_n, t, \delta) = 0.$$

Here, the oscillation function w is defined as

$$w(f, t, \delta) = \sup_{0\vee(t-\delta)\le t_1<t_2<t_3\le(t+\delta)\wedge t_\infty} \{|f(t_2) - [f(t_1), f(t_3)]|\}, \tag{I.4.1}$$

where for $b < a$ the interval $[a, b]$ is defined to be $[b, a]$, and $|a-[b, c]| := \inf_{d\in[b,c]} |a-d|$.

Proof of Theorem I.1.1 (i)

For the proof of Theorem I.1.1 (i) we follow the strategy from [12, Theorem 1.1 (i)]. We assume first that L is a compound Poisson process with jumps bounded in absolute value away from zero by some $a > 0$. Later on, we argue that the small jumps of L are asymptotically negligible. In order to show functional $\mathcal{F}$-stable convergence on $\mathbb{D}(\mathbb{R}_+; \mathbb{R})$ it is sufficient to show $\mathcal{F}$-stable convergence on $\mathbb{D}([0, t_\infty]; \mathbb{R})$, for arbitrary but fixed $t_\infty > 0$ (cf. [35, Chapter 3.3]). Throughout this subsection we therefore fix a $t_\infty > 0$, and denote by $\mathbb{D}$ the space $\mathbb{D}([0, t_\infty]; \mathbb{R})$ equipped with the Skorokhod M_1-topology, and by $\xrightarrow{\mathcal{L}_{M_1}-s}$ the $\mathcal{F}$-stable convergence of $\mathbb{D}$-valued processes.

Compound Poisson Case

Suppose that $(L_t)_{t\in\mathbb{R}}$ is a symmetric compound Poisson process with Lévy measure ν, satisfying $\nu([-a, a]) = 0$ for some $a > 0$. Let $0 \le T_1 < T_2 < ...$ denote the jump times of $(L_t)_{t\ge 0}$ in increasing order. For $\varepsilon > 0$ we define

$$\Omega_\varepsilon = \big\{\omega \in \Omega : \text{for all } m \text{ with } T_m(\omega) \in [0, t_\infty] \text{ we have } |T_m(\omega) - T_{m-1}(\omega)| > \varepsilon \\ \text{and } \Delta L_s(\omega) = 0 \text{ for all } s \in [-\varepsilon, 0]\big\}.$$

We note that $\Omega_\varepsilon \uparrow \Omega$, as $\varepsilon \downarrow 0$. Letting

$$M_{i,n,\varepsilon} := \int_{i/n-\varepsilon}^{i/n} g_{i,n}(s)\sigma_{s-}\, dL_s, \quad \text{and} \quad R_{i,n,\varepsilon} := \int_{-\infty}^{i/n-\varepsilon} g_{i,n}(s)\sigma_{s-}\, dL_s,$$

we have the decomposition $\Delta^n_{i,k} X = M_{i,n,\varepsilon} + R_{i,n,\varepsilon}$. It turns out that $M_{i,n,\varepsilon}$ is the asymptotically dominating term, whereas $R_{i,n,\varepsilon}$ is negligible as $n \to \infty$. We show that, on Ω_ε,

$$n^{\alpha p} \sum_{i=k}^{[nt]} |M_{i,n,\varepsilon}|^p \xrightarrow{\mathcal{L}_{M_1}-s} Z_t, \quad \text{where} \tag{I.4.2}$$

$$Z_t := |c_0|^p \sum_{m:T_m\in(0,t]} |\Delta L_{T_m} \sigma_{T_m-}|^p V_m,$$

where $(V_m)_{m\geq 1}$ are defined in Theorem I.1.1 (i). Denote by i_m the random index such that $T_m \in ((i_m-1)/n, i_m/n]$. Then, we have on Ω_ε

$$n^{\alpha p}\sum_{i=k}^{[nt]}|M_{i,n,\varepsilon}|^p = n^{\alpha p}\sum_{m:T_m\in(0,[nt]/n]}|\Delta L_{T_m}\sigma_{T_m-}|^p\left(\sum_{l=0}^{v_t^m}|g_{i_m+l,n}(T_m)|^p\right) := V_t^{n,\varepsilon}, \tag{I.4.3}$$

where the random index v_t^m is defined as

$$v_t^m = v_t^m(\varepsilon,n) = \begin{cases}[\varepsilon n]\wedge([nt]-i_m) & \text{if } T_m-([\varepsilon n]+i_m)/n > -\varepsilon,\\ ([\varepsilon n]-1)\wedge([nt]-i_m) & \text{if } T_m-([\varepsilon n]+i_m)/n \leq -\varepsilon.\end{cases}$$

For the proof of (I.4.2) we first show stable convergence of the finite dimensional distributions of $V^{n,\varepsilon}$. Thereafter, we show that the sequence $(V^{n,\varepsilon})_{n\geq 1}$ is tight and deduce the functional convergence $V^{n,\varepsilon} \xrightarrow{\mathcal{L}_{M_1}-s} Z$.

Lemma I.4.2. *For $r\geq 1$ and $0\leq t_1<\cdots<t_r\leq t_\infty$ we obtain on Ω_ε the $\mathcal{F}$-stable convergence*

$$(V_{t_1}^{n,\varepsilon},\ldots,V_{t_r}^{n,\varepsilon}) \xrightarrow{\mathcal{L}-s} (Z_{t_1},\ldots,Z_{t_r}), \qquad \text{as } n\to\infty.$$

Proof. Let $(U_i)_{i\geq 1}$ be i.i.d. $\mathcal{U}([0,1])$-distributed random variables, defined on an extension $(\Omega',\mathcal{F}',\mathbb{P}')$ of the original probability space, independent of $\mathcal{F}$. By arguing as in [12, Section 5.1], we deduce for any $d\geq 1$ the $\mathcal{F}$-stable convergence

$$\{n^\alpha g_{i_m+l,n}(T_m)\}_{l,m\leq d} \xrightarrow{\mathcal{L}-s} \{c_0 h_k(l+U_m)\}_{l,m\leq d}$$

as $n\to\infty$, where h_k is defined in (I.1.5). Defining

$$V_t^{n,\varepsilon,d} := n^{\alpha p}\sum_{m\leq d:T_m\in(0,[nt]/n]}|\Delta L_{T_m}\sigma_{T_m-}|^p\left(\sum_{l=0}^{d}|g_{i_m+l,n}(T_m)|^p\right)$$

$$Z_t^d := |c_0|^p\sum_{m\leq d:T_m\in(0,t]}|\Delta L_{T_m}\sigma_{T_m-}|^p\left(\sum_{l=0}^{d}|h_k(l+U_m)|^p\right),$$

the continuous mapping theorem for stable convergence yields

$$(V_{t_1}^{n,\varepsilon,d},\ldots,V_{t_r}^{n,\varepsilon,d}) \xrightarrow{\mathcal{L}-s} (Z_{t_1}^d,\ldots,Z_{t_r}^d), \qquad \text{for } n\to\infty, \tag{I.4.4}$$

for all $d\geq 1$. It follows by Lemma I.4.1 for all l with $k\leq l<[n\delta]$ that

$$n^{\alpha p}|g_{i_m+l,n}(T_m)|^p \leq C|l-k|^{(\alpha-k)p},$$

where we recall that $(\alpha-k)p<-1$. Consequently, we find a random variable $K>0$ such that for all $t\in[0,t_\infty]$

$$|V_t^{n,\varepsilon,d}-V_t^{n,\varepsilon}| \leq K\Bigg(\sum_{m>d:T_m\in[0,t_\infty]}|\Delta L_{T_m}\sigma_{T_m-}|^p + \sum_{m:T_m\in[0,t_\infty]}\sum_{l=v_t^m\wedge d}^{\infty}|l-k|^{(\alpha-k)p}\Bigg).$$

By definition, the random index $v_t^m = v_t^m(n,\omega)$ satisfies $\liminf_{n\to\infty} v_t^m(n,\omega) = \infty$ for all ω with $T_m(\omega) \neq t$. Consequently, we obtain that $\limsup_{n\to\infty} |V_t^{n,\varepsilon,d} - V_t^{n,\varepsilon}| \to 0$ almost surely as $d \to \infty$. It follows that on Ω_ε

$$\limsup_{n\to\infty} \left\{ \max_{t\in\{t_1,\ldots,t_r\}} |V_t^{n,\varepsilon} - V_t^{n,\varepsilon,d}| \right\} \to 0, \quad \text{almost surely, as } d \to \infty. \tag{I.4.5}$$

By monotone convergence theorem we obtain $\sup_{t\in[0,t_\infty]} |Z_t^d - Z_t| \to 0$ as $d \to \infty$. Together with (I.4.4) and (I.4.5), this implies the statement of the lemma by a standard approximation argument, see for example [15, Theorem 3.2]. $\square$

Recall that the stable convergence $V^{n,\varepsilon} \xrightarrow{\mathcal{L}_{M_1}-s} Z$ is equivalent to the joint convergence in law $(V^{n,\varepsilon}, Y) \xrightarrow{\mathcal{L}} (Z, Y)$ for all $\mathcal{F}$-measurable random variables Y, cf. [24, Proposition 5.33]. Consequently, Lemma I.4.2 and the following result together with Prokhorov's theorem imply (I.4.2), where we recall that $(\mathbb{D}([0,t_\infty]), M_1)$ is a Polish space.

Lemma I.4.3. *The sequence $(V^{n,\varepsilon})_{n\geq 1}$ of $(\mathbb{D}([0,t_\infty]), M_1)$-valued random variables is tight.*

Proof. The claim follows from [35, Theorem 12.12.3] if we verify that $(V^{n,\varepsilon})_{n\geq 1}$ satisfies the conditions of the theorem. Condition (i) follows since the processes $V^{n,\varepsilon}$ are increasing in t and from tightness of $\{V_{t_\infty}^{n,\varepsilon}\}_{n\in\mathbb{N}}$, which follows from Lemma I.4.2. For condition (ii) we need to verify that for all $\zeta, \xi > 0$ there is an $\eta > 0$ such that

$$\mathbb{P}\big(\sup_{t\in[0,t_\infty]} w(V^{n,\varepsilon}, t, \eta) > \xi \big) \leq \zeta, \quad \text{for all } n,$$

where the oscillation function w was defined in (I.4.1). This follows since the processes $V^{n,\varepsilon}$ are increasing, and consequently $w(V^{n,\varepsilon}, t, \eta) = 0$ for all n, all t and all η. $\square$

This concludes the proof of (I.4.2). Next we show that

$$n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |R_{i,n,\varepsilon}|^p \xrightarrow{\mathbb{P}} 0. \tag{I.4.6}$$

Recalling that $\alpha < k - 1/p$, it is sufficient to show that

$$\sup_{n\in\mathbb{N}} \sup_{i\in\{k,\ldots,[nt_\infty]\}} n^k |R_{i,n,\varepsilon}| < \infty, \quad \text{almost surely.}$$

It follows from Lemma I.4.1 that

$$n^k |g_{i,n}(s)\sigma_{s-}| \leq C(\mathbb{1}_{[-\delta,t_\infty]}(s) + |g^{(k)}(-s)\sigma_{s-}|\mathbb{1}_{(-\infty,-\delta)}(s)) := \psi_s.$$

Let $\widetilde{L} = (\widetilde{L}_t)_{t\in\mathbb{R}}$ denote the process defined by $\widetilde{L}_0 = 0$ and $\widetilde{L}_t - \widetilde{L}_u$ is the total variation of $v \mapsto L_v$ on $(u,t]$ for all $u < t$. Since L is a compound Poisson process, the process $\widetilde{L}$ is well-defined, finite and it follows from [32, Theorem 21.9] that $\widetilde{L}$ is a Lévy process with Lévy measure $\widetilde{\nu} = 2\nu_{|\mathbb{R}_+}$ and shift parameter η with respect to the truncation function $\tau : x \mapsto x\mathbb{1}_{\{|x|<1\}} + \mathrm{sign}(x)\mathbb{1}_{\{|x|\geq 1\}}$ given by $\eta = \int_{\mathbb{R}} \tau(x)\, \widetilde{\nu}(dx)$. Next we use the following estimate:

$$n^k|R_{i,n,\varepsilon}| \leq \int_{(-\infty,\frac{i}{n}-\varepsilon]} n^k|g_{i,n}(s)\sigma_{s-}|\, d\widetilde{L}_s \leq \int_{\mathbb{R}} \psi_s\, d\widetilde{L}_s.$$

The right-hand side is finite almost surely due to the following Lemma I.4.4, and the proof of (I.4.6) is complete.

Lemma I.4.4. *Let L be a symmetric compound Poisson process with Lévy measure ν satisfying $\nu([-a,a]) = 0$ for some $a \in (0,1]$ and let $\widetilde{L}$ and ψ be given as above. Suppose, in addition, that (B1) is satisfied. Then the stochastic integral $\int_{\mathbb{R}} \psi_s\, d\widetilde{L}_s$ exists and is finite almost surely.*

Proof. To show that the stochastic integral $\int_{\mathbb{R}} \psi_s\, d\widetilde{L}_s$ is well-defined it is enough to prove that $\Phi_{0,\widetilde{L}}(\psi) + \Psi_{0,\widetilde{L}}(\psi) < \infty$ almost surely (see (I.3.7) of Section I.3). For some $\beta' > \beta$ we have from (B1) that

$$\int_{\mathbb{R}} |\psi_s|^\theta \mathbb{1}_{\{|\psi_s|\leq 1\}} + |\psi_s|^{\beta'} \mathbb{1}_{\{|\psi_s|>1\}}\, ds < \infty, \quad \text{a.s.}$$

This implies that $\Phi_{0,\widetilde{L}}(\psi) < \infty$ almost surely (cf. Remark I.3.3). Next we note that

$$\Psi_{0,L}(\psi) = \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \tau(x\psi_s) - \tau(x)\psi_s \widetilde{\nu}(dx)\; + \eta\psi_s \Big|\, ds = \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \tau(x\psi_s)\widetilde{\nu}(dx) \Big|\, ds\ ,$$

where the second equality follows by definition of η above. Hence, to show that $\Psi_{0,L}(\psi) < \infty$ almost surely, it suffices according to (B1) to derive the following estimate. There exists a constant $C > 0$ such that for all $u \in \mathbb{R}$

$$\int_{\mathbb{R}} |\tau(ux)|\, \widetilde{\nu}(dx) \leq C\big(|u|^\rho \mathbb{1}_{\{|u|\leq 1\}} + \mathbb{1}_{\{|u|>1\}}\big). \tag{I.4.7}$$

where ρ is as in assumption (B1). By the definitions of τ and $\widetilde{\nu}$ we have that

$$\int_{\mathbb{R}} |\tau(ux)|\, \widetilde{\nu}(dx) = |u| \int_{\{|x|\leq |u|^{-1}\}} |x|\, \nu(dx) + \nu\big(x \in \mathbb{R} : |xu| > 1\big). \tag{I.4.8}$$

We recall that $\limsup_{t\to\infty} \nu([t,\infty))t^\theta < \infty$. Since ν is finite, there exists $C_0 > 0$ such that $\nu([t,\infty)) \leq C_0/t^\theta$ for all $t \geq a$. Consequently, we obtain for all $t \geq a$ and $f(u) = \mathbb{1}_{[t,\infty)}(u)$

$$\int_a^\infty f(x)\, \nu(dx) \leq \frac{C_0}{\theta} \int_a^\infty f(x) x^{-\theta-1}\, dx.$$

By monotone approximation, the inequality remains valid for all nondecreasing $f : [a,\infty) \to \mathbb{R}_+$. Therefore, the first term on the right-hand side of (I.4.8) is bounded by

$$\begin{aligned} |u| \int_{\{|x|\leq |u|^{-1}\}} |x|\, \nu(dx) &\leq (C_0/\theta) \mathbb{1}_{\{|u|\leq a^{-1}\}} |u| \int_a^{|u|^{-1}} |x|^{-\theta}\, dx \\ &\leq C\mathbb{1}_{\{|u|\leq a^{-1}\}} \begin{cases} |u|^\theta & \theta < 1, \\ |u|(\log(1/|u|) + \log(1/a)) & \theta = 1, \\ |u| & \theta > 1. \end{cases} \end{aligned}$$

For the second term on the right-hand side of (I.4.8) we use the following estimate

$$\nu\big(x \in \mathbb{R} : |xu| > 1\big) \leq C(\mathbb{1}_{\{|u|>1\}} + (|u|^{-1})^{-\theta}\mathbb{1}_{\{|u|\leq 1\}}) = C(\mathbb{1}_{\{|u|>1\}} + |u|^{\theta}\mathbb{1}_{\{|u|\leq 1\}})$$

for all $u \in \mathbb{R}$, which completes the proof of (I.4.7) and hence of the lemma. □

Recalling the decomposition $\Delta_{i,k}^n X = M_{i,n,\varepsilon} + R_{i,n,\varepsilon}$ we obtain by Minkowski's inequality

$$\sup_{t\in[0,t_\infty]} \left| \big(n^{\alpha p} V(p;k)_t^n\big)^{\frac{1}{p}} - \left(n^{\alpha p} \sum_{i=k}^{[nt]} |M_{i,n,\varepsilon}|^p\right)^{\frac{1}{p}} \right| \leq \left(n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |R_{i,n,\varepsilon}|^p\right)^{\frac{1}{p}}.$$

Therefore, by virtue of (I.4.2) and (I.4.6), we conclude that

$$n^{\alpha p} V(p;k)_t^n \xrightarrow{\mathcal{L}_{M_1}-s} Z_t \quad \text{on } \Omega_\varepsilon.$$

By letting $\varepsilon \to 0$ we conclude that Theorem I.1.1 (i) holds, when L is a compound Poisson process with jumps bounded away from 0.

Decomposition into big and small jumps

In this section we extend the proof of Theorem I.1.1 (i) to general symmetric Lévy processes $(L_t)_{t\in\mathbb{R}}$. We need the following preliminary result.

Lemma I.4.5. *Let $q \geq 1$ and $a \in (0,1]$. The function*

$$\xi(y) = \int_{-a}^{a} |yx|^2 \mathbb{1}_{\{|yx|\leq 1\}} + |yx|^q \mathbb{1}_{\{|yx|>1\}} \nu(dx)$$

satisfies $|\xi(y)| \leq C(|y|^2\mathbb{1}_{\{|y\leq 1|\}} + |y|^{\beta'\vee q}\mathbb{1}_{\{|y>1|\}})$ for any $\beta' > \beta$, where C does not depend on a.

Proof. Use the decomposition $\xi = \xi_1 + \xi_2$ with

$$\xi_1(y) = \int_{-a}^{a} |yx|^2 \mathbb{1}_{\{|yx|\leq 1\}}\, \nu(dx), \quad \text{and} \quad \xi_2(y) = \int_{-a}^{a} |yx|^q \mathbb{1}_{\{|yx|>1\}}\, \nu(dx).$$

We obtain

$$\xi_1(y)\mathbb{1}_{\{|y|\leq 1\}} \leq |y|^2 \int_{-1}^{1} x^2 \nu(dx) \mathbb{1}_{\{|y|\leq 1\}},$$

and $\xi_1(y)\mathbb{1}_{\{|y|>1\}} \leq C|y|^{\beta'\vee q}\mathbb{1}_{\{|y|>1\}}$ follows from (I.3.4), showing that ξ_1 satisfies the estimate given in the lemma. For $q > \beta$ we obtain

$$\xi_2(y) = 2|y|^q \mathbb{1}_{\{|y|>1/a\}} \int_{1/|y|}^{a} |x|^q \nu(dx) \leq C|y|^q \mathbb{1}_{\{|y|\geq 1\}}.$$

If $q \leq \beta$ we have similarly for any $\beta' > \beta$

$$\xi_2(y) \leq 2|y|^{\beta'} \mathbb{1}_{\{|y|>1/a\}} \int_{1/|y|}^{a} |x|^{\beta'} \nu(dx) \leq C|y|^{\beta'} \mathbb{1}_{\{|y|\geq 1\}},$$

which completes the proof. □

Now, given a general symmetric Lévy process $(L_t)_{t\in\mathbb{R}}$, consider for $a > 0$ the compound Poisson process $(L_t^{>a})_{t\in\mathbb{R}}$ defined by

$$L_0^{>a} = 0, \quad L_t^{>a} - L_s^{>a} = \sum_{s<u\leq t} \Delta L_u \mathbb{1}_{\{|\Delta L_u|>a\}}.$$

Moreover, let $(L_t^{\leq a})_{t\in\mathbb{R}}$ denote the Lévy process $(L_t - L_t^{>a})_{t\in\mathbb{R}}$. The key result of this section is showing that

$$\limsup_{n\to\infty} \Big\| n^{\alpha p} \sum_{i=k}^{[nt_\infty]} \Big| \int_{-\infty}^{i/n} g_{i,n}(s)\sigma_{s-}\, dL_s^{\leq a} \Big|^p \Big\|_1 \to 0, \quad \text{as } a \to 0. \tag{I.4.9}$$

We make the decomposition

$$\int_{-\infty}^{i/n} g_{i,n}(s)\sigma_{s-}\, dL_s^{\leq a} = A_{i,n} + B_{i,n},$$

where

$$A_{i,n} = \int_{-\delta}^{i/n} g_{i,n}(s)\sigma_{s-}\, dL_s^{\leq a} \quad \text{and} \quad B_{i,n} = \int_{-\infty}^{-\delta} g_{i,n}(s)\sigma_{s-}\, dL_s^{\leq a}.$$

Lemma I.3.1 shows that

$$\begin{aligned}
\Big\| n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |A_{i,n}|^p \Big\|_1 &= n^{-1} \sum_{i=k}^{[nt_\infty]} \Big\| \int_{-\delta}^{i/n} n^{\alpha+1/p} g_{i,n}(s)\sigma_{s-}\, dL_s^{\leq a} \Big\|_p^p \\
&\leq C n^{-1} \sum_{i=k}^{[nt_\infty]} \mathbb{E}\Big[\|F^{i,n}\|_{p,L^{\leq a}}^p \Big],
\end{aligned}$$

where the process $(F_t^{i,n})_{t\in\mathbb{R}}$ is defined as $F_t^{i,n} = n^{\alpha+1/p} g_{i,n}(t)\mathbb{1}_{(-\delta,i/n]}(t)\sigma_{t-}$. Since the random variable $\sup_{t\in[-\delta,\infty)} |\sigma_t|$ is uniformly bounded, we obtain by (I.3.6) and [12, Eq.(4.23)]

$$\begin{aligned}
\mathbb{E}\big[\|F^{i,n}\|_{p,L^{\leq a}}^p\big] &\leq C\|n^{\alpha+1/p} g_{i,n}\mathbb{1}_{[-\delta,i/n]}\|_{p,L^{\leq a}}^p \\
&\leq C|\Phi_{p,L^{\leq a}}(n^{\alpha+1/p} g_{k,n})|^{p/2} \vee |\Phi_{p,L^{\leq a}}(n^{\alpha+1/p} g_{k,n})| \\
&\leq C\Big(\int_{|x|\leq a} |x|^p + x^2 \nu(dx) \Big)^{p/2} \vee \Big(\int_{|x|\leq a} |x|^p + x^2 \nu(dx) \Big),
\end{aligned}$$

for all $n \in \mathbb{N}$ and $i \in \{k, \ldots, [nt_\infty]\}$. Since $p > \beta$ by assumption, we conclude that

$$\limsup_{n\to\infty} \Big\| n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |A_{i,n}|^p \Big\|_1 \to 0, \quad \text{as } a \to 0. \tag{I.4.10}$$

Next, we show that for all $a > 0$

$$\limsup_{n\to\infty} \Big\| n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |B_{i,n}|^p \Big\|_1 = 0. \tag{I.4.11}$$

Introducing the processes $(Y_t^{i,n})_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ defined as

$$Y_t^{i,n} = n^{\alpha+1/p} g_{i,n}(t)\sigma_{t-}\mathbb{1}_{(-\infty,-\delta]}(t), \quad \text{and} \quad Y_t = |g^{(k)}(-t)\sigma_{t-}\mathbb{1}_{(-\infty,-\delta]}(t)|,$$

we obtain by Lemma I.3.1 that

$$\Big\| n^{\alpha p} \sum_{i=k}^{[nt_\infty]} |B_{i,n}|^p \Big\|_1 \le C n^{-1} \sum_{i=k}^{[nt_\infty]} \mathbb{E}\big[\|Y^{i,n}\|^p_{p,L^{\le a}}\big].$$

Moreover, recalling that $|g^{(k)}|$ is decreasing on (δ,∞), an application of Lemma I.4.1 shows that

$$\mathbb{E}\big[\|Y^{i,n}\|^p_{p,L^{\le a}}\big] \le n^{p(\alpha+1/p-k)}\mathbb{E}\big[\|Y\|^p_{p,L^{\le a}}\big],$$

for all $i\in\{k,\dots,n\}$. Since $\alpha+1/p-k<0$, equation (I.4.11) follows if $\mathbb{E}\big[\|Y\|^p_{p,L^{\le a}}\big]<\infty$. Applying the estimate (I.3.6) shows that this is satisfied if $\mathbb{E}\big[\Phi^{1\vee\frac{p}{2}}_{p,L^{\le a}}(Y)\big]<\infty$, which is a consequence of (B1) and Lemma I.4.5, where we used that $p>\beta$. Now, (I.4.9) follows from (I.4.10) and (I.4.11).

We can complete the proof of Theorem I.1.1 (i) by combining (I.4.9) with the results of Subsection I.4. To this end, let

$$X_t^{>a} := \int_{-\infty}^t (g(t-s)-g_0(-s))\sigma_{s-}\, dL_s^{>a}, \quad X_t^{\le a} := \int_{-\infty}^t (g(t-s)-g_0(-s))\sigma_{s-}\, dL_s^{\le a},$$

and let $T_m^{>a} = T_m$ if $|\Delta L_{T_m}|>a$, and $T_m^{>a}=\infty$ else. The results of Subsection I.4 show that

$$n^{\alpha p} V(X^{>a},p;k)_t^n \xrightarrow{\mathcal{L}_{M_1}-s} Z_t^{>a} := \sum_{m: T_m^{>a}\in(0,t]} |\Delta L_{T_m^{>a}}\sigma_{T_m^{>a}-}|^p V_m$$

for all $a>0$, where $V(X^{>a},p;k)_t^n$ denotes the power variation of the process $X^{>a}$. Making the decomposition

$$\begin{aligned}
&\big(n^{\alpha p}V(p;k)_t^n\big)^{1/p} \\
&\qquad = \big(n^{\alpha p}V(X^{>a},p;k)_t^n\big)^{1/p} + \Big(\big(n^{\alpha p}V(p;k)_t^n\big)^{1/p} - \big(n^{\alpha p}V(X^{>a},p;k)_t^n\big)^{1/p}\Big) \\
&\qquad := U_t^{n,>a} + U_t^{n,\le a},
\end{aligned}$$

we have by Minkowski's inequality

$$\lim_{a\to 0}\limsup_{n\to\infty}\mathbb{P}\big(\sup_{t\in[0,t_\infty]}|U_t^{n,\le a}|>\varepsilon\big) \le \lim_{a\to 0}\limsup_{n\to\infty}\mathbb{P}\big(n^{\alpha p}V(X^{\le a},p;k)^n_{t_\infty}>\varepsilon^p\big)=0,$$

for all $\varepsilon>0$, which follows easily from (I.4.9). Since $U_t^{n,>a}\xrightarrow{\mathcal{L}_{M_1}-s} Z_t^{>a}$ as $n\to\infty$, and $\sup_{t\in[0,t_\infty]}|Z_t^{>a}-Z_t|\to 0$ almost surely, as $a\to 0$, Theorem I.1.1 (i) follows from [15, Theorem 3.2]. □

Remark I.4.6. A popular technique for proving limit theorems for volatility modulated processes is to freeze the volatility over blocks of length $1/n$ and derive a limit theorem for the resulting simpler process. However, in the framework of Theorem I.1.1 (i) this approach is not applicable, since the power variations of the two processes are not asymptotically equivalent if σ and L jump at the same times.

Proof of Theorem I.1.1 (ii)

Since $t \mapsto V(p;k)_t^n$ is increasing and the limiting function is continuous, uniform convergence on compact sets in probability follows if we show

$$n^{-1+p(\alpha+1/\beta)} V(p;k)_t^n \xrightarrow{\mathbb{P}} m_p \int_0^t |\sigma_s|^p ds$$

for a fixed $t > 0$, which we will do in the following. A crucial step in the proof is to show that the asymptotic behavior of the power variation does not change if we replace $\Delta_{i,k}^n X$ in (I.1.3) by $\sigma_{(i-k)/n} \Delta_{i,k}^n G$, where the process $(G_t)_{t \geq 0}$ is defined as the integral in (I.1.2) with $\sigma \equiv 1$. Note that assumption (A) ensures that G is well-defined. Thereafter, we divide the interval $[0,t]$ into subblocks of size $1/l$ and freeze σ at the beginning of each block. The limiting power variation for the resulting process can then be derived by applying part (ii) of [12, Theorem 1.1] on every block. The proof of Theorem I.1.1 (ii) is then completed by letting $l \to \infty$. The following lemma plays an important role for replacing $\Delta_{i,k}^n X$ in (I.1.3) by $\sigma_{(i-k)/n} \Delta_{i,k}^n G$. Here and in the following we denote by v_σ the modulus of continuity of σ defined as

$$v_\sigma(s,\eta) = \sup\{|\sigma_s - \sigma_r| : r \in [s-\eta, s+\eta]\}.$$

Lemma I.4.7. *Let $(\sigma_t)_{t \in \mathbb{R}}$ be a process with càdlàg or càglàd sample paths that is uniformly bounded on $[-\delta, \infty)$. For any $\alpha, q \in (0,\infty)$ we have*

$$\lim_{\varepsilon \to 0} \left[\limsup_{n \to \infty} \left(\frac{1}{n} \sum_{i=k}^{[nt]} \|v_\sigma(i/n, \varepsilon)\|_q^\alpha \right) \right] = 0.$$

Proof. Since v_σ is bounded and $x \mapsto x^\alpha$ is locally Lipschitz for $\alpha > 1$, we may assume w.l.o.g. that $\alpha \leq 1$ and $q \geq 1$. For $\kappa > 0$ we use the decomposition $\sigma = \sigma^{<\kappa} + \sigma^{\geq \kappa}$, where

$$\sigma_s^{\geq \kappa} = \sum_{-\delta < u \leq s} \Delta\sigma_u \mathbb{1}_{\{|\Delta \sigma_u| \geq \kappa\}},$$

and $\sigma_s^{<\kappa} = \sigma_s - \sigma_s^{\geq \kappa}$. Even though σ is uniformly bounded on $[-\delta, \infty)$, $\sigma^{\geq \kappa}$ and $\sigma^{<\kappa}$ might not be. For this reason we introduce the sets

$$\begin{aligned} \Omega_m := \big\{ \omega \,:\, & |\sigma_s^{<\kappa}(\omega)| + |\sigma_s^{\geq\kappa}(\omega)| \leq m \text{ for all } s \in [-\delta, t+\delta], \\ & \text{and } \sigma^{\geq \kappa}(\omega) \text{ has less than } m \text{ jumps in } [-\delta, t+\delta] \big\}. \end{aligned}$$

Note that $\Omega_m \uparrow \Omega$, as $m \to \infty$. By the triangular inequality we have

$$v_\sigma(s,\eta) \leq v_{\sigma^{<\kappa}}(s,\eta) \mathbb{1}_{\Omega_m} + v_{\sigma^{\geq\kappa}}(s,\eta)\mathbb{1}_{\Omega_m} + C \mathbb{1}_{\Omega_m^{\mathrm{c}}},$$

for all $s \in [0,t], \eta < \delta$ and $m \geq 1$. Since $\mathbb{P}(\Omega_m^{\mathrm{c}}) \to 0$ as $m \to \infty$, we can choose m sufficiently large such that

$$\frac{1}{n} \sum_{i=k}^{[nt]} \|v_\sigma(i/n,\varepsilon)\|_q^\alpha \leq \frac{1}{n} \sum_{i=k}^{[nt]} \left(\|v_{\sigma^{<\kappa}}(i/n,\varepsilon) \mathbb{1}_{\Omega_m}\|_q^\alpha + \|v_{\sigma^{\geq\kappa}}(i/n,\varepsilon)\mathbb{1}_{\Omega_m}\|_q^\alpha \right) + \kappa, \quad \text{(I.4.12)}$$

for all $n \in \mathbb{N}$ and $\varepsilon > 0$. We show that

$$\limsup_{\varepsilon\to 0} \limsup_{n\to\infty} \left(\frac{1}{n} \sum_{i=k}^{[nt]} \|v_{\sigma^{<\kappa}}(i/n, \varepsilon) \mathbb{1}_{\Omega_m}\|_q^\alpha \right) \leq 2\kappa^\alpha. \tag{I.4.13}$$

In order to do so, we assume the existence of sequences $(\varepsilon_l), (n_l), (i_l)$ with $\varepsilon_l \to 0$, $n_l \to \infty$ and $i_l \in \{1, ..., [tn_l]\}$ such that

$$\|v_{\sigma^{<\kappa}}(i_l/n_l, \varepsilon_l) \mathbb{1}_{\Omega_m}\|_q^\alpha > 2\kappa^\alpha \tag{I.4.14}$$

for all l, and derive a contradiction. Since $(i_l/n_l)_{l\geq 1}$ is a bounded sequence we may assume that i_l/n_l converges to some $s_0 \in [0, t]$ by considering a suitable subsequence $(l_k)_{k\geq 1}$. For all $\omega \in \Omega_m$ it holds that $\lim_{\gamma\to 0} v_{\sigma^{<\kappa}}(s_0, \gamma) = |\Delta\sigma_{s_0}^{<\kappa}| \leq \kappa$. Therefore, by the dominated convergence theorem, we can find a $\gamma > 0$ such that $\|v_{\sigma^{<\kappa}}(s_0, \gamma) \mathbb{1}_{\Omega_m}\|_q^\alpha \leq 2\kappa^\alpha$. This is a contradiction to (I.4.14), since for sufficiently large l we have $[i_l/n_l - \varepsilon_l, i_l/n_l + \varepsilon_l] \subset [s_0 - \gamma, s_0 + \gamma]$. This completes the proof of (I.4.13). Next, we show that

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left(\frac{1}{n} \sum_{i=k}^{[nt]} \|v_{\sigma^{\geq\kappa}}(i/n, \varepsilon) \mathbb{1}_{\Omega_m}\|_q^\alpha \right) = 0. \tag{I.4.15}$$

Recalling that $q/\alpha \geq 1$, an application of Jensen's inequality yields

$$\frac{1}{n} \sum_{i=k}^{[nt]} \|v_{\sigma^{\geq\kappa}}(i/n, \varepsilon) \mathbb{1}_{\Omega_m}\|_q^\alpha \leq \left\| t^{q/\alpha - 1} \frac{1}{n} \sum_{i=k}^{[nt]} \left(v_{\sigma^{\geq\kappa}}(i/n, \varepsilon) \mathbb{1}_{\Omega_m} \right)^q \right\|_1^{\alpha/q},$$

for all $n \in \mathbb{N}$, $\varepsilon > 0$. Now, (I.4.15) follows from the estimate

$$\frac{1}{n} \sum_{i=k}^{[nt]} \left(v_{\sigma^{\geq\kappa}}(i/n, \varepsilon) \mathbb{1}_{\Omega_m} \right)^q \leq \sup_{s\in[-\delta, t+\delta]} |\Delta\sigma_s^{\geq\kappa}|^q N \mathbb{1}_{\Omega_m} 2(\varepsilon) \leq C m^{q+1}(\varepsilon),$$

for all $n \in \mathbb{N}$. Here $N = N(\omega)$ denotes the number of jumps of $\sigma^{\geq\kappa}$ in $[-\delta, t + \delta]$. Using (I.4.13) and (I.4.15), the lemma now follows from (I.4.12) by letting $\kappa \to 0$. □

The proof of Theorem I.1.1 (ii) heavily relies on the estimate given in Lemma I.3.2. This lemma assumes the role that Itô's isometry typically plays in proofs of limit theorems for stochastic integral processes driven by a Brownian motion. In order to apply Lemma I.3.2, the following estimates will be crucial.

Lemma I.4.8. *Suppose that assumptions (A) and (B2) hold, and assume that $\alpha + 1/\beta < k$. For $\varepsilon > 0$ with $\varepsilon \leq \delta$ there is a constant $C > 0$ such that*

$$\mathbb{E}\left[\int_{\frac{i}{n}-\varepsilon}^{\frac{i}{n}} |g_{i,n}(s)\sigma_{s-}|^\beta \, ds \right] + \int_{\frac{i}{n}-\varepsilon}^{\frac{i}{n}} |g_{i,n}(s)|^\beta \, ds \leq C n^{-\alpha\beta - 1}, \quad \textit{and}$$
$$\mathbb{E}\left[\int_{-\infty}^{\frac{i}{n}-\varepsilon} |g_{i,n}(s)\sigma_{s-}|^\beta \, ds \right] + \int_{-\infty}^{\frac{i}{n}-\varepsilon} |g_{i,n}(s)|^\beta \, ds \leq C n^{-k\beta},$$

for all $i \in \{k, \ldots, n\}$.

Proof. By Lemma I.4.1 we have that

$$
\begin{aligned}
&|g_{i,n}(s)|^{\beta}\mathbb{1}_{[i/n-\varepsilon,i/n]}(s) \\
&\le C\big((i/n-s)^{\alpha\beta}\mathbb{1}_{[(i-k)/n,i/n]}(s)+n^{-k\beta}((i-k)/n-s)^{(\alpha-k)\beta}\mathbb{1}_{[i/n-\varepsilon,(i-k)/n]}(s)\big).
\end{aligned}
$$

Recalling that σ is bounded on $[-\delta,\infty)$, the first inequality follows by calculating the integral of the right hand side. The second inequality is a direct consequence of Lemma I.4.1 and assumptions (A) and (B2). □

A crucial step in the proof of Theorem I.1.1 (ii) is showing that

$$
n^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\|\Delta_{i,k}^{n}X-\sigma_{(i-k)/n}\Delta_{i,k}^{n}G\|_p^p\to 0, \tag{I.4.16}
$$

as $n\to\infty$, where the process $(G_t)_{t\ge 0}$ is defined as the integral in (I.1.2) with $\sigma\equiv 1$. We fix some $\varepsilon>0$ and make the decomposition

$$
\Delta_{i,k}^{n}X-\sigma_{(i-k)/n}\Delta_{i,k}^{n}G=A_i^{n,\varepsilon}+B_i^{n,\varepsilon}+C_i^{n,\varepsilon},
$$

where

$$
A_i^{n,\varepsilon}=\int_{i/n-\varepsilon}^{i/n}g_{i,n}(s)(\sigma_{s-}-\sigma_{i/n-\varepsilon})\,dL_s,\quad B_i^{n,\varepsilon}=(\sigma_{i/n-\varepsilon}-\sigma_{(i-k)/n})\int_{i/n-\varepsilon}^{i/n}g_{i,n}(s)\,dL_s,
$$

$$
C_i^{n,\varepsilon}=\int_{-\infty}^{i/n-\varepsilon}g_{i,n}(s)\sigma_{s-}\,dL_s-\sigma_{(i-k)/n}\int_{-\infty}^{i/n-\varepsilon}g_{i,n}(s)\,dL_s.
$$

We deduce (I.4.16) by showing that

$$
\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left(n^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\|A_i^{n,\varepsilon}\|_p^p\right)=0,
$$

and the same for $B_i^{n,\varepsilon}$ and $C_i^{n,\varepsilon}$, respectively. For $A_i^{n,\varepsilon}$ we obtain by Lemma I.3.2

$$
\begin{aligned}
&n^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\|A_i^{n,\varepsilon}\|_p^p \\
&\qquad\le Cn^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\left\{\mathbb{E}\left[\int_{i/n-\varepsilon}^{i/n}\left|g_{i,n}(s)(\sigma_{s-}-\sigma_{i/n-\varepsilon})\right|^{\beta}ds\right]\right\}^{p/\beta} \\
&\qquad\le Cn^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\|v_\sigma(i/n,\varepsilon+1/n)\|_\beta^p\left(\int_{i/n-\varepsilon}^{i/n}|g_{i,n}(s)|^{\beta}\,ds\right)^{p/\beta}.
\end{aligned}
$$

By Lemma I.4.7 and Lemma I.4.8 we conclude that

$$
\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left(n^{-1+p(\alpha+1/\beta)}\sum_{i=k}^{[nt]}\|A_i^{n,\varepsilon}\|_p^p\right)=0. \tag{I.4.17}
$$

For $B_i^{n,\varepsilon}$ we apply Hölder's inequality with p' and q' satisfying $1/p' + 1/q' = 1$ and $pq' < \beta$, which is possible due to our assumption $p < \beta$. This yields

$$\begin{aligned}
&n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} \|B_i^{n,\varepsilon}\|_p^p \\
&\qquad \leq n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} \|(\sigma_{i/n-\varepsilon} - \sigma_{(i-k)/n})\|_{pp'}^p \bigg\| \int_{i/n-\varepsilon}^{i/n} g_{i,n}(s)\, dL_s \bigg\|_{pq'}^p, \\
&\qquad \leq Cn^{-1} \sum_{i=k}^{[nt]} \big\| v_\sigma(i/n, \varepsilon + k/n) \big\|_{pp'}^p.
\end{aligned}$$

Here we have used that, as a consequence of Lemma I.3.2 and Lemma I.4.8, whenever $pq' < \beta$ there exists a $C > 0$ such that $\|n^{\alpha+1/\beta} \int_{i/n-\varepsilon}^{i/n} g_{i,n}(s)\, dL_s\|_{pq'} < C$ for all $n \in \mathbb{N}$, $i \in \{k, ..., [nt]\}$. Thus, by Lemma I.4.7

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \left(n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} \|B_i^{n,\varepsilon}\|_p^p \right) = 0. \tag{I.4.18}$$

Moreover, by Lemma I.3.2 and Lemma I.4.8 it follows that for all $\varepsilon > 0$

$$\limsup_{n \to \infty} \left(n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} \|C_i^{n,\varepsilon}\|_p^p \right) \leq C \limsup_{n \to \infty} (n^{p(\alpha+1/\beta-k)}) = 0,$$

which together with (I.4.17) and (I.4.18) completes the proof of (I.4.16).

By Minkowski's inequality for $p \geq 1$ and subadditivity for $p < 1$, it is now sufficient to show that

$$n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} |\sigma_{(i-1/n)} \Delta_{i,k}^n G|^p \xrightarrow{\mathbb{P}} m_p \int_0^t |\sigma_s|^p\, ds, \tag{I.4.19}$$

in order to prove Theorem I.1.1 (ii).

Intuitively, replacing $|\Delta_{i,k}^n X|$ by $|\sigma_{(i-k)/n} \Delta_{i,k}^n G|$ corresponds to freezing the process $(\sigma_t)_{t \in \mathbb{R}}$ over blocks of length $1/n$. For the proof of (I.4.19) we freeze σ now over small blocks with block size $1/l$ that does not depend on n. This will allow us to apply [12, Theorem 1.1(ii)] on every block. Thereafter, (I.4.19) follows by letting $l \to \infty$. For $l > 0$ we decompose

$$\begin{aligned}
&n^{-1+p(\alpha+1/\beta)} \sum_{i=k}^{[nt]} |\sigma_{(i-k)/n} \Delta_{i,k}^n G|^p - m_p \int_0^t |\sigma_s|^p\, ds \\
&\quad = n^{-1+p(\alpha+1/\beta)} \left(\sum_{i=k}^{[nt]} |\Delta_{i,k}^n G|^p \big(|\sigma_{(i-k)/n}|^p - |\sigma_{(j_{l,i}-1)/l}|^p \big) \right) \\
&\qquad + \left(\sum_{j=1}^{[tl]+1} |\sigma_{(j-1)/l}|^p \left(n^{-1+p(\alpha+1/\beta)} \sum_{i \in I_l(j)} |\Delta_{i,k}^n G|^p - m_p l^{-1} \right) \right) \\
&\qquad + \left(m_p l^{-1} \sum_{j=1}^{[tl]} |\sigma_{(j-1)/l}|^p - m_p \int_0^t |\sigma_s|^p ds \right) \quad := D_{n,l} + E_{n,l} + F_l.
\end{aligned}$$

Here, $j_{l,i}$ denotes the index $j \in \{1, ..., [tl]+1\}$ such that $(i-k)/n \in ((j-1)/l, j/l]$ and $I_l(j)$ is the set of indices i such that $(i-k)/n \in ((j-1)/l, j/l]$. We show that

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}(|D_{n,l} + E_{n,l} + F_l| > \varepsilon) = 0$$

for any $\varepsilon > 0$. Note that $F_l \xrightarrow{\text{a.s.}} 0$ as $l \to \infty$, since the Riemann integral of any càdlàg function exists. For every $l \in \mathbb{N}$ we have $\limsup_{n\to\infty} \mathbb{P}(|E_{n,l}| > \varepsilon) = 0$ by [12, Theorem 1.1(ii)]. For $\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}(|D_{n,l}| > \varepsilon) = 0$ we argue as follows. Choose some $p' > 1$ such that $pp' < \beta$ and let q' be such that $1/p' + 1/q' = 1$. We find

$$\begin{aligned}
\|D_{n,l}\|_1 &= \left\| n^{-1+p(\alpha+1/\beta)} \left(\sum_{i=k}^{[nt]} |\Delta_{i,k}^n G|^p (|\sigma_{(i-k)/n}|^p - |\sigma_{(j_{l,n,i}-1)/l}|^p) \right) \right\|_1 \\
&\le n^{-1} \sum_{i=k}^{[nt]} \| |n^{\alpha+1/\beta} \Delta_{i,k}^n G|^p \|_{p'} \| |\sigma_{(i-k)/n}|^p - |\sigma_{(j_{l,n,i}-1)/l}|^p \|_{q'} \\
&\le \left(n^{-1} \sum_{i=k}^{[nt]} \| n^{\alpha+1/\beta} \Delta_{i,k}^n G \|_{pp'}^{2/p} \right)^{1/2} \left(n^{-1} \sum_{i=k}^{[nt]} \| |\sigma_{(i-k)/n}|^p - |\sigma_{(j_{l,n,i}-1)/l}|^p \|_{q'}^2 \right)^{1/2}.
\end{aligned}$$

The first factor is bounded by Lemmas I.3.2 and I.4.8. For the second factor we can apply Lemma I.4.7, since the process $(|\sigma_t|^p)_{t\in\mathbb{R}}$ is càdlàg and bounded on $[-\delta, \infty)$, and conclude that $\lim_{l\to\infty} \limsup_{n\to\infty} \|D_{n,l}\|_1 = 0$. This completes the proof of (I.4.19), and hence of Theorem I.1.1 (ii). □

Proof of Theorem I.1.1 (iii)

For the proof of Theorem I.1.1 (iii) we show that under the conditions of the theorem the process X admits a modification with k-times differentiable sample paths with k-th derivative F, as defined in the theorem. Then the result follows by an application of the following stochastic Fubini theorem. For a proof we refer to [1, Theorem 3.1], where a similar Fubini theorem was shown for deterministic integrands. The generalisation towards predictable integrands is straightforward.

Lemma I.4.9. *Let $f : \mathbb{R} \times \mathbb{R} \times \Omega \to \mathbb{R}$ be a random field that is measurable with respect to the product σ-algebra $\mathcal{B}(\mathbb{R}) \otimes \Pi$, where Π denotes the $(\mathcal{F}_t)_{t\in\mathbb{R}}$-predictable σ-algebra on $\mathbb{R} \times \Omega$. That is, Π is the σ-algebra generated by all sets $A \times (s,t]$, where $s < t$ and $A \in \mathcal{F}_s$. Let $(L_t)_{t\in\mathbb{R}}$ be a symmetric Lévy process that has finite first moment. Assume that we have*

$$\mathbb{E}\left[\int_{\mathbb{R}} \|f(u,\cdot)\|_{1,L} \, du \right] < \infty.$$

Then, we obtain

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(u,s) \, du \right) dL_s = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(u,s) \, dL_s \right) du \quad \textit{almost surely,}$$

and all the integrals are well-defined.

The following auxiliary result ensures that the conditions of this lemma are satisfied in our framework.

Lemma I.4.10. *Suppose that assumption (B1) holds. Let $q \in \{1, p\}$ which in particular implies $\alpha > k - 1/(\beta \vee q)$. If $q > 1$ assume additionally that the jumps of the Lévy process L are bounded in absolute value by 1. For any $t > 0$, the random field $f_t(u,s) := g^{(k)}(u-s)\sigma_{s-}\mathbb{1}_{[0,t]}(u)\mathbb{1}_{(-\infty,u)}(s)$ satisfies*

$$\int_0^t \mathbb{E}[\|f_t(u,\cdot)\|_{q,L}^q]\,du < \infty.$$

Proof. We decompose

$$\begin{aligned}
&\int_0^t \mathbb{E}[\|f_t(u,\cdot)\|_{q,L}^q]\,du \\
&\qquad \le C\int_0^t \mathbb{E}[\|f_t(u,\cdot)\mathbb{1}_{(-\delta,t]}\|_{q,L}^q]\,du + C\int_0^t \mathbb{E}[\|f_t(u,\cdot)\mathbb{1}_{(-\infty,-\delta]}\|_{q,L}^q]\,du \\
&\qquad := I_1 + I_2,
\end{aligned}$$

and show that both summands are finite. For I_1 we use that σ is bounded on $[-\delta,\infty)$. Thus, denoting $e_t(u,s) = g^{(k)}(u-s)\mathbb{1}_{[0,t]}(u)\mathbb{1}_{(-\delta,u)}(s)$, we obtain using (I.3.6)

$$I_1 \le C\int_0^t \Phi_{q,L}(e_t(u,\cdot)) + \Phi_{q,L}^{\frac{q}{2}}(e_t(u,\cdot))\,du \le Ct\big(\Phi_{q,L}(e_t(t,\cdot)) + \Phi_{q,L}^{\frac{q}{2}}(e_t(t,\cdot))\big),$$

where in the second inequality we used $|e_t(u,s)| \le |e_t(t, s+t-u)|$, and that $\Phi_{q,L}(f)$ is invariant under shifting the argument of the function f. For I_1 to be finite it is therefore sufficient to show that the following term is finite

$$\int_{-\delta}^t \int_{\mathbb{R}} |g^{(k)}(t-s)x|^2 \mathbb{1}_{\{|g^{(k)}(t-s)x|\le 1\}} + |g^{(k)}(t-s)x|^q \mathbb{1}_{\{|g^{(k)}(t-s)x|>1\}}\,\nu(dx)\,ds := J_1 + J_2.$$

We fix $\beta' \in (\beta\vee 1, 1/(k-\alpha))$ and $q' \in [q, 1/(k-\alpha))$ such that the Lévy process satisfies $\mathbb{E}[|L_1|^{q'}] < \infty$. Indeed, the former is possible by the conditions $\alpha > k - 1/(\beta \vee p)$ and $p \ge 1$ in Theorem I.1.1(iii). The latter is possible for $q = 1$ by the assumption $\theta > 1$ in Theorem I.1.1 (iii), and for $q = p > 1$ by the assumption of bounded jumps in the lemma. Recalling that $|g^{(k)}(t)| \le C|t|^{\alpha-k}$ for all $t \in (0,\delta)$, in order to show $J_1 + J_2 < \infty$, it is then sufficient to show

$$J_1 + J_2 \le C\bigg(1 + \int_{-\delta}^t |g^{(k)}(t-s)|^{\beta'}\,ds + \int_{-\delta}^t |g^{(k)}(t-s)|^{q'}\,ds\bigg). \tag{I.4.20}$$

For $q = p > 1$, this estimate follows easily from Lemma I.4.5, where we use the assumption that L has jumps bounded by 1. For $q = 1$ the estimate follows for J_1 by (I.3.4). For J_2 we obtain

$$\begin{aligned}
J_2 &\le \int_{-\delta}^t \int_{-1}^1 |g^{(k)}(t-s)x|^{\beta'} \mathbb{1}_{\{|g^{(k)}(t-s)x|>1\}}\,\nu(dx)\,ds \\
&\quad + 2\int_{-\delta}^t |g^{(k)}(t-s)|^{q'}ds \int_1^\infty |x|^{q'}\nu(dx) \\
&\le C\int_{-\delta}^t |g^{(k)}(t-s)|^{\beta'} \mathbb{1}_{\{|g^{(k)}(t-s)|>1\}} + |g^{(k)}(t-s)|^{q'}\,ds,
\end{aligned}$$

which concludes the proof of (I.4.20) and of $I_1 < \infty$. For I_2 we use that $|g^{(k)}|$ is decreasing on (δ, ∞), which implies that $I_2 \leq Ct\mathbb{E}[\|f_t(0,\cdot)\mathbb{1}_{(-\infty,-\delta]}\|^q_{q,L}]$. By (I.3.6) the latter is finite if $\Phi^{1\vee\frac{q}{2}}_{q,L}(f_t(0,\cdot)\mathbb{1}_{(-\infty,-\delta]}) \in L^1(\Omega)$. This follows easily from Assumption (B2) (recall that $q \leq p$) and (I.3.4). $\square$

With these preliminaries at hand, we can finally prove Theorem I.1.1 (iii). As remarked at the beginning of Subsection I.4, it is sufficient to show convergence in probability for a fixed $t > 0$ in order to obtain uniform convergence on compacts in probability. Therefore, the theorem is an immediate consequence of the following result and Lemma 4.3 in [12].

Lemma I.4.11. *Under the conditions of Theorem I.1.1 (iii), there is a process $(Z_t)_{t\geq 0}$ that satisfies almost surely $V(Z,p;k)^n_t = V(X,p;k)^n_t$ for all $n \in \mathbb{N}$ and $t \geq 0$, has almost surely k-times absolutely continuous sample paths and satisfies for Lebesgue almost all $t \geq 0$ that*

$$\frac{\partial^k Z_t}{(\partial t)^k} = \int_{-\infty}^t g^{(k)}(t-s)\sigma_{s-}\, dL_s := F_t,$$

and $F \in L^p([0,t_0])$ for any $t_0 > 0$.

Proof. For ease of notation we only consider $k = 1$. The general case follows by similar arguments. We let $a \in (0,1]$ and define the processes $(F^{\leq a}_u)_{u\in\mathbb{R}}$ and $(F^{>a}_u)_{u\in\mathbb{R}}$ by

$$F^{\leq a}_u = \int_{-\infty}^u g'(u-s)\sigma_{s-}\, dL^{\leq a}_s, \quad \text{and} \quad F^{>a}_u = \sum_{s\in(-\infty,u)} g'(u-s)\sigma_{s-}\Delta L_s \mathbb{1}_{\{|\Delta L_s|>a\}},$$

where the process $(L^{\leq a}_t)_{t\in\mathbb{R}}$ is the truncated Lévy process introduced in Section I.4. We show that both processes $F^{\leq a}_u$ and $F^{>a}_u$ are well-defined and that they both admit a modification with sample paths in $L^p([0,t])$. Then, we define the process

$$Z_t := \int_0^t (F^{\leq a}_u + F^{>a}_u)\, du,$$

and show that it satisfies the properties given in the lemma.

We begin by analysing $F^{\leq a}_u$. It is well-defined, since, as a consequence of Lemma I.4.10, $f_{t_0}(u,s) = g'(u-s)\sigma_{s-}\mathbb{1}_{[0,t_0]}(u)\mathbb{1}_{(-\infty,u)}(s)$ is integrable in s with respect to $L^{\leq a}$ for Lebesgue almost all u. Applying Lemmas I.3.1 and I.4.10 we obtain $F^{\leq a} \in L^p([0,t])$, almost surely, since

$$\mathbb{E}\Big[\int_0^t |F^{\leq a}_u|^p\, du\Big] \leq C\int_0^t \mathbb{E}[\|f_t(u,\cdot)\|^p_{p,L^{\leq a}}]\, du < \infty.$$

For the process $F^{>a}_u$ we make the decomposition

$$\begin{aligned} F^{>a}_u &= F^{>a,\leq-\delta}_u + F^{>a,>-\delta}_u \\ &= \sum_{s\in(-\infty,-\delta]} g'(u-s)\sigma_{s-}\Delta L_s\mathbb{1}_{\{|\Delta L_s|>a\}} + \sum_{s\in(-\delta,u)} g'(u-s)\sigma_{s-}\Delta L_s\mathbb{1}_{\{|\Delta L_s|>a\}}. \end{aligned}$$

We argue first that $F^{>a,\leq -\delta}$ is well-defined and in $L^p([0,t])$ almost surely. Applying Lemma I.4.4 we obtain that

$$\sum_{s\in(-\infty,-\delta]} |g'(-s)\sigma_{s-}\Delta L_s|\mathbb{1}_{\{|\Delta L_s|>a\}} < \infty$$

almost surely. Since $|g'|$ is decreasing on $[\delta,\infty)$, this implies that $F^{>a,\leq -\delta}$ is well-defined and uniformly bounded in u. For $F_u^{>a,>-\delta}$ we use that L has only finitely many jumps of size $> a$ on $[-\delta,t]$. Therefore, $F^{>a,>-\delta}$ is well-defined and we find a positive random variable $K<\infty$ such that

$$\begin{aligned}\int_0^t |F_u^{>a,>-\delta}|^p \, du &\leq K\int_0^t \sum_{s\in(-\delta,u)} \big|g'(u-s)\sigma_{s-}\Delta L_s \mathbb{1}_{\{|\Delta L_s|>a\}}\big|^p \, du \\ &\leq K \sum_{s\in(-\delta,t)} \big|\sigma_{s-}\Delta L_s \mathbb{1}_{\{|\Delta L_s|>a\}}\big|^p \int_0^t |g'(u-s)|^p \, du,\end{aligned}$$

which is finite since $|g'(s)|\leq Cs^{\alpha-1}$ for $s\in(0,\delta)$ and $(\alpha-1)p>-1$. All that remains to show is that $V(X,p;1)_t^n = V(Z,p;1)_t^n$ for all $n\in\mathbb{N}$ and all $t>0$ with probability 1. For any $t>0$ it holds with probability 1 that

$$X_t - X_0 = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f_t(u,s)\, du\right) dL_s = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f_t(u,s)\, dL_s\right) du = Z_t,$$

where we have applied Lemmas I.4.9 and I.4.10. Consequently, it holds that $\mathbb{P}[X_t = Z_t + X_0 \text{ for all } t\in\mathbb{Q}_+] = 1$ which implies $V(X,p;1)_t^n = V(Z,p;1)_t^n$ for all $n\in\mathbb{N}$ and all $t>0$ almost surely. $\square$

Bibliography

[1] Barndorff-Nielsen, O. and A. Basse-O'Connor (2011). Quasi Ornstein-Uhlenbeck processes. *Bernoulli 17*(3), 916–941.

[2] Barndorff-Nielsen, O., F. Benth, and A. Veraart (2011). Modelling electricity forward markets by ambit fields. *Available at https://ssrn.com/abstract=1938704*.

[3] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2009). Power variation for Gaussian processes with stationary increments. *Stochastic Process. Appl. 119*(6), 1845–1865.

[4] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2011). Multipower variation for Brownian semistationary processes. *Bernoulli 17*(4), 1159–1194.

[5] Barndorff-Nielsen, O., J. Corcuera, M. Podolskij, and J. Woerner (2009). Bipower variation for Gaussian processes with stationary increments. *J. Appl. Probab. 46*(1), 132–150.

[6] Barndorff-Nielsen, O., J. M. Corcuera, and M. Podolskij (2013). Limit theorems for functionals of higher order differences of Brownian semi-stationary processes. In *Prokhorov and contemporary probability theory*, Volume 33 of *Springer Proc. Math. Stat.*, pp. 69–96. Springer, Heidelberg.

[7] Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From stochastic calculus to mathematical finance*, pp. 33–68. Springer, Berlin.

[8] Barndorff-Nielsen, O., M. Pakkanen, and J. Schmiegel (2014). Assessing relative volatility/intermittency/energy dissipation. *Electron. J. Stat. 8*(2), 1996–2021.

[9] Barndorff-Nielsen, O. and J. Schmiegel (2007). Ambit processes; with applications to turbulence and tumour growth. In *Stochastic analysis and applications*, pp. 93–124. Springer.

[10] Barndorff-Nielsen, O. and J. Schmiegel (2008). Time change, volatility, and turbulence. In *Mathematical control theory and finance*, pp. 29–53. Springer, Berlin.

[11] Barndorff-Nielsen, O. and J. Schmiegel (2009). Brownian semistationary processes and volatility/intermittency. In *Advanced financial modelling*, Volume 8 of *Radon Ser. Comput. Appl. Math.*, pp. 1–25. Walter de Gruyter, Berlin.

[12] Basse-O'Connor, A., R. Lachiéze-Rey, and M. Podolskij (2016). Power variation for a class of stationary increments levy driven moving averages. *Annals of Probability*. To appear.

[13] Basse-O'Connor, A. and M. Podolskij (2017). On critical cases in limit theory for stationary increments Lévy driven moving averages. *Stochastics 89*(1), 360–383.

[14] Benassi, A., S. Cohen, and J. Istas (2004). On roughness indices for fractional fields. *Bernoulli 10*(2), 357–373.

[15] Billingsley, P. (1999). *Convergence of probability measures* (Second ed.). John Wiley & Sons, Inc., New York.

[16] Chronopoulou, A., F. Viens, and C. Tudor (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electron. J. Stat. 3*, 1393–1435.

[17] Coeurjolly, J. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process. 4*(2), 199–227.

[18] Dang, T. and J. Istas (2015). Estimation of the hurst and the stability indices of a h-self-similar stable process. Working paper. Available at arXiv:1506.05593.

[19] Gärtner, K. and M. Podolskij (2015). On non-standard limits of Brownian semi-stationary processes. *Stochastic Process. Appl. 125*(2), 653–677.

[20] Grahovac, D., N. Leonenko, and M. Taqqu (2015). Scaling properties of the empirical structure function of linear fractional stable motion and estimation of its parameters. *J. Stat. Phys. 158*(1), 105–119.

[21] Guyon, X. and J. León (1989). Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbf{R}$. *Ann. Inst. H. Poincaré Probab. Statist. 25*(3), 265–282.

[22] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic processes and their applications 118*(4), 517–559.

[23] Jacod, J. and P. Protter (2012). *Discretization of processes*. Springer, Heidelberg.

[24] Jacod, J. and A. Shiryaev (2003). *Limit theorems for stochastic processes* (Second ed.). Springer-Verlag, Berlin.

[25] Jensen, E., K. Jonsdottir, J. Schmiegel, and O. Barndorff-Nielsen (2006). Spatio-temporal modelling-with a view to biological growth. *Monographs on statistics and applied probability 107*, 47.

[26] Kwapień, S. and W. Woyczyński (1992). *Random series and stochastic integrals: single and multiple*. Birkhäuser Boston, Inc., Boston, MA.

[27] Musielak, J. (1983). *Orlicz spaces and modular spaces.* Springer-Verlag, Berlin.

[28] Nourdin, I. and A. Réveillac (2009). Asymptotic behavior of weighted quadratic variations of fractional Brownian motion: the critical case $H = 1/4$. *Ann. Probab. 37*(6), 2200–2230.

[29] Podolskij, M. and M. Vetter (2010). Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl. 64*(3), 329–351.

[30] Rajput, B. and J. Rosiński (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields 82*(3), 451–487.

[31] Rosiński, J. and W. Woyczyński (1986). On Itô stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab. 14*(1), 271–286.

[32] Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge.

[33] Skorohod, A. (1956). Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen. 1*, 289–319.

[34] Tudor, C. A. and F. G. Viens (2009). Variations and estimators for self-similarity parameters via Malliavin calculus. *Ann. Probab. 37*(6), 2093–2134.

[35] Whitt, W. (2002). *Stochastic-process limits.* Springer-Verlag, New York.

Paper II

On limit theory for functionals of stationary increments Lévy driven moving averages

Andreas Basse-O'Connor[1], *Claudio Heinrich*[1] *and Mark Podolskij*[1]

[1] *Department of Mathematics, Aarhus University, Denmark*

Abstract: We present several limit theorems for a class of variation functionals obtained by applying a continuous function f on the kth order differences of stationary increments Lévy driven moving average process. The limiting behavior of such functionals depends not only on the function f at hand but also on the Blumenthal-Getoor index β of the driving Lévy process and on the behavior of the kernel at 0, which is specified by the power α. For the first order asymptotic theory, we show that at least three different cases occur, depending on the interplay of α, β and k as well as on certain properties of the function f. In connection with one of the three cases we prove a second order limit theorem when the function f is bounded, with two different limits; a central limit theorem and, when the Appell rank of f is greater 1, convergence in distribution to a $(k-\alpha)\beta$-stable random variable.

II.1 Introduction and main results

The last years have seen an increasing interest in the limit theory for various classes of stochastic processes. Limit theorems in the high frequency setting are an important tool for analysing the small scale behaviour of stochastic processes and have manifold applications in statistical inference, such as parameter estimation or testing for jumps. For Itô semimartingales the existing limit theory includes power and multipower variation [5] as well as related variation functionals [21, 20]. We refer to [2, 3, 4] for the limit theory for multipower variation of fractional Brownian motion and a class of related processes, and to [14, 15] for power variation of the Rosenblatt process. In the recent publication [7], the authors consider power variations of stationary increments Lévy driven moving averages and derive the first order limit theory as well as a partial second order limit theory. This article builds on their results and extends the limit theory to include more general variation functionals obtained by applying a continuous function to the kth order increments of the process.

We consider an infinitely divisible process with stationary increments $(X_t)_{t\geq 0}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, given as

$$X_t = \int_{-\infty}^{t} \big\{ g(t-s) - g_0(-s) \big\}\, dL_s, \tag{II.1.1}$$

where $L = (L_t)_{t\in\mathbb{R}}$ is a symmetric Lévy process on $\mathbb{R}$ with $L_0 = 0$. That is, for all $u \in \mathbb{R}$, $(L_{t+u} - L_u)_{t\geq 0}$ is a Lévy process indexed by $\mathbb{R}_+$ the distribution of which is invariant under multiplication with -1. Furthermore, g and g_0 are continuous functions from $\mathbb{R}$ into $\mathbb{R}$ vanishing on $(-\infty, 0)$. The class of stationary increments Lévy driven moving averages contains in particular the (symmetric) linear fractional stable motions, which is the model (II.1.1) with $g(s) = g_0(s) = s_+^{\alpha}$ driven by a symmetric stable Lévy process. These processes have been considered by many authors. Recent research addresses various topics such as, among others, semimartingale property [8], fine scale behavior [9, 17], simulation techniques [16] and statistical inference [1].

In this paper we consider for a continuous function $f : \mathbb{R} \to \mathbb{R}$ the variation functional

$$V(f;k)_t^n := a_n \sum_{i=k}^{[nt]} f(b_n \Delta_{i,k}^n X), \tag{II.1.2}$$

where $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ are suitable (nonrandom) normalising sequences, and $[nt]$ denotes the integer part of nt. The kth order increments $\Delta_{i,k}^n X$ of X, $k \in \mathbb{N}$ are defined as

$$\Delta_{i,k}^n X := \sum_{j=0}^{k} (-1)^j \binom{k}{j} X_{(i-j)/n}, \qquad i \geq k.$$

For instance, we have that $\Delta_{i,1}^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ and $\Delta_{i,2}^n X = X_{\frac{i}{n}} - 2X_{\frac{i-1}{n}} + X_{\frac{i-2}{n}}$. We recall that the *Blumenthal–Getoor index* of L is defined as

$$\beta := \inf \Big\{ r \geq 0 : \int_{-1}^{1} |x|^r \,\nu(dx) < \infty \Big\} \in [0,2],$$

where ν denotes the Lévy measure of L. It is well-known that $\sum_{s\in[0,1]} |\Delta L_s|^p$ is finite when $p > \beta$, while it is infinite for $p < \beta$. Here $\Delta L_s = L_s - L_{s-}$ where $L_{s-} = \lim_{u\uparrow s,\, u<s} L_u$. If L is stable with index of stability $\beta \in (0,2)$, the index of stability and the Blumenthal-Getoor index coincide, and both will be denoted β. The asymptotic theory is investigated under the following conditions on g, g_0 and ν that were introduced in [7].

Assumption (A): *The function $g\colon \mathbb{R} \to \mathbb{R}$ satisfies*

$$g(t) \sim c_0 t^\alpha \qquad \text{as } t \downarrow 0 \quad \text{for some } \alpha > 0 \text{ and } c_0 \neq 0,$$

where $g(t) \sim f(t)$ as $t \downarrow 0$ means that $\lim_{t\downarrow 0} g(t)/f(t) = 1$. For some $\theta \in (0,2]$, $\limsup_{t\to\infty} \nu(x\colon |x| \geq t)t^\theta < \infty$ and $g - g_0$ is a bounded function in $L^\theta(\mathbb{R}_+)$. Furthermore, g is k-times continuously differentiable on $(0,\infty)$ and there exists a $\delta > 0$ such that $|g^{(k)}(t)| \leq Ct^{\alpha-k}$ for all $t \in (0,\delta)$, and such that both $|g'|$ and $|g^{(k)}|$ are in $L^\theta((\delta,\infty))$ and are decreasing on (δ,∞).

This assumption ensures in particular that the integral X_t is well-defined in the sense of [24], see [7, Section 2.4]. When L is a β-stable Lévy process, we may and do always choose $\theta = \beta$. For Theorem II.1.1 (i) below, we need to strengthen Assumption (A) slightly if $\theta = 1$ and assume the following

Assumption (A-log): *In addition to (A) suppose that*

$$\int_\delta^\infty |g^{(k)}(s)|^\theta \log(1/|g^{(k)}(s)|)\, ds < \infty.$$

In order to formulate our main results, we require some more notation. For $p > 0$ we denote by $\mathcal{C}^p$ the space of $r := [p]$-times continuous differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $f^{(r)}$ is locally $p - r$ Hölder continuous if $p \notin \mathbb{N}$. Let $h_k\colon \mathbb{R} \to \mathbb{R}$ be given by

$$h_k(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (x-j)_+^\alpha, \qquad x \in \mathbb{R},$$

where $y_+ = \max\{y, 0\}$ for all $y \in \mathbb{R}$. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{R}}$ be the filtration generated by the Lévy process and $(T_m)_{m\geq 1}$ be a sequence of $\mathbb{F}$-stopping times that exhausts the jumps of $(L_t)_{t\geq 0}$. That is, $\{T_m(\omega) : m \geq 1\} \cap [0,\infty) = \{t \geq 0 : \Delta L_t(\omega) \neq 0\}$ and $T_m(\omega) \neq T_n(\omega)$ for all $m \neq n$ with $T_m(\omega) < \infty$. Let $(U_m)_{m\geq 1}$ be independent and uniform $[0,1]$-distributed random variables, defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of the original probability space, which are independent of $\mathcal{F}$. We recall that a sequence $(Z^n)_{n\in\mathbb{N}}$ of random variables defined on $(\Omega, \mathcal{F})$ with values in a Polish space $(E, \mathcal{E})$ converges $\mathcal{F}$-stably in law to Z, which is defined on the extended space $(\Omega', \mathcal{F}')$ if for all bounded continuous $g : E \to \mathbb{R}$ and for all bounded $\mathcal{F}$-measurable random variables Y it holds that $\mathbb{E}[g(Z^n)Y] \to \mathbb{E}'[g(Z)Y]$, where $\mathbb{E}'$ denotes the expectation on the extended space. We denote $\mathcal{F}$-stable convergence in law by $Z^n \xrightarrow{\mathcal{L}-s} Z$, and refer to [25, 18] for more details. By $\xrightarrow{\text{u.c.p.}}$ we denote uniform convergence on compact sets in probability of stochastic processes. That is, $(Z_t^n)_{t\geq 0} \xrightarrow{\text{u.c.p.}} (Z_t)_{t\geq 0}$ as $n \to \infty$ means that $\mathbb{P}(\sup_{t\in[0,N]} |Z_t^n - Z_t| > \varepsilon) \to 0$ for all $N \in \mathbb{N}$ and all $\varepsilon > 0$. A definition of the Skorokhod M_1-topology, which was introduced in [27], will be given in Section II.2. For a more detailed exposition we refer to [31].

Theorem II.1.1. *Suppose (A) is satisfied and assume that the Blumenthal–Getoor index satisfies $\beta < 2$. We have the following three cases:*

(i) Let $k > \alpha$ and suppose that (A-log) holds if $\theta = 1$. Assume $f(0) = 0$ and that $f \in \mathcal{C}^p$ for some $p > \beta \vee \frac{1}{k-\alpha}$. With the normalising sequences $a_n = 1$ and $b_n = n^\alpha$ we obtain the $\mathcal{F}$-stable convergence of finite dimensional distributions

$$V(f;k)_t^n \xrightarrow{\mathcal{L}-s} V(f;k)_t := \sum_{m:T_m\in[0,t]} \sum_{l=0}^{\infty} f\big(c_0 \Delta L_{T_m} h_k(l+U_m)\big),$$

for all $t > 0$. Moreover the sequence of càdlàg processes $\big(V(f;k)_t^n\big)_{t\geq 0}$ converges stably in law to $\big(V(f;k)_t\big)_{t\geq 0}$ with respect to the Skorokhod M_1-topology if f satisfies additionally the following condition:

(FC) Each of the two functions $x \mapsto f(x)\mathbb{1}_{\{x\geq 0\}}$ and $x \mapsto f(x)\mathbb{1}_{\{x<0\}}$ is either nonnegative or nonpositive.

(ii) Suppose that L is a symmetric β-stable Lévy process with scale parameter $\rho_L > 0$. Assume that $H = \alpha + 1/\beta < k$ and $\mathbb{E}[|f(L_1)|] < \infty$. Then, setting $a_n = 1/n$ and $b_n = n^H$, we obtain

$$V(f;k)_t^n \xrightarrow{u.c.p.} t\mathbb{E}[f(S)],$$

where S is a symmetric β-stable random variable with scale parameter $\rho_L \|h_k\|_{L^\beta(\mathbb{R})}$.

(iii) Suppose that $(1 \vee \beta)(k - \alpha) < 1$ and assume that $f(x) \leq C(1 \vee |x|^q)$ for some q with $q(k-\alpha) < 1$, and some finite constant C. With the normalising sequences $a_n = 1/n$ and $b_n = n^k$ it holds that

$$V(f;k)_t^n \xrightarrow{u.c.p.} \int_0^t f(F_u)\,du$$

where $(F_u)_{u\in\mathbb{R}}$ is a version with measurable sample paths of the process defined by

$$F_u = \int_{-\infty}^u g^{(k)}(u-s)\,dL_s \quad \text{a.s. for all } u \in \mathbb{R} \tag{II.1.3}$$

which necessarily satisfies $\int_0^t |f(F_u)|du < \infty$, almost surely.

The limiting random variable in (i) is indeed well-defined, as we show in Lemma II.2.2 below. The three cases of the theorem are closely related to the three possible limits for the realised power variation derived in [7, Theorem 1.1]. We remark that [7, Theorem 1.1] shows only the convergence of the realised power variation at a fixed time $t > 0$, the functional convergence was shown in [6]. Unlike for the power variation, the conditions of Theorem II.1.1 (i) are not in conflict with the conditions of (ii) or (iii). As a consequence, the functional $V(f;k)_t^n$ can converge to different limits for different choices of the normalising sequences (a_n) and (b_n). This phenomenon should not be surprising, however, since it also occurs for other classes of stochastic processes. As an example, consider a β-stable Lévy process L and the function $f(x) = \sin^2(x)$. Then, for the functional $V(f;1)_t^n$ with the normalising sequences $a_n = b_n = 1$ we obtain the almost sure convergence

$$\sum_{i=1}^{[tn]} \sin^2(\Delta_{i,1}^n L) \xrightarrow{\text{a.s.}} \sum_{m:T_m\in[0,t]} \sin^2(\Delta L_{T_m}).$$

The right hand side is indeed finite since $\sin^2$ is bounded and satisfies $\sin^2(x) \sim x^2$ as $x \to 0$. However, for the choice of normalising sequences $a_n = n^{-1}$ and $b_n = n^{1/\beta}$ we obtain by self-similarity of L

$$\frac{1}{n}\sum_{i=1}^{[tn]} \sin^2(n^{1/\beta}\Delta_{i,1}^n L) \stackrel{\text{d}}{=} \frac{1}{n}\sum_{i=1}^{[tn]} \sin^2(\Delta_{i,1}^1 L) \xrightarrow{\text{a.s.}} \mathbb{E}[\sin^2(L_1)],$$

showing that the functional $V(f;k)_t^n$ may have different limits for different normalising sequences, when applied to a Lévy process.

For Theorem II.1.1 (ii) we give a second order limit theorem when the function f is bounded. To this end we introduce the notion of the Appell rank of f. Let

$$\Phi_\rho(x) = \mathbb{E}[f(x+\rho S)] - \mathbb{E}[f(\rho S)],$$

where S is a symmetric β-stable random variable with scale parameter 1, and $\rho > 0$. From boundedness of f it follows that Φ_ρ is infinitely differentiable. The Appell rank of f at $\rho > 0$ is then defined as

$$k_\rho^* := \min\{r \in \mathbb{N} \,:\, \Phi_\rho^{(r)}(0) \neq 0\}.$$

The Appell rank has been introduced in [19] and is known to have major impact on the second order asymptotic behaviour of $V(f;k)_t^n$, the Appell rank of f plays an important role, similar as in the limit theory for discrete time moving averages driven by stable non-Gaussian noise. In Theorem II.1.2 (i) we consider only functions f with $k_\rho^* > 1$, which is for example satisfied when the function is even. Moreover, we restrict ourselves to fixed $t > 0$, without loss of generality $t = 1$, and set $V(f;k)^n := V(f;k)_t^n$. Let us mention that in [23] and [28], where the authors derive similar limit theorems in the low frequency setting, they show functional convergence towards a limiting process. By multiplying the Lévy process with a constant we may and do assume without loss of generality $c_0 = 1$ where the constant c_0 was introduced in Assumption (A). Moreover, we strengthen our basic assumption as follows.

Assumption (A2): *Suppose that in addition to Assumption (A) we have $|g^{(k)}(t)| \leq Ct^{\alpha-k}$ for all $t > 0$. For the function $\zeta : [0,\infty) \to \mathbb{R}$ defined as $\zeta(t) = g(t)t^{-\alpha}$ the limit $\lim_{t\downarrow 0} \zeta^{(j)}$ exists in $\mathbb{R}$ for all $j = 0, ..., k$.*

For Theorem II.1.1 (ii) we obtain the following second order limit theorem.

Theorem II.1.2. *Suppose assumption (A2) is satisfied and that f is bounded. Let L be a symmetric β-stable Lévy process with scale parameter ρ_L and set $H = \alpha + \frac{1}{\beta}$.*

(i) Assume that $\alpha \in (k - 2/\beta, k - 1/\beta)$, and suppose additionally that $k_\rho^* > 1$ for all $\rho > 0$. Then it holds that

$$n^{1-\frac{1}{(k-\alpha)\beta}}\left(n^{-1}\sum_{i=k}^{n}\left\{f\left(n^H\Delta_{i,k}^n X\right) - \mathbb{E}\left[f\left(n^H\Delta_{i,k}^n X\right)\right]\right\}\right) \xrightarrow{\mathcal{L}} S, \quad \text{(II.1.4)}$$

where S is a $(k-\alpha)\beta$-stable random variable with location parameter 0, scale parameter ρ_S and skewness parameter η_S, which are specified in (II.3.62).

(ii) Assume that $\alpha \in (0, k-2/\beta)$. It holds that

$$\sqrt{n}\left(n^{-1}\sum_{i=k}^{n}\left\{f(n^H\Delta_{i,k}^n X) - \mathbb{E}[f(n^H\Delta_{i,k}^n X)]\right\}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\eta^2). \quad \text{(II.1.5)}$$

where the variance is given as $\eta^2 := \lim_{m\to\infty}\eta_m^2$ with η_m defined in (II.3.72).

We remark that the condition $k_\rho^* > 1$ is stronger than the conditions for comparable results for discrete time moving averages, e.g. [28], where it is typically sufficient to control the Appell rank at the scale parameter of the stable random variable $X_t - X_{t-1}$. In Theorem II.4.6 we show that the condition $k_\rho^* > 1$ for all $\rho > 0$ is satisfied if both the positive and negative part of f have Appell rank greater 1 at $\rho = 1$.

Throughout all our proofs we denote by C a generic positive constant that does not depend on n or ω, but may change from line to line. proof we denote all positive constants that do not depend on n or ω by C, even though they may change from line to line. For a random variable Y and $q > 0$ we denote $\|Y\|_q = \mathbb{E}[|Y|^q]^{1/q}$. We abbreviate 'symmetric β-stable' by SβS and denote $Y \sim$ SβS(ρ) if Y is symmetric β-stable distributed with scale parameter ρ, i.e. if its characteristic function is given as

$$\mathbb{E}[\exp(i\eta Y)] = e^{-|\rho\eta|^\beta}, \qquad \eta \in \mathbb{R}.$$

We use frequently the notation

$$g_{i,n}(s) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}g((i-j)/n - s),$$

which leads to the expression

$$\Delta_{i,k}^n X = \int_{-\infty}^{i/n} g_{i,n}(s)\, dL_s$$

for the the kth order increments of X.

II.2 Proof of Theorem II.1.1

In this section we present the proof of Theorem (II.1.1). We begin by briefly recalling the definition and some properties of the Skorokhod M_1-topology, as it is not widely used. It was originally introduced by Skorokhod [27] by defining a metric on the completed graphs of càdlàg functions, where the completed graph of f is defined as

$$\Gamma_f = \{(x,t) \in \mathbb{R}\times\mathbb{R}_+ \,:\, x = \alpha f(t-) + (1-\alpha)f(t),\ \text{for some } \alpha \in [0,1]\}.$$

The M_1-topology is weaker than the more commonly used J_1-topology but still strong enough to make many important functionals, such as sup and inf, continuous. It can be shown that the stable convergence in Theorem II.1.1 (i) does not hold with respect to the J_1-topology. Since M_1 is metrisable, it is entirely defined by characterising convergence of sequences, which we do in the following. A sequence f_n of functions in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ converges to $f \in \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ with respect to the Skorokhod M_1-topology if and only if $f_n(t) \to f(t)$ for all t in a dense subset of $[0, \infty)$, and for all $t_\infty \in [0, \infty)$ it holds that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \sup_{0 \le t \le t_\infty} w(f_n, t, \delta) = 0.$$

Here, the oscillation function w is defined as

$$w(f, t, \delta) = \sup_{0 \vee (t-\delta) \le t_1 < t_2 < t_3 \le (t+\delta) \wedge t_\infty} \{ |f(t_2) - [f(t_1), f(t_3)]| \},$$

where for $b < a$ the interval $[a, b]$ is defined to be $[b, a]$, and $|a - [b, c]| := \inf_{d \in [b,c]} |a - d|$.

For the functions $g_{i,n}$ we obtain the the following important estimates.

Lemma II.2.1. *Suppose that assumption (A) is satisfied. It holds that*

$$\begin{aligned}
&|g_{i,n}(s)| \le C(i/n - s)^\alpha && \text{for } s \in [(i-k-1)/n, i/n], \\
&|g_{i,n}(s)| \le Cn^{-k}((i-k)/n - s)^{\alpha-k} && \text{for } s \in (i/n - \delta, (i-k-1)/n), \text{ and} \\
&|g_{i,n}(s)| \le Cn^{-k}\big(\mathbb{1}_{[(i-k)/n-\delta, i/n-\delta]}(s) + g^{(k)}((i-k)/n - s)\mathbb{1}_{(-\infty,(i-k)/n-\delta)}(s)\big), \\
& && \text{for } s \in (-\infty, i/n - \delta].
\end{aligned}$$

Proof. The first inequality follows directly from (A). The second inequality follows from Taylor expansion of order k and the condition $|g^{(k)}(t)| \le Ct^{\alpha-k}$ for $t \in (0, \delta)$. The third inequality follows again through Taylor expansion and the fact that the function $g^{(k)}$ is decreasing on (δ, ∞). $\square$

Proof of Theorem II.1.1 (i)

The proof is divided into three parts. First, we assume that L is a compound Poisson process and show the stable convergence for fixed $t > 0$. Thereafter we argue that the convergence holds functional with respect to the M_1-topology, when f satisfies condition (FC). Finally, the results are extended to general Lévy processes by truncation. For this step, an isometry for Lévy integrals that is due to [24] plays a key role.

Since $\mathcal{C}^q \subset \mathcal{C}^p$ for $p < q$ we may and do assume that $p \notin \mathbb{N}$. Note that $f \in \mathcal{C}^p$ implies that for any $N > 0$ there is a constant C_N such that

$$|f^{(j)}(x)| \le C_N |x|^{p-j}, \quad \text{for all } x \in [-N, N], \text{ and } j = 0, \dots r. \tag{II.2.6}$$

By the assumption $p > \frac{1}{k-\alpha}$ this implies the following estimate to be used in the proof below. For all $N > 0$ there is a constant C_N such that

$$|f^{(j)}(x)| \le C_N |x|^{\gamma_j}, \quad \text{for all } x \in [-N, N], \text{ and } j = 0, \dots r, \tag{II.2.7}$$

where $\gamma_j = \frac{p-j}{p(k-\alpha)}$. The following Lemma ensures in particular that the limit in Theorem II.1.1 (i) exists.

Lemma II.2.2. *Let $t > 0$ be fixed. Under the conditions of Theorem II.1.1 (i) there is a finite random variable $K > 0$ such that*

$$\sum_{m:T_m\in[0,t]} \sum_{l=0}^{\infty} \big|f\big(c_0\Delta L_{T_m} h_k(l+U_m)\big)\big| \le K, \quad \textit{and}$$

$$\sum_{m:T_m\in[0,t]} \sum_{l=0}^{n-1} \big|f\big(\Delta L_{T_m} n^\alpha g_{i_m+l,n}(T_m)\big)\big| \le K, \quad \textit{for all } n,$$

where i_m denotes the random index such that $T_m \in \big(\frac{i_m-1}{n}, \frac{i_m}{n}\big]$.

Proof. Throughout the proof, K denotes a positive random variable that may change from line to line. For the first inequality note that $|h_k(l+U_m)| \le C(l-k)^{\alpha-k}$ for all $l > k$ and $|h_k(l+U_m)| \le C$ for $l \in \{0,...,k\}$. This implies in particular

$$|c_0\Delta L_{T_m}(\omega)h_k(l+U_m)| < \begin{cases} C(l-k)^{\alpha-k}\sup_{s\in[0,t]}\{|\Delta L_s|\}, & \text{for } l > k \\ C\sup_{s\in[0,t]}\{|\Delta L_s|\}, & \text{for } l \in \{0,...,k\}. \end{cases}$$

Therefore, we find by (II.2.6) a random variable K such that

$$\big|f\big(c_0\Delta L_{T_m} h_k(l+U_m)\big)\big| \le K\big|c_0\Delta L_{T_m} h_k(l+U_m)\big|^p$$

for all $l \ge 0$ and all m. Consequently, the first sum in the lemma is dominated by

$$K\bigg(\sum_{m:T_m\in[0,t]} |\Delta L_{T_m}|^p + \sum_{m:T_m\in[0,t]} |\Delta L_{T_m}|^p \sum_{l=k+1}^{\infty} (l-k)^{(\alpha-k)p}\bigg) < K,$$

where we used that $(\alpha-k)p < -1$, and that $\sum |\Delta L_{T_m}|^p < \infty$ since $p > \beta$. The second inequality follows by the same arguments since Lemma II.2.1 implies the existence of a constant $C > 0$ such that for all $n \in \mathbb{N}$

$$\begin{aligned} &n^\alpha g_{i_m+l,n}(T_m) \le C && \text{for } l \in \{0,...,k\}, \text{ and} \\ &n^\alpha g_{i_m+l,n}(T_m) \le C(l-k)^{\alpha-k}, && \text{for } l \in \{k+1,...,n-1\}. \end{aligned}$$

□

Compound Poisson process as driving process

In this subsection, we show the convergence of $V(f;k)^n_{t_\infty}$ for some fixed $t_\infty > 0$ under the assumption that L is a compound Poisson process. The extension to functional convergence when condition (FC) is satisfied follows in the next subsection, the extension to general L thereafter.

Let $0 \le T_1 < T_2 < ...$ denote the jump times of $(L_t)_{t\ge 0}$. For $\varepsilon > 0$ we define

$$\begin{aligned} \Omega_\varepsilon = \big\{\omega\in\Omega : &\text{ for all } m \text{ with } T_m(\omega)\in[0,t_\infty] \text{ we have } |T_m(\omega)-T_{m-1}(\omega)| > \varepsilon \\ &\text{and } \Delta L_s(\omega) = 0 \text{ for all } s\in[-\varepsilon,0] \text{ and } |\Delta L_s| \le \varepsilon^{-1} \text{ for all } s\in[0,t_\infty]\big\}. \end{aligned}$$

We note that $\Omega_\varepsilon \uparrow \Omega$, as $\varepsilon \downarrow 0$. Letting

$$M_{i,n,\varepsilon} := \int_{i/n-\varepsilon}^{i/n} g_{i,n}(s)\, dL_s, \quad \text{and} \quad R_{i,n,\varepsilon} := \int_{\infty}^{i/n-\varepsilon} g_{i,n}(s)\, dL_s,$$

we have the decomposition $\Delta_{i,k}^n X = M_{i,n,\varepsilon} + R_{i,n,\varepsilon}$. It turns out that $M_{i,n,\varepsilon}$ is the asymptotically dominating term, whereas $R_{i,n,\varepsilon}$ is negligible as $n \to \infty$. We show that, on Ω_ε,

$$\sum_{i=k}^{[nt_\infty]} f(n^\alpha M_{i,n,\varepsilon}) \xrightarrow{\mathcal{L}-s} Z_{t_\infty}, \quad \text{where} \quad Z_{t_\infty} := \sum_{m:T_m\in[0,t_\infty]} \sum_{l=0}^{\infty} f(c_0 \Delta L_{T_m} h_k(l+U_m)). \tag{II.2.8}$$

Here, $(U_m)_{m\geq 1}$ are independent identically $\mathcal{U}([0,1])$-distributed random variables, defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of the original probability space, that are independent of $\mathcal{F}$. For this step, the following expression for the left hand side is instrumental. On Ω_ε it holds that

$$\sum_{i=k}^{[nt]} f(n^\alpha M_{i,n,\varepsilon}) = V_t^{n,\varepsilon}, \tag{II.2.9}$$

where

$$V_t^{n,\varepsilon} := \sum_{m:T_m\in(0,[nt]/n]} \sum_{l=0}^{v_t^m} f(n^\alpha \Delta L_{T_m} g_{i_m+l,n}(T_m)). \tag{II.2.10}$$

Here, i_m denotes the random index such that $T_m \in ((i_m-1)/n, i_m/n]$, and v_t^m is defined as

$$v_t^m = v_t^m(\varepsilon, n) := \begin{cases} [\varepsilon n] \wedge ([nt]-i_m) & \text{if } T_m - ([\varepsilon n]+i_m)/n > -\varepsilon, \\ [\varepsilon n]-1 \wedge ([nt]-i_m) & \text{if } T_m - ([\varepsilon n]+i_m)/n \leq -\varepsilon. \end{cases} \tag{II.2.11}$$

Additionally, we set $v_t^m = \infty$ if $T_m > [nt]/n$. The following Lemma proves (II.2.8) in a slightly more general fashion, since the proof of functional convergence in the next subsection requires convergence of finite dimensional distributions.

Lemma II.2.3. *For $r \geq 1$ and $0 \leq t_1 < \cdots < t_r \leq t_\infty$ we obtain on Ω_ε the $\mathcal{F}$-stable convergence*

$$(V_{t_1}^{n,\varepsilon}, \ldots, V_{t_r}^{n,\varepsilon}) \xrightarrow{\mathcal{L}-s} (Z_{t_1}, \ldots, Z_{t_r}), \qquad \text{as } n \to \infty.$$

Proof. By arguing as in [7, Section 5.1], we deduce for any $d \geq 1$ the $\mathcal{F}$-stable convergence

$$\{n^\alpha g_{i_m+l,n}(T_m)\}_{l,m\leq d} \xrightarrow{\mathcal{L}-s} \{c_0 h_k(l+U_m)\}_{l,m\leq d}$$

as $n \to \infty$. Defining

$$V_t^{n,d} := \sum_{m\leq d:T_m\in(0,[nt]/n]} \sum_{l=0}^{d} f(n^\alpha \Delta L_{T_m} g_{i_m+l,n}(T_m)) \quad \text{and}$$

$$Z_t^d := \sum_{m\leq d:T_m\in(0,t]} \sum_{l=0}^{d} f(c_0 \Delta L_{T_m} h_k(l+U_m)),$$

we obtain by the continuous mapping theorem for stable convergence

$$(V_{t_1}^{n,d}, \ldots, V_{t_r}^{n,d}) \xrightarrow{\mathcal{L}-s} (Z_{t_1}^d, \ldots, Z_{t_r}^d), \qquad \text{as } n \to \infty, \tag{II.2.12}$$

for all $d \geq 1$. Therefore, by a standard approximation argument (cf. [11, Thm 3.2]), it is sufficient to show that

$$\limsup_{n\to\infty} \left\{ \max_{t\in\{t_1,\ldots,t_r\}} |V_t^{n,\varepsilon} - V_t^{n,d}| \right\} \xrightarrow{\text{a.s.}} 0, \qquad \text{as } d \to \infty, \text{ and} \tag{II.2.13}$$

$$\sup_{t\in[0,t_\infty]} |Z_t^d - Z_t| \xrightarrow{\text{a.s.}} 0, \qquad \text{as } d \to \infty. \tag{II.2.14}$$

For sufficiently large n we have

$$\begin{aligned}
|V_t^{n,d} - V_t^{n,\varepsilon}| \leq & \sum_{m\leq d: T_m\in(0,[nt]/n]} \sum_{l=d\wedge v_t^m}^{d\vee v_t^m} |f(\Delta L_{T_m} n^\alpha g_{i_m+l,n}(T_m))| \\
& + \sum_{m>d: T_m\in(0,[nt]/n]} \sum_{l=0}^{v_t^m} |f(\Delta L_{T_m} n^\alpha g_{i_m+l,n}(T_m))| \\
\leq & \sum_{m: T_m\in(0,t_\infty]} \sum_{l=d\wedge v_t^m}^{n-1} |f(\Delta L_{T_m} n^\alpha g_{i_m+l,n}(T_m))| \\
& + \sum_{m>d: T_m\in(0,[nt]/n]} \sum_{l=0}^{n-1} |f(\Delta L_{T_m} n^\alpha g_{i_m+l,n}(T_m))|,
\end{aligned}$$

for all $t \in [0, t_\infty]$. Therefore, (II.2.13) follows from Lemma II.2.2 by an application of the dominated convergence theorem since the random index $v_t^m = v_t^m(n,\omega)$ satisfies $\liminf_{n\to\infty} v_t^m(n,\omega) = \infty$, almost surely. Lemma II.2.2 also implies (II.2.14), since

$$\begin{aligned}
\sup_{t\in[0,t_\infty]} |Z_t^d - Z_t| \leq & \sum_{m\leq d: T_m\in(0,t_\infty]} \sum_{l=d+1}^{\infty} |f(c_0 \Delta L_{T_m} h_k(l+U_m))| \\
& + \sum_{m>d: T_m\in(0,t_\infty]} \sum_{l=0}^{\infty} |f(c_0 \Delta L_{T_m} h_k(l+U_m))|.
\end{aligned}$$

The Lemma now follows from (II.2.12), (II.2.13) and (II.2.14). □

Recalling the decomposition (II.2.8) and applying the triangle inequality, the proof can be completed by showing that

$$J_n := \sum_{i=k}^{[nt_\infty]} |f(n^\alpha \Delta_{i,k}^n X) - f(n^\alpha M_{i,n,\varepsilon})| \xrightarrow{\text{a.s.}} 0, \qquad \text{as } n \to \infty. \tag{II.2.15}$$

We first argue that the random variables $\{n^\alpha M_{i,n,\varepsilon}, n^\alpha \Delta_{i,k}^n X\}_{n\in\mathbb{N}, i\in\{k,\ldots,[nt_\infty]\}}$ are on Ω_ε uniformly bounded by a constant, which will allow us to apply the estimate (II.2.6). The random variables $M_{i,n,\varepsilon}$ satisfy by construction either $|n^\alpha M_{i,n,\varepsilon}| = 0$ or $|n^\alpha M_{i,n,\varepsilon}| = |n^\alpha g_{i,n}(T_m)\Delta L_{T_m}|$ for some m, where we recall that on Ω_ε it holds that $T_m - T_{m-1} > \varepsilon$. Consequently, they are uniformly bounded by Lemma II.2.1, where we used that $k > \alpha$ and that the jumps of L are bounded on Ω_ε. The uniform boundedness of $n^\alpha \Delta_{i,k}^n X = n^\alpha(M_{i,n,\varepsilon} + R_{i,n,\varepsilon})$ follows by [7, (4.8),(4.12)] which implies that for any $\eta > 0$

$$\sup_{n\in\mathbb{N},\ i\in\{k,\dots,[nt_\infty]\}} \left\{ n^{k-\eta} |R_{i,n,\varepsilon}| \right\} < \infty, \quad \text{almost surely.} \tag{II.2.16}$$

In order to show (II.2.15) we apply Taylor expansion for f at $n^\alpha M_{i,n,\varepsilon}$, and bound the terms in the Taylor expansion using (II.2.6) and the following result.

Lemma II.2.4. *Let $\psi : \mathbb{R} \to \mathbb{R}$ be continuous and such that $|\psi(x)| \le C|x|^\gamma$ for all $x \in [-1, 1]$ for some $\gamma \in (0, 1/(k - \alpha))$. It holds on Ω_ε that*

$$\limsup_{n\to\infty} \left\{ n^{(k-\alpha)\gamma-1} \sum_{i=k}^{[nt_\infty]} |\psi(n^\alpha M_{i,n,\varepsilon})| \right\} \le C, \quad a.s.$$

Proof. We have on Ω_ε

$$\sum_{i=k}^{[nt_\infty]} |\psi(n^\alpha M_{i,n,\varepsilon})| = W_{t_\infty}^{n,\varepsilon},$$

where

$$W_{t_\infty}^{n,\varepsilon} := \sum_{m:T_m\in(0,[nt_\infty]/n]} \sum_{l=0}^{v_{t_\infty}^m} |\psi(n^\alpha \Delta L_{T_m} g_{i_m+l,n}(T_m))|,$$

and $v_{t_\infty}^m$ is the random index defined in (II.2.11). By Lemma II.2.1 the random variables $n^\alpha g_{i_m+l,n}(T_m)$ are bounded for $l = 0, ..., k$. For $l \in \{k+1, ..., n-1\}$, Lemma II.2.1 implies that $n^\alpha g_{i_m+l,n}(T_m) \le C(l-k)^{\alpha-k}$. Since the random index $v_{t_\infty}^m$ satisfies $v_{t_\infty}^m < n$ for all m, we obtain on Ω_ε

$$\sum_{i=k}^{[nt_\infty]} |\psi(n^\alpha M_{i,n,\varepsilon})| \le C \sum_{m:T_m\in(0,t_\infty]} \left(\sum_{l=0}^{k} |n^\alpha g_{i_m+l,n}(T_m)|^\gamma + \sum_{l=k+1}^{n} |(l-k)^{\alpha-k}|^\gamma \right).$$

It follows by comparison with the integral $\int_{k+1}^n (s-k)^{(\alpha-k)\gamma}\, ds$ that the right hand side multiplied with $n^{(k-\alpha)\gamma-1}$ is convergent, where we used that $(\alpha - k)\gamma \in (-1, 0)$ and that the number of jumps of $L(\omega)$ in $[0, t_\infty]$ is uniformly bounded for $\omega \in \Omega_\varepsilon$. $\square$

Considering the sum J_n in (II.2.15), Taylor expansion up to order $r = [p]$ shows that

$$\begin{aligned} J_n &\le \sum_{i=k}^{[nt_\infty]} \left| n^\alpha R_{i,n,\varepsilon} f'(n^\alpha M_{i,n,\varepsilon}) \right| + \cdots + \frac{1}{r!} \sum_{i=k}^{[nt_\infty]} \left| (n^\alpha R_{i,n,\varepsilon})^r f^{(r)}(n^\alpha M_{i,n,\varepsilon}) \right| + TR_r \\ &:= T_1 + \cdots + T_r + TR_r, \end{aligned} \tag{II.2.17}$$

where TR_r denotes the Taylor rest term. Recalling the estimate (II.2.7), we can now for $j = 0, \dots, [p]$ estimate the jth Taylor monomial T_j by applying Lemma II.2.4 on

$\psi = f^{(j)}$, where we remark that $\gamma_j = \frac{p-j}{p(k-\alpha)} \in (0, 1/(k-\alpha))$. Using (II.2.16) and recalling that $p > k - \alpha$, we obtain that for sufficiently small $\eta > 0$

$$\begin{aligned}\frac{1}{j!}\sum_{i=k}^{[nt_\infty]} \big|(n^\alpha R_{i,n,\varepsilon})^j f^{(j)}(n^\alpha M_{i,n,\varepsilon})\big| &\le Cn^{-j/p-\eta}\sum_{i=k}^{[nt_\infty]} \big|f^{(j)}(n^\alpha M_{i,n,\varepsilon})\big| \\ &\le Cn^{-\eta}, \end{aligned} \tag{II.2.18}$$

where the second inequality follows from Lemma II.2.4 since $(k-\alpha)\gamma_j - 1 = -j/p$. For the Taylor rest term TR_r we have by the mean value theorem the expression

$$TR_r = \frac{1}{r!}\sum_{i=k}^{[nt_\infty]} \big|(n^\alpha R_{i,n,\varepsilon})^r\big(f^{(r)}(\xi_{i,n}) - f^{(r)}(n^\alpha M_{i,n,\varepsilon})\big)\big|,$$

with $\xi_{i,n} \in (n^\alpha|M_{i,n,\varepsilon}|, n^\alpha|X_{i,n,\varepsilon}|)$ where we set $(a,b) := (b,a)$ for $a > b$. Since $n^\alpha|M_{i,n,\varepsilon}|$ and $n^\alpha|X_{i,n,\varepsilon}|$ are bounded and $f^{(r)}$ is locally $p-r$-Hölder continuous, it follows that

$$TR_r \le Cn \sup_{n\in\mathbb{N},\ i\in\{k,\ldots,[nt_\infty]\}} |n^\alpha R_{i,n,\varepsilon}|^p.$$

From (II.2.16) it follows that $TR_r \to 0$ as $n \to \infty$, where we recall that $(\alpha - k)p < -1$.

Together with (II.2.17) and (II.2.18) this implies $J_n \xrightarrow{\text{a.s.}} 0$, and it follows that

$$\sup_{t\in[0,t_\infty]}\left\{\left|V(f;k)^n_t - \sum_{i=k}^{[tn]} f(n^\alpha M_{i,n,\varepsilon})\right|\right\} \xrightarrow{\text{a.s.}} 0$$

on Ω_ε. Now, the theorem follows from Lemma II.2.3 by letting $\varepsilon \to 0$.

Functional convergence

In this subsection we show that if f satisfies (FC) and under the assumption that L is a compound Poisson process, the convergence in Theorem II.1.1 holds functional with respect to the Skorokhod $M1$-topology. To this end, we denote by $\xrightarrow{\mathcal{L}_{M_1}-s}$ the $\mathcal{F}$-stable convergence of càdlàg processes, regarded as $\mathbb{D}([0,t_\infty];\mathbb{R})$-valued random variables, where t_∞ is some fixed positive time horizon and $\mathbb{D}([0,t_\infty];\mathbb{R})$ is equipped with the Skorokhod M_1-topology. We first replace (FC) by the following stronger auxiliary assumption.

(FC') It holds that f is either nonnegative or nonpositive.

This assumption puts us into the comfortable situation that our limiting process is monotonic. Recall the definition of the processes $V^{n,\varepsilon}$ and Z introduced in (II.2.8) and (II.2.10), respectively. In Lemma II.2.3 the stable convergence of the finite dimensional distributions of $V^{n,\varepsilon}$ to Z was shown. By Prokhorov's theorem the functional convergence $V^{n,\varepsilon} \xrightarrow{\mathcal{L}_{M_1}-s} Z$ on Ω_ε follows thus from the following Lemma.

Lemma II.2.5. *The sequence of $\mathbb{D}([0,t_\infty])$-valued random variables $(V^{n,\varepsilon}\mathbb{1}_{\{\Omega_\varepsilon\}})_{n\ge1}$ is tight if $\mathbb{D}([0,t_\infty])$ is equipped with the Skorokhod M_1-topology.*

Proof. It is sufficient to show that the conditions of [31, Theorem 12.12.3] are satisfied. Condition (i) is satisfied, since the family of real valued random variables $(V^{n,\varepsilon}_{t_\infty})_{n\geq 1}$ is tight by Lemma II.2.3. Condition (ii) is satisfied, since the oscillating function w_s introduced in [31, chapter 12, (5.1)] satisfies $w_s(V^{n,\varepsilon}, \theta) = 0$ for all $\theta > 0$ and all n, since $V^{n,\varepsilon}$ is monotonic by assumption (FC'). □

Recalling the identity (II.2.9) and the asymptotic equivalence of $\sum_{i=k}^{[tn]} f(n^\alpha M_{i,n,\varepsilon})$ and $V(f;k)^n$ shown in (II.2.15) and thereafter, the functional convergence in Theorem II.1.1 follows.

Now, for general f satisfying condition (FC) we decompose $f = f_+ + f_-$ with $f_+(x) = f(x)\mathbb{1}_{\{x>0\}}$ and $f_-(x) = f(x)\mathbb{1}_{\{x<0\}}$. Both functions f_+ and f_- satisfy (FC'), and the functional convergence of $V(f_+;k)^n$ and $V(f_-;k)^n$ follows, with the corresponding limits denoted by Z^+ and Z^-. Note that Z^+ jumps exactly at those times, where the Lévy process L jumps up, and Z^- at those, where it jumps down. In particular, Z^+ and Z^- do not jump at the same time, which implies that summation is continuous at (Z^+, Z^-) with respect to the M_1-topology (cf. [31, Thm. 12.7.3]). Thus, an application of the continuous mapping theorem yields the convergence of $V(f;k)^n = V(f_+;k)^n + V(f_-;k)^n$ towards $Z = Z^+ + Z^-$. Let us stress that indeed the sole reason why the extra condition (FC) is required for functional convergence is that summation is not continuous on the Skorokhod space, and consequently the convergence of $V(f_+;k)^n$ and $V(f_-;k)^n$ does not generally imply the convergence of $V(f;k)^n$.

Extension to infinite activity Lévy processes

In this section we extend the results of Theorem II.1.1 (i) to moving averages driven by a general Lévy process L, by approximating L by a sequence of compound Poisson processes $(\hat{L}(j))_{j\geq 1}$. To this end we introduce the following notation. Let N be the jump measure of L, that is $N(A) := \#\{t : (t, \Delta L_t) \in A\}$ for measurable $A \subset \mathbb{R} \times (\mathbb{R} \setminus \{0\})$, and define for $j \in \mathbb{N}$

$$X_t(j) := \int_{(-\infty,t]\times[-\frac{1}{j},\frac{1}{j}]} \{(g(t-s) - g_0(-s))x\} N(ds, dx).$$

Denote $\hat{X}_t(j) := X_t - X_t(j)$. The results of the last section show that Theorem II.1.1 holds for $\hat{X}(j)$, since it is a moving average driven by a compound Poisson process. By letting $j \to \infty$ we will show that the theorem remains valid for X by deriving the following approximation result

Lemma II.2.6. *Suppose that f satisfies the conditions of Theorem II.1.1 (i). It holds for all $\varepsilon > 0$ that*

$$\lim_{j\to\infty} \limsup_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,t_\infty]} |V(X, f; k)^n_t - V(\hat{X}(j), f; k)^n_t| > \varepsilon \Big) = 0. \quad \text{(II.2.19)}$$

Proof. In the following we call a family of random variables $\{Y_{n,j}\}_{n,j\in\mathbb{N}}$ *asymptotically tight* if for any $\varepsilon > 0$ there is an $N > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}(|Y_{n,j}| > N) < \varepsilon, \quad \text{for all } j \in \mathbb{N}.$$

We deduce first for $p > \beta \vee \frac{1}{k-\alpha}$ the asymptotic tightness of the families

$$\left\{ \sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n X|^p, \sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n \hat{X}(j)|^p \right\}_{n,j\in\mathbb{N}}, \quad \text{and} \tag{II.2.20}$$

$$\left\{ \max_{i\in k,\dots,[nt_\infty]} (|n^\alpha \Delta_{i,k}^n X|), \max_{i\in k,\dots,[nt_\infty]} (|n^\alpha \Delta_{i,k}^n \hat{X}(j)|) \right\}_{n,j\in\mathbb{N}}.$$

The authors of [7] show the stable convergences in law

$$\sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n \hat{X}(j)|^p \xrightarrow{\mathcal{L}-s} Z_j, \quad \sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n X|^p \xrightarrow{\mathcal{L}-s} Z,$$

where Z_j and Z are defined as in [7, (4.34)]. The asymptotic tightness of the first family follows thus from the tightness of the family $\{Z_j, Z\}_{j\in\mathbb{N}}$, see [7, (4.35)]. The second family follows from the first by the estimate $\max_{i=1,\dots,n}(|a_i|) \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p}$ for $a_1, ..., a_n \in \mathbb{R}$. The asymptotic tightness of the second family allows us for the proof of (II.2.19) to assume that $|\Delta_{i,k}^n \hat{X}(j)|$ and $|\Delta_{i,k}^n X|$ are uniformly bounded by some $N > 0$. Consider first the case $p < 1$. By local Hölder-continuity of f of order p we have that

$$\sup_{t\in[0,t_\infty]} |V(f, X; k)_t^n - V(f, \hat{X}(j); k)_t^n| \le C_N \sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n X(j)|^p,$$

and (II.2.19) follows from [7, Lemma 4.2], where we used that $p > \beta \vee \frac{1}{(k-\alpha)}$. Let now $p > 1$. We can find $\xi_{i,n,j} \in [n^\alpha \Delta_{i,k}^n \hat{X}(j), n^\alpha \Delta_{i,k}^n X]$ such that $|f(n^\alpha \Delta_{i,k}^n \hat{X}(j)) - f(n^\alpha \Delta_{i,k}^n X)| = |n^\alpha \Delta_{i,k}^n X(j) f'(n^\alpha \xi_{i,n,j})|$ and obtain by (II.2.6)

$$\begin{aligned} |f(n^\alpha \Delta_{i,k}^n \hat{X}(j)) - f(n^\alpha \Delta_{i,k}^n X)| &\le C|n^\alpha \Delta_{i,k}^n X(j)||\xi_{i,n,j}|^p \\ &\le C|n^\alpha \Delta_{i,k}^n X(j)||\xi_{i,n,j}|^\gamma \\ &\le C|n^\alpha \Delta_{i,k}^n X(j)|^{\gamma+1} + C|n^\alpha \Delta_{i,k}^n X(j)||n^\alpha \Delta_{i,k}^n X|^\gamma, \end{aligned}$$

with $\gamma = \frac{p-1}{p}\left(\beta \vee \frac{1}{k-\alpha}\right)$ satisfying $\gamma < p-1$ by assumption. Thus, in order to complete the proof of (II.2.19), it is sufficient to show that for all $\varepsilon > 0$ we obtain

$$\lim_{j\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n X(j)|^{\gamma+1} > \varepsilon \right) = 0, \quad \text{and} \tag{II.2.21}$$

$$\lim_{j\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\sum_{i=k}^{[nt_\infty]} |n^\alpha \Delta_{i,k}^n X(j)||n^\alpha \Delta_{i,k}^n X|^\gamma > \varepsilon \right) = 0. \tag{II.2.22}$$

By definition it holds that $\gamma + 1 > \beta \vee \frac{1}{k-\alpha}$, and (II.2.21) follows from [7, Lemma 4.2]. For (II.2.22) we choose Hölder conjugates θ_1 and $\theta_2 = \theta_1/(\theta_1 - 1)$ with $\theta_1 \in \left(\beta \vee \frac{1}{k-\alpha}, p\right)$, where we used that $p > 1$. Hölders inequality and the estimate

$$\mathbb{P}(|XY| > \varepsilon) \le \mathbb{P}(|X| > \varepsilon/N) + P(|Y| > N) \qquad \text{for any } N > 0$$

lead to the decomposition

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{i=k}^{[nt_\infty]}|n^\alpha\Delta_{i,k}^n X(j)||n^\alpha\Delta_{i,k}^n X|^\gamma>\varepsilon\Big)\\
&\qquad\le\mathbb{P}\Big(\sum_{i=k}^{[nt_\infty]}|n^\alpha\Delta_{i,k}^n X(j)|^{\theta_1}>\Big(\frac{\varepsilon}{N}\Big)^{\theta_1}\Big)+\mathbb{P}\Big(\sum_{i=k}^{[nt_\infty]}|n^\alpha\Delta_{i,k}^n\hat{X}(j)|^{\gamma\theta_2}>N^{\theta_2}\Big)\\
&\qquad:=J_{n,j,N}^1+J_{n,j,N}^2.
\end{aligned}$$

Since $\theta_1>\beta\vee\frac{1}{k-\alpha}$, yet another application of [7, Lemma 4.2] yields that

$$\lim_{j\to\infty}\limsup_{n\to\infty}J_{n,j,N}^1=0\quad\text{ for all } N>0.$$

Moreover, $\theta_1<p$ implies $\gamma\theta_2>\beta\vee\frac{1}{k-\alpha}$. Therefore, it follows from the asymptotic tightness of the family (II.2.20) that

$$\limsup_{j\to\infty}\limsup_{n\to\infty}J_{n,j,N}^2\to 0,\quad\text{ as } N\to\infty.$$

This shows (II.2.22) which completes the proof of the Lemma. □

Finally, the proof of Theorem II.1.1 (i) can be completed by letting $j\to\infty$. More precisely, we introduce for $j\in\mathbb{N}$ the stopping times

$$T_{m,j}:=\begin{cases}T_m & \text{if } |\Delta L_{T_m}|>1/j,\\ \infty & \text{else.}\end{cases}$$

The results of the last two subsections show that

$$V(\hat{X}(j),f;k)_t^n\xrightarrow{\mathcal{L}-s}Z_t^j:=\sum_{m:T_{m,j}\in[0,t]}\sum_{l=0}^{\infty}f(c_0\Delta L_{T_{m,j}}h_k(l-U_m)),\quad\text{ for all } t>0,$$

and that the convergence holds functional with respect to the M_1-topology if f satisfies (FC). From Lemma II.2.2 and an application of the dominated convergence theorem it follows that

$$\sup_{t\in[0,t_\infty]}|Z_t-Z_t^j|\xrightarrow{\text{a.s.}}0,\quad\text{ as } j\to\infty.$$

Theorem II.1.1 (i) follows therefore from Lemma II.2.6 and a standard approximation argument (cf. [11, Thm 3.2]). □

Proof of Theorem II.1.1 (ii)

Let us first remark that it is sufficient to show convergence in probability for fixed $t>0$ in order to obtain u.c.p.-convergence by the following standard argument. Making the decomposition $f=f_+-f_-$ with $f_+(x)=f(x)\mathbb{1}_{\{f(x)>0\}}$ and $f_-(x)=-f(x)\mathbb{1}_{\{f(x)<0\}}$, the statistics $V(f_+;k)_t^n$ and $V(f_-;k)_t^n$ are increasing in t and converge to the (non-random) limiting processes $(t\mathbb{E}[f_+(S)])_{t\ge0}$ and $(t\mathbb{E}[f_-(S)])_{t\ge0}$, respectively. Since the limiting processes are continuous in t, u.c.p.-convergence follows from convergence in probability for all $t > 0$, see for example [21, Equation (2.2.16)].

The proof relies on replacing the increments of X by the increments of its tangent process, which is the linear fractional stable motion Y, defined as

$$Y_t = \int_{-\infty}^{t} \big\{(t-s)^{\alpha} - (-s)_+^{\alpha}\big\} dL_s,$$

where $x_+ := \max\{x, 0\}$. It is well known that the process Y is self-similar of index $H = \alpha + 1/\beta$, i.e. $(Y_{at})_{t\geq 0} \overset{\mathrm{d}}{=} (a^H Y_t)$ for any $a > 0$, see [29]. Moreover, the discrete time stationary sequence $(Y_r)_{r\in\mathbb{Z}}$ is mixing and hence ergodic, see for example [13]. Denoting by $V(f;Y)_t^n$ the variation functional (II.1.2) with $a_n = n^{-1}$ and $b_n = n^H$ applied on the process Y, it follows from Birkhoff's ergodic theorem, see [22, Theorem 10.6], that

$$V(f;Y)_t^n = \frac{1}{n}\sum_{i=k}^{[nt]} f(n^H \Delta_{i,k}^n Y) \overset{\mathrm{d}}{=} \frac{1}{n}\sum_{i=k}^{[nt]} f(\Delta_{i,k}^1 Y) \to t\mathbb{E}[f(\Delta_{k,k}^1 Y)], \quad \text{almost surely.}$$

Here we used that the expectation on the right hand side is well-defined by assumption. By (II.3.45), the random variable $\Delta_{k,k}^1 Y$ is SβS distributed with scale parameter $\rho_L \|h_k\|_{L^\beta(\mathbb{R})}$, and the right hand side is the limiting expression in the theorem. It is therefore sufficient to argue that

$$\mathbb{E}\big[|V(X;f)_t^n - V(Y;f)_t^n|\big] \to 0, \quad \text{as } n\to\infty. \tag{II.2.23}$$

For $N > 0$ and $\varepsilon > 0$ we denote by $w_f(\varepsilon, N)$ the modulus of continuity

$$w_f(\varepsilon, N) := \sup\{|f(x) - f(y)| \,:\, x, y \in [-N, N],\ |x-y| < \varepsilon\}.$$

We obtain the estimate

$$\begin{aligned}
&\mathbb{E}\big[|V(X;f)_t^n - V(Y;f)_t^n|\big] \leq \frac{1}{n}\sum_{i=k}^{[nt]} \mathbb{E}\big[|f(n^H\Delta_{i,k}^n X) - f(n^H \Delta_{i,k}^n Y)|\big] \\
&\leq \frac{1}{n}\sum_{i=k}^{[nt]} \Big(w_f(\varepsilon, N)\mathbb{P}\big(n^H|\Delta_{i,k}^n X - \Delta_{i,k}^n Y| < \varepsilon,\ |n^H\Delta_{i,k}^n X| \vee |n^H \Delta_{i,k}^n Y| \leq N\big) \\
&\qquad + C_N \mathbb{P}\big(n^H|\Delta_{i,k}^n X - \Delta_{i,k}^n Y| > \varepsilon,\ |n^H\Delta_{i,k}^n X| \vee |n^H \Delta_{i,k}^n Y| \leq N\big) \\
&\qquad + \mathbb{E}\big[\mathbb{1}_{\{|n^H\Delta_{i,k}^n X| > N\}} |f(n^H\Delta_{i,k}^n X)|\big] + \mathbb{E}\big[\mathbb{1}_{\{|n^H\Delta_{i,k}^n Y| > N\}} |f(n^H\Delta_{i,k}^n Y)|\big]\Big) \\
&=: \frac{1}{n}\sum_{i=k}^{[nt]} (J_{i,\varepsilon,N}^{n,1} + J_{i,\varepsilon,N}^{n,2} + J_{i,N}^{n,3}),
\end{aligned} \tag{II.2.24}$$

where $C_N = 2\sup_{|x|\leq N}|f(x)|$. For the first summand we have that for any $N > 0$

$$\frac{1}{n}\sum_{i=k}^{[nt]} J_{i,\varepsilon,N}^{n,1} \leq w_f(\varepsilon, N) \to 0 \qquad \text{as } \varepsilon \to 0.$$

For the second summand an application of Markov's inequality with some $p < \beta$ shows that for all $\varepsilon, N > 0$

$$\frac{1}{n}\sum_{i=k}^{[nt]} J_{i,\varepsilon,N}^{n,2} \leq \frac{C_N \varepsilon^{-p}}{n}\sum_{i=1}^{[nt]} \mathbb{E}\big[|n^H(\Delta_{i,k}^n X - \Delta_{i,k}^n Y)|^p\big] \to 0 \qquad \text{as } n \to \infty,$$

where the convergence follows from [7, (4.45)]. Hence, by stationarity of $(n^H \Delta_{i,k}^n X)_{i \geq k}$, it is sufficient to argue that

$$\lim_{N\to\infty} \limsup_{n\to\infty} \mathbb{E}\big[\mathbb{1}_{\{|n^H \Delta_{k,k}^n X| > N\}} |f(n^H \Delta_{k,k}^n X)|\big] = 0, \tag{II.2.25}$$

which we do in the following. From [7, (4.45)] it follows that $n^H \Delta_{k,k}^n X \xrightarrow{\mathcal{L}} \Delta_{k,k}^1 Y$, implying that $\rho_n \to \rho$ where ρ_n and ρ denote the scale parameters of the SβS random variables $n^H \Delta_{k,k}^n X$ and $\Delta_{k,k}^1 Y$, respectively. In particular there are constants $c, C \in (0, \infty)$ such that $c < \rho_n < C$ for all n. Recalling that the density ψ of a standard SβS random variable S satisfies $\psi(y) \leq C(1+|y|)^{-1-\beta}$, it follows that

$$\begin{aligned}
\mathbb{E}\big[\mathbb{1}_{\{|n^H \Delta_{k,k}^n X| > N\}} |f(n^H \Delta_{k,k}^n X)|\big] &= \mathbb{E}\big[\mathbb{1}_{\{|\rho_n S| > N\}} |f(\rho_n S)|\big] \\
&\leq C \int_{\mathbb{R}} (1+|y|)^{-1-\beta} \mathbb{1}_{\{|\rho_n y| > N\}} |f(\rho_n y)|\, dy \\
&= C\rho_n^{-1} \int_{\mathbb{R}} (1+|\rho_n^{-1} y|)^{-1-\beta} \mathbb{1}_{\{|y| > N\}} |f(y)|\, dy \\
&\leq C\rho_n^{\beta} \int_{\mathbb{R}} (c+|y|)^{-1-\beta} \mathbb{1}_{\{|y| > N\}} |f(y)|\, dy \\
&\leq C \int_{\mathbb{R}} (c+|y|)^{-1-\beta} \mathbb{1}_{\{|y| > N\}} |f(y)|\, dy
\end{aligned}$$

Now (II.2.25) follows from $\mathbb{E}[|f(S)|] < \infty$, and the decomposition (II.2.24) implies (II.2.23) by letting $N \to \infty$ and $\varepsilon \to 0$. This completes the proof of Theorem II.1.1 (ii). □

Proof of Theorem II.1.1 (iii)

By the argument given at the beginning of the last subsection, u.c.p.-convergence follows if we show convergence in probability of $V(f;k)_t^n$ for arbitrary $t > 0$.

Let us first remark that the growth condition $|f(x)| \leq C(1 \vee |x|^q)$ for some q with $q(k-\alpha) < 1$ is weaker for larger q and can therefore be thought of as

$$|f(x)| \leq C|x|^{\frac{1}{k-\alpha} - \varepsilon} \quad \text{ for } |x| \to \infty,$$

if $k > \alpha$, whereas for $k \leq \alpha$ we require only that f is of polynomial growth. Since by assumption of the theorem we have $k - \alpha < 1$, we may and do always assume that $q > 1$. We recall that a function $\xi : \mathbb{R} \to \mathbb{R}$ is absolutely continuous if there is a function ξ' such that

$$\xi(t) - \xi(s) = \int_s^t \xi'(u)\, du, \quad \text{for all } s < t.$$

This implies that ξ is differentiable almost everywhere and the derivative coincides with ξ' almost everywhere. If ξ' can again be chosen absolutely continuous with derivative $\xi^{(2)}$ we say that ξ is two times absolutely continuous, and similarly we define k-times absolute continuity.

By an application of [12, Theorem 5.1] it has been shown in [7, Lemma 4.3] that under the condition $(k-\alpha)(1 \vee \beta) > 1$ the process X admits a k-times absolutely continuous version and the k-th derivative is a version of the process $(F_u)_{u\in\mathbb{R}}$ defined in (II.1.3). Moreover, [7, Lemma 4.3] shows that for every $q \geq 1, q \neq \theta$ with $q(k-\alpha) < 1$ the process F admits a version with sample paths in $L^q([0,t])$, almost surely, which implies $\int_0^t |f(F_u)|du < \infty$. The intuition behind the convergence in Theorem II.1.1 (iii) is that by the mean value theorem we have $n^k \Delta_{i,k}^n X \approx F_{\frac{i-1}{n}}$ which implies

$$V(f;k)_t^n = \frac{1}{n}\sum_{i=k}^n f(n^k \Delta_{i,k}^n X) \approx \frac{1}{n}\sum_{i=k}^n f(F_{\frac{i-1}{n}}) \to \int_0^t f(F_u)du, \quad \text{as } n \to \infty,$$

by convergence of Riemann sums to the integral. The remainder of this section is dedicated to formalising this statement. This requires some work, mainly due to the fact that the kth derivative process F is not necessarily continuous, which compromises the intuition $n^k \Delta_{i,k}^n X \approx F_{\frac{i-1}{n}}$. The proof of Theorem II.1.1 (iii) is complete by the following result, where we denote by $W^{k,q}$ the space of k-times absolutely continuous functions ξ on $[0,t]$ satisfying $\xi^{(k)} \in L^q([0,t])$.

Lemma II.2.7. *Let $\xi \in W^{k,q}$, and suppose that $|f(x)| \leq C(1 \vee |x|^q)$ for some $q \geq 1$ and some C. It holds that*

$$V(\xi; f,k)_t^n := n^{-1}\sum_{i=k}^{[nt]} f(n^k \Delta_{i,k}^n \xi) \to \int_0^t f(\xi_s^{(k)})\, ds,$$

as $n \to \infty$.

Proof. Assume first $\xi \in \mathcal{C}^{k+1}([0,t])$. Taylor approximation shows that

$$n^k \Delta_{i,k}^n \xi = \xi_{\frac{i-k}{n}}^{(k)} + a_{i,n},$$

where $|a_{i,n}| \leq C/n$ for all $n \geq 1, k \leq i \leq n$. We can therefore assume w.l.o.g. that f has compact support and admits a concave modulus of continuity ω_f, i.e. a continuous increasing function $\omega_f : [0,\infty) \to [0,\infty)$ with $\omega_f(0) = 0$ such that $|f(x) - f(y)| \leq \omega_f(|x-y|)$ for all x, y. We have by Jensen's inequality that

$$\limsup_{n\to\infty}\left| V(\xi,f,k)_t^n - \frac{1}{n}\sum_{i=k}^{[tn]} f\big(\xi_{\frac{i-k}{n}}^{(k)}\big)\right| \leq \limsup_{n\to\infty}\left\{\frac{[tn]}{n}\,\omega_f\left(\frac{1}{[tn]}\sum_{i=k}^{[tn]}|a_{i,n}|\right)\right\} = 0.$$

The result follows by the convergence of Riemann sums

$$\frac{1}{n}\sum_{i=k}^{[tn]} f\big(\xi_{\frac{i-k}{n}}^{(k)}\big) \to \int_0^t f(\xi_s^{(k)})\, ds.$$

In the following we extend the result to general $\xi \in W^{k,q}$ by approximating ξ with a sequence $(\xi^m)_{m\geq 1}$ of functions in $\mathcal{C}^{k+1}$. To this end, choose ξ^m such that

$$\int_0^t |\xi_s^{(k)} - \xi_s^{m,(k)}|^q \, ds \leq 1/m, \qquad \text{for all } m.$$

Indeed, the existence of such a sequence follows since continuous functions are dense in $L^q([0,t])$. Note that II.2 implies that $\int_0^t |\xi_s^{(k)} - \xi_s^{m,(k)}| \, ds \leq C/m^{1/q}$, since we assumed $q \geq 1$. The proof of the lemma will now be completed by showing that for a sequence $(\xi^m)_{m\geq 1}$ satisfying II.2 it holds that

$$\limsup_{m\to\infty} \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \, ds = 0, \tag{II.2.26}$$

and that

$$\limsup_{m\to\infty} \sup_{n\in\mathbb{N}} |V(\xi; f, k)_t^n - V(\xi^m; f, k)_t^n| = 0. \tag{II.2.27}$$

Proof of (II.2.26)*:*
Since $\xi^{m,(k)}$ converges in $L^q([0,t])$, the family $(|\xi^{m,(k)}|^q)_{m\geq 1}$ and consequently also the family $\{f(\xi^{m,(k)})_{m\geq 1}\}$ are uniformly integrable. Therefore, given $\varepsilon > 0$, there is a N such that

$$\begin{aligned}
&\int_0^t |\xi_s^{m,(k)}|^q \mathbb{1}_{\{|\xi_s^{m,(k)}|>N\}} \, ds < \varepsilon \quad \text{for all } m, \text{ and} \\
&\int_0^t |\xi_s^{(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}|>N\}} \, ds < \varepsilon.
\end{aligned} \tag{II.2.28}$$

Choosing a continuous function $\widetilde{f}_N$ with compact support such that $\widetilde{f}_N = f$ on $[-N, N]$, and denoting by ω_N a concave modulus of continuity of $\widetilde{f}_N$, we have by Jensen's inequality

$$\begin{aligned}
&\limsup_{m\to\infty} \left\{ \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \, ds \right\} \\
&\quad \leq \limsup_{m\to\infty} \left\{ t\omega_N \left(t^{-1} \int_0^t |\xi_s^{(k)} - \xi_s^{m,(k)}| \mathbb{1}_{\{|\xi_s^{(k)}| \vee |\xi_s^{m,(k)}| \leq N\}} \, ds \right) \right. \\
&\qquad\qquad \left. + \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| \vee |\xi_s^{m,(k)}| > N\}} \, ds \right\} \\
&\quad = \limsup_{m\to\infty} \left\{ \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| \vee |\xi_s^{m,(k)}| > N\}} \, ds \right\} \\
&\quad \to 0, \quad \text{as } n \to \infty.
\end{aligned}$$

Hence, (II.2.26) follows by letting $\varepsilon \to 0$ from the estimate

$$\limsup_{m\to\infty} \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| \vee |\xi_s^{m,(k)}| > N\}} \, ds < \varepsilon (4 + 2 \sup_{|x|\leq 1} |f(x)|), \tag{II.2.29}$$

which we derive in the following. Note that

$$\begin{aligned}
&\int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| \vee |\xi_s^{m,(k)}| > N\}} \, ds \\
&\leq \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| > N\}} \, ds + \int_0^t |f(\xi_s^{(k)}) - f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{m,(k)}| > N\}} \, ds \\
&= I_1 + I_2.
\end{aligned}$$

For $N > 1$ we have by (II.2.28) and II.2 that

$$\begin{aligned}
I_1 &\leq \varepsilon + \int_0^t |f(\xi_s^{m,(k)})| \mathbb{1}_{\{|\xi_s^{(k)}| > N, |\xi_s^{m,(k)}| \leq 1\}} \, ds + \int_0^t |\xi_s^{m,(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > N, |\xi_s^{m,(k)}| > 1\}} \, ds \\
&\leq \varepsilon + \sup_{|x| \leq 1} |f(x)| \int_0^t \mathbb{1}_{\{|\xi_s^{(k)}| > N\}} \, ds + \int_0^t |\xi_s^{m,(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > N\}} \, ds \\
&\leq \varepsilon(1 + \sup_{|x| \leq 1} |f(x)|) + \left\{ \left(\int_0^t |\xi_s^{m,(k)} - \xi_s^{(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > N\}} \, ds \right)^{1/q} \right. \\
&\qquad \left. + \left(\int_0^t |\xi_s^{(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > N\}} \, ds \right)^{1/q} \right\}^q \\
&\leq \varepsilon(1 + \sup_{|x| \leq 1} |f(x)|) + \left\{ 1/m^{1/q} + \varepsilon^{1/q} \right\}^q,
\end{aligned}$$

and consequently

$$\limsup_{m \to \infty} I_1 \leq \varepsilon(2 + \sup_{|x| \leq 1} |f(x)|).$$

By a similar argument it follows that $\limsup_{m \to \infty} I_2 \leq \varepsilon(2 + \sup_{|x| \leq 1} |f(x)|)$, and we obtain (II.2.29), which completes the proof of (II.2.26).

Proof of (II.2.27)*:*

In order to show (II.2.27) we split the sum

$$|V(\xi; f, k)_t^n - V(\xi^m; f, k)_t^n| \leq \frac{1}{n} \sum_{i=k}^{[tn]} \left| f(n^k \Delta_{i,k}^n \xi) - f(n^k \Delta_{i,k}^n \xi^m) \right|$$

into sums over the following sets of indices, where N and M are positive constants:

$$\begin{aligned}
A_n^N &= \{i \in \{k, ..., [tn]\} \, : \, n^k |\Delta_{i,k}^n \xi| > N\} \\
B_{m,n}^{N,M} &= \{i \in \{k, ..., [tn]\} \, : \, n^k |\Delta_{i,k}^n \xi| \leq N, \; n^k |\Delta_{i,k}^n \xi^m| > M\} \\
C_{m,n}^{N,M} &= \{i \in \{k, ..., [tn]\} \, : \, n^k |\Delta_{i,k}^n \xi| \leq N, \; n^k |\Delta_{i,k}^n \xi^m| \leq M\}.
\end{aligned}$$

and estimate the corresponding sums separately. The following relationship between $\Delta_{i,k}^n \xi$ and $\xi^{(k)}$ will be essential. For all $\xi \in W^{k,q}$ we have

$$\Delta_{i,k}^n \xi = \int_{\frac{i-1}{n}}^{i/n} \int_{s_1 - 1/n}^{s_1} \cdots \int_{s_{k-1} - 1/n}^{s_{k-1}} \xi_{s_k}^{(k)} ds_k \ldots ds_1.$$

In particular, it follows that

$$
\begin{aligned}
|n^k\Delta^n_{i,k}\xi| &\leq \int_{[0,t]^k} n^k|\xi^{(k)}_{s_k}|\mathbb{1}_{\{(s_1,\ldots,s_k)\in[(i-k)/n,i/n]^k\}}\,ds_k\ldots\,ds_1\\
&= k^{k-1}\int_{\frac{i-k}{n}}^{i/n} n|\xi^{(k)}_s|\,ds. \qquad\text{(II.2.30)}
\end{aligned}
$$

The A_n^N term: We show that for given $\varepsilon>0$ we can find sufficiently large N such that

$$
\begin{aligned}
&\limsup_{m\to\infty}\sup_{n\in\mathbb{N}}\left\{n^{-1}\sum_{i\in A_n^N}\left|f(n^k\Delta^n_{i,k}\xi)-f(n^k\Delta^n_{i,k}\xi^m)\right|\right\}\\
&\leq \limsup_{m\to\infty}\sup_{n\in\mathbb{N}}\Bigg\{n^{-1}\sum_{i\in A_n^N}|n^k\Delta^n_{i,k}\xi|^q\\
&\qquad + n^{-1}\sum_{i\in A_n^N}|n^k\Delta^n_{i,k}\xi^m|^q\mathbb{1}_{\{|n^k\Delta^n_{i,k}\xi^m|>1\}}\\
&\qquad + n^{-1}\sum_{i\in A_n^N}|f(n^k\Delta^n_{i,k}\xi^m)|\mathbb{1}_{\{|n^k\Delta^n_{i,k}\xi^m|\leq 1\}}\Bigg\}\\
&:= \limsup_{m\to\infty}\sup_{n\in\mathbb{N}}\left\{I_{1,n,N}+I_{2,n,m,N}+I_{3,n,m,N}\right\}\leq\varepsilon, \qquad\text{(II.2.31)}
\end{aligned}
$$

First we consider $I_{1,n,N}$. By (II.2.30) we have for all $i\in A_n^N$

$$
N< k^{k-1}\int_{\frac{i-k}{n}}^{i/n}|\xi^{(k)}_s|n\,ds\leq k^{k-1}\int_{\frac{i-k}{n}}^{i/n} n|\xi^{(k)}_s|\mathbb{1}_{\{|\xi^{(k)}_s|>C_{0,k}\}}\,ds+\frac{N}{2},
$$

where $C_{0,k}:=N(2k^k)^{-1}$. Therefore, again by (II.2.30), it follows that

$$
\begin{aligned}
|n^k\Delta^n_{i,k}\xi| &\leq k^{k-1}\int_{\frac{i-k}{n}}^{i/n}|\xi^{(k)}_s|n\,ds\\
&\leq 2k^{k-1}\int_{\frac{i-k}{n}}^{i/n}|\xi^{(k)}_s|n\,ds-N\\
&\leq 2k^{k-1}\int_{\frac{i-k}{n}}^{i/n}|\xi^{(k)}_s|\mathbb{1}_{\{|\xi^{(k)}_s|>C_{0,k}\}}n\,ds. \qquad\text{(II.2.32)}
\end{aligned}
$$

Consequently, recalling that $q\geq 1$, we have by Jensen's inequality

$$
\begin{aligned}
n^{-1}\sum_{i\in A_n^N}|n^k\Delta^n_{i,k}\xi|^q &\leq (2k^{k-1})^q k^{q-1}n^{-1}\sum_{i\in A_n^N}\int_{\frac{i-k}{n}}^{i/n}|\xi^{(k)}_s|^q\mathbb{1}_{\{|\xi^{(k)}_s|>C_{0,k}\}}n\,ds\\
&\leq (2k^k)^q\int_0^t|\xi^{(k)}_s|^q\mathbb{1}_{\{|\xi^{(k)}_s|>C_{0,k}\}}\,ds. \qquad\text{(II.2.33)}
\end{aligned}
$$

It follows for sufficiently large $N>0$ that

$$
\limsup_{m\to\infty}\sup_{n\in\mathbb{N}}\{I_{1,n,N}\}\leq\varepsilon. \qquad\text{(II.2.34)}
$$

Next, we argue that the same holds for the $I_{2,n,m,N}$ term. By II.2 and Minkowski's inequality it follows for any $A \in \mathcal{B}([0,t])$ that $\int_A |\xi_s^{m,(k)}|^q \, ds \le 2^{q-1} \int_A |\xi_s^{(k)}|^q \, ds + C/m$. Consequently, it holds that

$$\begin{aligned} n^{-1} \sum_{i \in A_n^N} |n^k \Delta_{i,k}^n \xi^m|^q \mathbb{1}_{\{|n^k \Delta_{i,k}^n \xi^m| > 1\}} &\le C n^{-1} \sum_{i \in A_n^N} \int_{\frac{i-k}{n}}^{i/n} |\xi_s^{m,(k)}|^q n \, ds \\ &\le C \sum_{i \in A_n^N} \int_{\frac{i-k}{n}}^{i/n} |\xi_s^{(k)}|^q \, ds + \frac{C}{m} \\ &\le C \sum_{i \in A_n^N} \int_{\frac{i-k}{n}}^{i/n} |\xi_s^{(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > C_{0,k}\}} \, ds + \frac{C}{m} \\ &\le C \int_0^t |\xi_s^{(k)}|^q \mathbb{1}_{\{|\xi_s^{(k)}| > C_{0,k}\}} \, ds + \frac{C}{m}, \end{aligned}$$

where the first inequality follows from (II.2.30), and the third from (II.2.32) in the third inequality. This shows that for sufficiently large N it holds that

$$\limsup_{m \to \infty} \sup_{n \in \mathbb{N}} \{I_{2,n,m,N}\} \le \varepsilon. \tag{II.2.35}$$

Next, we estimate the $I_{3,n,m,N}$ term. Introducing the notation

$$D_{m,n} = \{i \in \{k, ..., [tn]\} \, : \, n^k |\Delta_{i,k}^n \xi^{(m)}| \le 1\}$$

we have

$$I_{3,n,m,N} = n^{-1} \sum_{i \in A_n^N \cap D_{m,n}} |f(n^k \Delta_{i,k}^n \xi^{(m)})| \le n^{-1} |A_n^N \cap D_{m,n}| \sup_{\{|x|<1\}} |f(x)| \tag{II.2.36}$$

where $|A_n^N \cap D_{m,n}|$ denotes the number of elements of $A_n^N \cap D_{m,n}$. Using (II.2.30) we have for all $i \in A_n^N \cap D_{m,n}$

$$N - 1 \le n^k |\Delta_{i,k}^n (\xi^{(k)} - \xi^{m,(k)})| \le k^{k-1} \int_{\frac{i-k}{n}}^{i/n} |\xi_s^{(k)} - \xi_s^{m,(k)}| n \, ds,$$

and it follows that

$$|A_n^N \cap D_{m,n}| \le \frac{nk^k}{N-1} \int_0^t |\xi_s^{(k)} - \xi_s^{m,(k)}| n \, ds \le \frac{nk^k t}{(N-1) m^{1/q}},$$

where we recall II.2. With (II.2.36) it follows that for all $N > 1$ we have

$$\limsup_{m \to \infty} \sup_{n \in \mathbb{N}} \{I_{3,n,m,N}\} = 0. \tag{II.2.37}$$

Combining (II.2.34), (II.2.35) and (II.2.37) we conclude that (II.2.31) holds for sufficiently large N.

The $B^{N,M}_{m,n}$ term: We show that for any $\varepsilon > 0$ and any $N > 0$ we can find a sufficiently large M such that

$$\begin{aligned}
&\limsup_{m\to\infty} \sup_{n\in\mathbb{N}} \Big\{ n^{-1} \sum_{i\in B^{N,M}_{m,n}} \big| f(n^k\Delta^n_{i,k}\xi) - f(n^k\Delta^n_{i,k}\xi^m)\big| \Big\} \\
&\quad\le \limsup_{m\to\infty} \sup_{n\in\mathbb{N}} \Big\{ n^{-1} \sum_{i\in B^{N,M}_{m,n}} \big| f(n^k\Delta^n_{i,k}\xi)\big| + n^{-1} \sum_{i\in B^{N,M}_{m,n}} |n^k\Delta^n_{i,k}\xi^m|^q \Big\} \\
&\quad:= \limsup_{m\to\infty} \sup_{n\in\mathbb{N}} \{J^1_{n,m,N,M} + J^2_{n,m,N,M}\} < \varepsilon. \qquad \text{(II.2.38)}
\end{aligned}$$

The argument for $J^1_{n,m,N,M}$ is similar to the one used for $I_{3,m,n,N}$ above. We assume that $M > N$. For $i \in B^{N,M}_{m,n}$ it holds by (II.2.30) that

$$M - N < n^k |\Delta^n_{i,k}(\xi - \xi^m)| \le k^{k-1} n \int_{\frac{i-k}{n}}^{i/n} |\xi_s - \xi^m_s|\, ds.$$

Consequently, we have for all $m \in \mathbb{N}$

$$|B^{N,M}_{m,n}| \le \frac{k^k n}{M-N} \int_0^t |\xi_s - \xi^m_s|\, ds \le \frac{k^k n t}{(M-N) m^{1/q}},$$

where $|B^{N,M}_{m,n}|$ denotes the number of elements in $B^{N,M}_{m,n}$. Then, it follows that for all $M > N$

$$\begin{aligned}
&\limsup_{m} \sup_{n} \{J^1_{n,m,N,M}\} \\
&\quad\le \limsup_{m} \sup_{n} \{n^{-1} |B^{N,M}_{m,n}| \sup_{s\in[-N,N]} |f(s)|\} \\
&\quad\le \limsup_{m} \sup_{n} \Big\{ \frac{k^k}{(M-N) m^{1/q}} \sup_{s\in[-N,N]} |f(s)| \Big\} = 0. \qquad \text{(II.2.39)}
\end{aligned}$$

For $J^2_{n,m,N,M}$ we obtain by arguing as in (II.2.33) with $\xi^{(k)}$ replaced by $\xi^{m,(k)}$ and N replaced by M that

$$J^2_{n,m,N,M} \le (2k^k)^q \int_0^t |\xi^{m,(k)}_s|^q \mathbb{1}_{\{|\xi^{m,(k)}_s| > M/2k^k\}}\, ds,$$

for all m, n, N. Since $(|\xi^{m,(k)}|^q)_{m\ge 1}$ is uniformly integrable we can for $\varepsilon > 0$ find sufficiently large M such that

$$\limsup_{m} \sup_{n} \{J^2_{n,m,N,M}\} \le \varepsilon. \qquad \text{(II.2.40)}$$

Now, (II.2.38) follows from (II.2.39) and (II.2.40).

The $C^{N,M}_{m,n}$ term: We show that for all $N, M > 0$ we have that

$$\limsup_{m\to\infty} \sup_{n\in\mathbb{N}} \Big\{ n^{-1} \sum_{i\in C^{N,M}_{m,n}} \big| f(n^k\Delta^n_{i,k}\xi) - f(n^k\Delta^n_{i,k}\xi^m)\big| \Big\} = 0. \qquad \text{(II.2.41)}$$

Since $|n^k \Delta_{i,k}^n \xi| \leq N$ and $|n^k \Delta_{i,k}^n \xi^m| \leq M$ for all $i \in C_{m,n}^{N,M}$, we can replace f by a continuous function $\widetilde{f}_{N,M}$ with compact support, such that $f(x) = \widetilde{f}_{N,M}(x)$ for all $x \in [-(N \vee M), N \vee M]$. Denote by $\widetilde{\omega}_{N,M}$ a concave modulus of continuity for $\widetilde{f}_{N,M}$. It holds that

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \Big\{ n^{-1} \sum_{i \in C_{m,n}^{N,M}} \big| f(n^k \Delta_{i,k}^n \xi) - f(n^k \Delta_{i,k}^n \xi^m) \big| \Big\} \\
&\quad = \sup_{n\in\mathbb{N}} \Big\{ n^{-1} \sum_{i \in C_{m,n}^{N,M}} \big| \widetilde{f}_{N,M}(n^k \Delta_{i,k}^n \xi) - \widetilde{f}_{N,M}(n^k \Delta_{i,k}^n \xi^m) \big| \Big\} \\
&\quad \leq \sup_{n\in\mathbb{N}} \Big\{ [tn]/n \, \widetilde{\omega}_{N,M} \Big([tn]^{-1} \sum_{j=k}^{[tn]} n^k |\Delta_{i,k}^n \xi - \Delta_{i,k}^n \xi^m| \Big) \Big\} \\
&\quad \leq t \, \widetilde{\omega}_{N,M} \Big(t^{-1} k^k \int_0^t |\xi_s^{(k)} - \xi_s^{m,(k)}| \, ds \Big),
\end{aligned}$$

where we used (II.2.30) in the last inequality. Now (II.2.41) follows by II.2.

Finally, by (II.2.31), (II.2.38) and (II.2.41) we can for any $\varepsilon > 0$ find sufficiently large N, M such that

$$\limsup_{m\to\infty} \sup_{n\to\infty} \Big(n^{-1} \sum_{i=k}^{[tn]} \big| f(n^k \Delta_{i,k}^n \xi) - f(n^k \Delta_{i,k}^n \xi^m) \big| \Big) < \varepsilon.$$

By letting $\varepsilon \to 0$ we obtain (II.2.27) and the proof of the lemma is complete. □

II.3 Proof of Theorem II.1.2

Throughout this section we assume that the conditions of Theorem II.1.2 are satisfied. We begin by introducing some notation followed by a brief outline of the proofs.

For any function ψ on the real line we denote

$$D^k \psi(y) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \psi(y - j).$$

The following functions and processes will be frequently used throughout the proofs of both parts of the theorem.

$$g_n(s) := n^\alpha g(s/n), \qquad \phi_t^n(s) := D^k g_n(t-s), \quad \text{and} \quad Y_t^n := \int_{-\infty}^t \phi_t^n(s) dL_s, \tag{II.3.42}$$

for $n \in \mathbb{N}$. By our conditions on the function g it holds that $g_n(s) \to s_+^\alpha$, and consequently $\phi_t^n(s) \to h_k(t-s)$ as $n \to \infty$, where h_k was defined in Section II.1. Therefore, we complement (II.3.42) by defining

$$\phi_t^\infty(s) := h_k(t-s), \quad \text{and} \quad Y_t^\infty := \int_{-\infty}^t h_k(t-s) dL_s.$$

By self-similarity of L it holds that $\{n^H \Delta_{r,k}^n X\}_{r=k,\dots,n} \overset{\mathrm{d}}{=} \{Y_r^n\}_{r=k,\dots,n}$, and to deduce the theorem we show convergence in distribution under proper scaling of

$$S_n := \sum_{r=k}^{n} \big(f(Y_r^n) - \mathbb{E}[f(Y_r^n)]\big) = \sum_{r=k}^{n} V_r^n$$

where we denoted $V_r^n := f(Y_r^n) - \mathbb{E}[f(Y_r^n)]$ for brevity. In order to outline the strategy for the proof of Theorem II.1.2 (i) we recall that $(\mathcal{F}_t)_{t\in\mathbb{R}}$ denotes the filtration generated by L and introduce additionally the σ-algebras

$$\mathcal{G}_s^1 := \sigma(L_r - L_u \,|\, s \le r, u \le s+1),$$

remarking that $(\mathcal{G}_s^1)_{s\in\mathbb{R}}$ is not a filtration. For $n \ge 1, m, r \ge 0$ we denote

$$\zeta_{r,j}^n := \mathbb{E}[V_r^n|\mathcal{F}_{r-j+1}] - \mathbb{E}[V_r^n|\mathcal{F}_{r-j}] - \mathbb{E}[V_r^n|\mathcal{F}_{r-j}^1],$$
$$R_r^n := \sum_{j=1}^{\infty} \zeta_{r,j}^n \quad \text{and} \quad Q_r^n := \sum_{j=1}^{\infty} \mathbb{E}[V_r^n \,|\, \mathcal{G}_{r-j}^1]. \tag{II.3.43}$$

The sums R_r^n and Q_r^n converge almost surely, as we argue in Remark 1. We obtain the decomposition

$$S_n = \sum_{r=k}^{n} R_r^n + \sum_{r=k}^{n} (Q_r^n - Z_r) + \sum_{r=k}^{n} Z_r, \tag{II.3.44}$$

where $(Z_r)_{r\ge k}$ is a sequence of i.i.d. random variables, to be defined in (II.3.49) below. In the proof of Theorem II.1.2 we argue that the first two sums are asymptotically negligible and that the random variables Z_r are in the domain of attraction of a $(k-\alpha)\beta$-stable random variable with location parameter 0, scale parameter ρ_S and skewness parameter η_S as defined in (II.3.62) in the proof. We remark that similar decompositions have been successful to derive stable limit theorems for discrete time moving averages, see for example [19].

For the proof of Theorem II.1.2 (ii) we approximate S_n by

$$S_{n,m} = \sum_{r=k}^{n} \big(f(Y_r^{n,m}) - \mathbb{E}[f(Y_r^{n,m})]\big), \quad \text{where} \quad Y_r^{n,m} := \int_{r-m}^{r} \phi_r^n(s) dL_s.$$

More precisely, the main part of the proof is to derive the identity

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}[n^{-1}(S_n - S_{m,n})^2] = 0.$$

It is then sufficient to establish asymptotic normality of $(S_{n,m})_{n\in\mathbb{N}}$, which follows by the central limit theorem for m-dependent sequences of random variables. This general approach to deriving central limit theorems is popular in the literature, see [23] for an example.

Throughout the proof we will frequently use that for a deterministic function ψ and $a < b \in \mathbb{R}$ the integral $\int_a^b \psi(s) dL_s$ is symmetric β-stable distributed with scale parameter

$$\rho_L \bigg(\int_a^b |h(s)|^\beta ds \bigg)^{1/\beta} = \rho_L \|h\|_{L^\beta([a,b])}, \tag{II.3.45}$$

see [26, Proposition 3.4.1]. Moreover, we recall that for a symmetric β-stable random variable S with scale parameter 1 and $\gamma > \beta$ there is a $C > 0$ such that

$$\mathbb{E}[(\rho S)^\gamma \mathbb{1}_{\{|\rho S| \le 1\}}] \le C\rho^\beta, \quad \text{and} \quad \mathbb{P}(|\rho S| > 1) \le C\rho^\beta \quad \text{for all } \rho \in (0, 1]. \tag{II.3.46}$$

For the proof of this result we refer to [7, Lemma 5.5]. The function ϕ_j^n introduced above satisfies the estimate

$$\|\phi_j^n\|_{L^\beta([0,1])} \le Cj^{\alpha-k}, \tag{II.3.47}$$

for all $j \in \mathbb{N}$, which follows from Taylor approximation and the condition (A2) in Section II.1. Moreover, it satisfies the following estimate that was derived in [7, Eq. (5.92)]. There is a $C > 0$ such that for all $n \in \mathbb{N}$ and $j \in \mathbb{N}$

$$\|\phi_j^n - \phi_j^\infty\|_{L^\beta([0,1])} \le Cn^{-1}j^{\alpha-k+1}. \tag{II.3.48}$$

Recalling the definition of Φ_ρ and k_ρ^* in Section II.1 we have the following important equivalence.

Lemma II.3.1. *Let $K \subset (0, \infty)$ be bounded away from 0 by ε, i.e. $K \cap [0, \varepsilon) = \emptyset$. The following statements are equivalent.*

(i) *$k_\rho^* > 1$ for all $\rho \in K$.*

(ii) *There is a constant $C_\varepsilon > 0$ such that for all $x, y \in \mathbb{R}$ and for all $\rho \in K$ it holds that*

$$|\Phi_\rho(x) - \Phi_\rho(y)| \le C_\varepsilon\big\{(1 \wedge |x| + 1 \wedge |y|)|x - y|\mathbb{1}_{\{|x-y| \le 1\}} + \mathbb{1}_{\{|x-y|>1\}}\big\}.$$

(iii) *There is a constant $C_\varepsilon > 0$ such that for all $x \in \mathbb{R}$ and for all $\rho \in K$ it holds that*

$$|\Phi_\rho(x)| \le C_\varepsilon(1 \wedge x^2).$$

Proof. We first derive (i) $\Rightarrow$ (ii). By [28, Lemma 3.1] all derivatives of Φ_ρ are uniformly bounded by some C_ε, for all $\rho \in K$, since K is bounded away from 0. In particular, $|\Phi_\rho(x) - \Phi_\rho(y)|\mathbb{1}_{\{|x-y|>1\}} \le C_\varepsilon \mathbb{1}_{\{|x-y|>1\}}$ follows immediately. For $x < y, |x - y| \le 1$ we have $|\Phi_\rho(x) - \Phi_\rho(y)| \le \int_x^y |\Phi_\rho'(z)|dz \le C_\varepsilon|x - y|$. Moreover, as $\Phi_\rho'(0) = 0$, it holds that

$$|\Phi_\rho(x) - \Phi_\rho(y)| \le \int_x^y \int_0^z |\Phi''_{\rho^\infty}(u)|du\, dz \le C_\varepsilon|x - y|||x| + |y||,$$

and (ii) follows. (ii) $\Rightarrow$ (iii) follows by letting $y = 0$. (iii) $\Rightarrow$ (i) follows by Taylor expansion of Φ_ρ. $\square$

Proof of Theorem II.1.2 (i)

In order to define the sequence $(Z_r)_{r\geq k}$ used in (II.3.44) we let

$$U^n_{j,r} := \int_r^{r+1} \phi^n_j(s)dL_s, \quad \text{where } n\in\mathbb{N}\cup\{\infty\} \text{ and } j\geq k,$$

and denote

$$\rho^n_j := \|\phi^n_j\|_{L^\beta(\mathbb{R}\setminus[0,1])}, \quad \text{and} \quad \rho^n := \|\phi^n_1\|_{L^\beta(\mathbb{R})}.$$

Then, Z_r is defined as

$$Z_r := \sum_{j=1}^{\infty} \big\{\Phi_{\rho^\infty_j}(U^\infty_{j+r,r}) - \mathbb{E}[\Phi_{\rho^\infty_j}(U^\infty_{j+r,r})]\big\}, \tag{II.3.49}$$

where the sum is almost surely absolutely convergent, see Remark 1. Since for all $j\geq 0$ the sequence $(U^\infty_{j+r,r})_{r\geq k}$ is i.i.d., so is $(Z_r)_{r\geq k}$. By the decomposition (II.3.44), the proof of Theorem II.1.2 (i) is divided into three parts. First we show that

$$n^{\frac{1}{(\alpha-k)\beta}} \sum_{r=k}^{n} R^n_r \xrightarrow{\mathbb{P}} 0. \tag{II.3.50}$$

Thereafter, we argue that

$$n^{\frac{1}{(\alpha-k)\beta}} \sum_{r=k}^{n} (Q^n_r - Z_r) \xrightarrow{\mathbb{P}} 0. \tag{II.3.51}$$

In the third part of the proof we show that the random variables $(Z_r)_{r\geq k}$ are in the domain of attraction of a $(k-\alpha)\beta$-stable distributed with location parameter 0, scale parameter ρ_S and skewness parameter η_S, as defined in (II.3.62), which then implies the convergence (II.1.4).

Proof of (II.3.50)*:* Define for $l\geq j$ the random variables

$$\begin{aligned} \vartheta^n_{r,j,l} &= \mathbb{E}[\zeta^n_{r,j}\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l}] - \mathbb{E}[\zeta^n_{r,j}\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l-1}] \\ &= \mathbb{E}[f(Y^n_r)\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l}] - \mathbb{E}[f(Y^n_r)\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l-1}] \\ &\quad - \big\{\mathbb{E}\big[\mathbb{E}[f(Y^n_r)\,|\,\mathcal{G}_{r-j}]\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l}\big] - \mathbb{E}\big[\mathbb{E}[f(Y^n_r)\,|\,\mathcal{G}_{r-j}]\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l-1}\big]\big\}. \end{aligned} \tag{II.3.52}$$

Note that $\mathbb{E}[\vartheta^n_{r,j,l}|\mathcal{G}^1_{r-j}]$ for all $l\geq j$. It holds that $\lim_{l\to\infty}\mathbb{E}[\zeta^n_{r,j}\,|\,\mathcal{G}^1_{r-j}\vee\mathcal{G}_{r-l}]=0$, a.s., which implies the decomposition

$$\zeta^n_{r,j} = \sum_{l=j}^{\infty} \vartheta^n_{r,j,l}, \tag{II.3.53}$$

Using that the sequence $(\vartheta^n_{r,j,l})_{l=j,\ldots}$ is orthogonal, i.e. $\mathbb{E}[\vartheta^n_{r,j,l}\vartheta^n_{r,j,l'}]=0$ for $l\neq l'$, and applying Lemma II.4.2 with $\gamma=2$ we obtain

$$\mathbb{E}[|\zeta^n_{r,j}|^2] = \sum_{l=j}^{\infty}\mathbb{E}[|\vartheta^n_{r,j,l}|^2] \leq Cj^{2(\alpha-k)\beta+1}. \tag{II.3.54}$$

We can now rewrite

$$\sum_{r=k}^{n} R_r^n = \sum_{s=-\infty}^{n-1} M_s^n, \quad \text{where} \quad M_s^n = \sum_{r=1\vee(s+1)}^{n} \zeta_{r,r-s}^n$$

are martingale differences. Exploiting the orthogonality of martingale differences, it follows from the estimate (II.3.54) that

$$\mathbb{E}\Bigg[\Big(\sum_{r=k}^{n} R_r^n\Big)^2\Bigg] \leq C(n + n^{2(\alpha-k)\beta+4}).$$

For details we refer to the proof of [7, Equation (5.22)]. Therefore, (II.3.50) follows by the assumption $\alpha \in (k-2/\beta, k-1/\beta)$, which implies $1 + \frac{2}{\beta(\alpha-k)} < 0$ and

$$2(\alpha-k)\beta + 4 + \frac{2}{\beta(\alpha-k)} = \frac{2}{\beta(\alpha-k)}((\alpha-k)\beta+1)^2 < 0.$$

Proof of (II.3.51)*:* The estimation of this term uses similar methods as the proof of [7, Proposition 5.2]. Substituting $s = r - j$ in Q_r^n we obtain the expression

$$\sum_{r=k}^{n} Q_r^n = \sum_{s=-\infty}^{n-1} \sum_{j=(k-s)\vee 1}^{n-s} \mathbb{E}[V_{s+j}^n | \mathcal{G}_s^1].$$

Therefore we can make the decomposition

$$\sum_{r=k}^{n} (Q_r^n - Z_r) = H_n^{(1)} + H_n^{(2)},$$

where

$$H_n^{(1)} = \sum_{s=-\infty}^{k-1} \sum_{j=k-s}^{n-s} \Big\{\mathbb{E}[V_{s+j}^n|\mathcal{G}_s^1] - \big\{\Phi_{\rho_j^\infty}(U_{s+j,s}^\infty) - \mathbb{E}[\Phi_{\rho_j^\infty}(U_{s+j,s}^\infty)]\big\}\Big\}$$

$$H_n^{(2)} = \sum_{s=k}^{n-1} \sum_{j=1}^{n-s} \Big\{\mathbb{E}[V_{s+j}^n|\mathcal{G}_s^1] - \big\{\Phi_{\rho_j^\infty}(U_{s+j,s}^\infty) - \mathbb{E}[\Phi_{\rho_j^\infty}(U_{s+j,s}^\infty)]\big\}\Big\}.$$

We use that by definition of V_{s+j}^n and $\mathcal{G}_s^1$ it holds that

$$\mathbb{E}[V_{s+j}^n|\mathcal{G}_s^1] = \Phi_{\rho_j^n}(U_{s+j,s}^n) - \mathbb{E}[\Phi_{\rho_j^n}(U_{s+j,s}^n)]. \tag{II.3.55}$$

We argue first that for sufficiently large N the set $\{\rho_j^n : n \in \{N, ..., \infty\}, j \in \mathbb{N}\}$ is bounded away from 0. Choose $\varepsilon > 0$ such that $\rho^\infty > \varepsilon$ and $\rho_j^\infty > \varepsilon$ for all $j \in \mathbb{N}$. By Lemma II.3.47 it holds that $\rho^n \to \rho^\infty$ and we can choose N sufficiently large such that $|\rho^n - \rho^\infty| < \varepsilon/3$ for all $n > N$. By (II.3.47) we can find a $J > 0$ such that for all $j > J$ and all n it holds that $|\rho_j^n - \rho^n| < \varepsilon/3$, implying that $\rho_j^n > \varepsilon/3$ for all $j > J$ and $n > N$. For $j \in \{1, ..., J\}$ we use that $\rho_j^n \to \rho_j^\infty > \varepsilon$ as $n \to \infty$, which again follows from (II.3.47). Therefore, choosing N larger if necessary, we obtain $\rho_j^n > \varepsilon/3$ for all $j \in \mathbb{N}$ and $n > N$. Now with Lemma II.3.1 we obtain for $H_n^{(1)}$ the estimate

$$\begin{aligned}
\mathbb{E}[|H_n^{(1)}|] &\leq 2 \sum_{s=-\infty}^{k-1} \sum_{j=k-s}^{n-s} \left\{\mathbb{E}[|\Phi_{\rho_j^n}(U_{s+j,s}^n)|] + \mathbb{E}[|\Phi_{\rho_j^\infty}(U_{s+j,s}^n)|]\right\} \\
&\leq C \sum_{s=-\infty}^{k-1} \sum_{j=k-s}^{n-s} \left\{\mathbb{E}[(U_{s+j,s}^n)^2 \wedge 1] + \mathbb{E}[(U_{s+j,s}^n)^2 \wedge 1]\right\} \\
&\leq C \sum_{s=-k+1}^{\infty} \sum_{j=k+s}^{n+s} j^{(\alpha-k)\beta} \\
&= C\Big(\sum_{s=-k+1}^{n} \sum_{j=k+s}^{n+s} j^{(\alpha-k)\beta} + \sum_{s=n+1}^{\infty} \sum_{j=k+s}^{n+s} j^{(\alpha-k)\beta} \Big) \\
&\leq C\Big(\sum_{s=-k+1}^{n} s^{(\alpha-k)\beta+1} + \sum_{s=n+1}^{\infty} n s^{(\alpha-k)\beta} \Big) \leq C n^{(\alpha-k)\beta+2}.
\end{aligned}$$

The third inequality uses (II.3.46) and (II.3.47), and the last two inequalities follow from $-1 < (\alpha-k)\beta < -2$. Since $(\alpha-k)\beta + 2 + \frac{1}{(\alpha-k)\beta} = \frac{1}{(\alpha-k)\beta}((\alpha-k)\beta+1)^2 < 0$, we obtain

$$n^{\frac{1}{(\alpha-k)\beta}} H_n^{(1)} \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \to \infty. \tag{II.3.56}$$

For the estimation of $H_n^{(2)}$ we use that $H_n^{(2)}$ is of the form $H_n^{(2)} = \sum_{s=k^{n-1}} Z_s^{(n)}$ where for each fixed n, $\{Z_s^{(n)} : s = k, ..., n-1\}$ are martingale differences. Since $(k-\alpha)\beta \in (1,2)$, we can choose $q \in [1,2] \setminus \{\beta\}$ with $(k-\alpha-1)\beta < q < (k-\alpha)\beta$. It follows from the von Bahr-Esseen inequality [30, Theorem 1] that

$$\begin{aligned}
\mathbb{E}[|H_n^{(2)}|^q] &\leq C \sum_{s=k}^{n-1} \mathbb{E}[|Z_s^{(n)}|^q] \\
&\leq C \sum_{s=k}^{n-1} \Big(\sum_{j=1}^{n-s} \Big\| \mathbb{E}[V_{s+j}^n | \mathcal{G}_s^1] - \Big\{ \Phi_{\rho_j^\infty}(U_{s+j,s}^\infty) - \mathbb{E}[\Phi_{\rho_j^\infty}(U_{s+j,s}^\infty)] \Big\} \Big\|_q \Big)^q \\
&\leq C \sum_{s=1}^{n-k} \Big(\sum_{j=1}^{s} \Big\| \mathbb{E}[V_{n-s+j}^n | \mathcal{G}_{n-s}^1] - \Big\{ \Phi_{\rho_j^\infty}(U_{n-s+j,n-s}^\infty) - \mathbb{E}[\Phi_{\rho_j^\infty}(U_{n-s+j,n-s}^\infty)] \Big\} \Big\|_q \Big)^q \\
&\leq C n \Big(\sum_{j=1}^{n} \Big\| \Phi_{\rho_j^n}(U_{j,0}^n) - \Phi_{\rho_j^\infty}(U_{j,0}^\infty) \Big\|_q \Big)^q \leq C n \Big(\sum_{j=1}^{n} \|C_j^n\|_q + \sum_{j=1}^{n} \|D_j^n\|_q \Big)^q,
\end{aligned} \tag{II.3.57}$$

where $C_j^n = \Phi_{\rho_j^n}(U_{j,0}^n) - \Phi_{\rho_j^n}(U_{j,0}^\infty)$, and $D_j^n = \Phi_{\rho_j^n}(U_{j,0}^\infty) - \Phi_{\rho_j^\infty}(U_{j,0}^\infty)$. In the fourth inequality we used the representation (II.3.55) and $\|Z - \mathbb{E}[Z]\|_q \leq 2\|Z\|_q$ for any random variable Z. For the estimation of the first sum we use [7, Lemma 5.4], (II.3.47) and (II.3.48) to obtain that for $\varepsilon > 0$ sufficiently small

$$\begin{aligned}\sum_{j=1}^{n} \|C_j^n\|_q \le C \sum_{j=1}^{n} \Big\{ &\Big(\|\phi_j^n\|_{L^\beta([0,1])}^{(\beta-q)/q-\varepsilon} + \|\phi_j^\infty\|_{L^\beta([0,1])}^{(\beta-q)/q-\varepsilon} \Big) \|\phi_j^n - \phi_j^\infty\|_{L^\beta([0,1])}^{1-\varepsilon} \mathbb{1}_{\{\beta>q\}} \\ &+ \|\phi_j^n - \phi_j^\infty\|_{L^\beta([0,1])}^{\beta/q} \Big\} \\ \le C \sum_{j=1}^{n} &\Big\{ j^{(\alpha-k)\{(\beta-q)/q-\varepsilon\}} n^{-1+\varepsilon} j^{(\alpha-k+1)(1-\varepsilon)} \mathbb{1}_{\{\beta>q\}} \Big\} + (n^{-1} j^{\alpha-k+1})^{\beta/q} \\ \le C &\Big(n^{-1+\varepsilon} \sum_{j=1}^{n} j^{(\alpha-k)\beta/q+1+\varepsilon'\}} \mathbb{1}_{\{\beta>q\}} + n^{-\beta/q} \sum_{j=1}^{n} j^{(\alpha-k+1)\beta/q} \Big) \\ \le C &n^{(\alpha-k)\beta/q+1+\varepsilon+\varepsilon'}, \end{aligned} \tag{II.3.58}$$

where $\varepsilon' = \varepsilon(2(k-\alpha)-1)$. In the last inequality we used that $q \ge 1$ implies that $(\alpha-k)\beta/q + 1 + \varepsilon' > -1$ for all $\varepsilon > 0$, and that $(\alpha-k+1)\beta/q > -1$.

For the D_j^n term we apply [7, Lemma 5.3] to obtain the estimate

$$\big|(\rho_j^n)^\beta - (\rho_j^\infty)^\beta\big| \le 2\big|\|\phi_j^n\|_{L^\beta(\mathbb{R})} - \|\phi_j^\infty\|_{L^\beta(\mathbb{R})}\big| \le C n^{(\alpha-k)\beta+1}.$$

Applying Corollary II.4.7 we have that

$$\begin{aligned}\|D_j^n\|_q &\le C\big|(\rho_j^n)^\beta - (\rho_j^\infty)^\beta\big| \||U_{j,0}^n|^2 \wedge 1\|_q \\ &\le C\big|(\rho_j^n)^\beta - (\rho_j^\infty)^\beta\big| \|\phi_j^n\|_{L^\beta([0,1])}^{\beta/q} \le C n^{(\alpha-k)\beta+1} j^{(\alpha-k)\beta/q},\end{aligned}$$

where we used (II.3.46) and (II.3.47). Since $(\alpha-k)\beta/q < -1$, we obtain

$$\sum_{j=1}^{n} \|D_j^n\|_q \le C n^{(\alpha-k)\beta+1} \le n^{(\alpha-k)\beta/q+1}, \tag{II.3.59}$$

where we used $q \ge 1$. From (II.3.57), (II.3.58) and (II.3.59) we deduce that for any $\varepsilon > 0$ there is a constant C such that

$$n^{\frac{1}{(\alpha-k)\beta}} \mathbb{E}[|H_n^{(2)}|] \le n^{\frac{1}{(\alpha-k)\beta}} \|H_n^{(2)}\|_q \le C n^{(\alpha-k)\beta/q+1+1/q+\frac{1}{(\alpha-k)\beta}+\varepsilon} := C n^{\eta+\varepsilon}. \tag{II.3.60}$$

We show that $\eta < 0$. Since $q \ge 1$, the function $\xi : x \mapsto x^2 + (q+1)x + q$ is decreasing on $(-\infty, -q]$ and satisfies $\xi(-q) = 0$. Recalling that $q < \beta(k-\alpha)$, this implies $\xi((\alpha-k)\beta) > 0$ and thus $\eta = \frac{\xi((\alpha-k)\beta)}{(\alpha-k)\beta q} < 0$. Now, (II.3.51) follows from (II.3.56) and (II.3.60).

We turn now to the third step of the proof, i.e. we show that Z_r is in the domain of attraction of a $(k-\alpha)\beta$-stable random variable. This part is divided into two steps. First we define the random variable

$$Q := \overline{\Phi}(L_{k+1} - L_k) - \mathbb{E}[\overline{\Phi}(L_{k+1} - L_k)], \quad \text{where} \quad \overline{\Phi}(x) := \sum_{j=1}^{\infty} \Phi_{\rho_j^\infty}(\phi_j^\infty(0)x)$$

and show that it is in the domain of attraction of a $(k-\alpha)\beta$-stable random variable S with scale parameter ρ_S and skewness parameter η_S. Thereafter we argue that for some $r > (k-\alpha)\beta$ we have that

$$\mathbb{P}(|Z_k - Q| > x) \le C x^{-r}, \quad \text{for all } x \ge 1. \tag{II.3.61}$$

By an application of Markov's inequality it follows then that Z_k is in the domain of attraction of S as well, and an application of [26, Theorem 1.8.1] shows the convergence (II.1.4).

Let us first remark that the function $\overline{\Phi}$ and the random variable Q are well-defined. Indeed, since $\rho_j^\infty \to \rho^\infty$, the set $\{\rho_j^\infty\}_{j\in\mathbb{N}}$ is bounded away from 0 and it follows from Lemma II.3.1 that

$$|\overline{\Phi}(x)| \le C\sum_{j=1}^{\infty}(|\phi_j^\infty(0)x|^2 \wedge 1) \le C\sum_{j=1}^{\infty}(|j^{(\alpha-k)}x|^2 \wedge 1) \le C|x|^2\sum_{j=1}^{\infty} j^{2(\alpha-k)}.$$

Since $2 > \frac{1}{k-\alpha}$, it follows that $\overline{\Phi}$ and Q are well-defined. Moreover, an application of the dominated convergence theorem shows that $\overline{\Phi}$ is continuous. In order to show that Q is in the domain of attraction of a $(k-\alpha)\beta$-stable random variable with scale parameter ρ_S and skewness parameter η_S we now determine constants c_-, c_+ such that

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q<-x) = c_-, \quad \lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q>x) = c_+.$$

Indeed, it follows then from [26, Theorem 1.8.1] that Q is in the domain of attraction of a $(k-\alpha)\beta$-stable with scale parameter ρ_S and skewness parameter η_S, given by

$$\rho_S := \left(\frac{c_+ + c_-}{\tau_{(k-\alpha)\beta}}\right)^{1/(k-\alpha)\beta}, \qquad \text{and} \qquad \eta_S := \frac{c_+ - c_-}{c_+ + c_-}. \tag{II.3.62}$$

Here the constant τ_γ is for $\gamma \in (0,2)$ defined as

$$\tau_\gamma := \begin{cases} \frac{1-\gamma}{\Gamma(2-\gamma)\cos(\pi\gamma/2)} & \text{if } \gamma \neq 1, \\ \pi/2 & \text{if } \gamma = 1. \end{cases} \tag{II.3.63}$$

See (II.3.66) and (II.3.67) below for the definition of c_+ and c_-. respectively.

In order to derive c_+ and c_- explicitly, we remark that for $x>0$ it holds by substituting $t = (x/u)^{1/(k-\alpha)}$ that

$$\begin{aligned} x^{1/(\alpha-k)}\overline{\Phi}(x) &= x^{1/(\alpha-k)}\int_0^\infty \Phi_{\rho^\infty_{1+[t]}}(\phi^\infty_{1+[t]}(0)x)dt \\ &= \frac{1}{k-\alpha}\int_0^\infty \Phi_{\rho^\infty_{1+[(x/u)^{1/(k-\alpha)}]}}\big(\phi^\infty_{1+[(x/u)^{1/(k-\alpha)}]}(0)x\big)u^{-1+1/(\alpha-k)}du \\ &\to \frac{1}{k-\alpha}\int_0^\infty \Phi_{\rho^\infty}(k_\alpha u)u^{-1+1/(\alpha-k)}du := \kappa_+, \quad \text{as } x\to\infty, \end{aligned} \tag{II.3.64}$$

where $k_\alpha = \alpha(\alpha-1)\ldots(\alpha-k+1)$. In the last line we use that $\{\Phi_{\rho_j^\infty}(x) : j \in \mathbb{N}\cup\{\infty\}, x\in\mathbb{R}\}$ is a bounded set by Lemma II.3.1 since ρ_j^∞ is bounded away from 0. Therefore, the convergence follows from the dominated convergence theorem, where we remark that for all $t\in\mathbb{R}$ there is by the mean value theorem a $\xi_t \in [t-k-1, t]$ such that

$$\phi^\infty_{[t]}(0) = h_k([t]) = k_\alpha(\xi_t)_+^{\alpha-k},$$

which implies the convergence

$$\phi^\infty_{1+[(x/u)^{1/(k-\alpha)}]}(0)x \to k_\alpha u, \quad \text{as } x\to\infty.$$

Similarly we obtain for $x < 0$ that

$$|x|^{1/(\alpha-k)}\overline{\Phi}(x) \to \frac{1}{k-\alpha}\int_{-\infty}^{0} \Phi_{\rho^\infty}(k_\alpha u)|u|^{-1+1/(\alpha-k)}du := \kappa_-, \quad \text{as } x \to -\infty. \tag{II.3.65}$$

We argue next that

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q > x) = \tau_\beta \rho_L\big(\kappa_+^{k-\alpha}\mathbb{1}_{\{\kappa_+>0\}} + \kappa_-^{k-\alpha}\mathbb{1}_{\{\kappa_->0\}}\big) := c_+, \tag{II.3.66}$$

where τ_β was defined in (II.3.63). To this end we make the decomposition

$$\mathbb{P}(Q > x) = \mathbb{P}(Q > x, L_{k+1} - L_k > 0) + \mathbb{P}(Q > x, L_{k+1} - L_k < 0),$$

and analyse the two summands separately. Consider the first summand and assume $\kappa_+ > 0$. By (II.3.64) it follows that $\overline{\Phi}(y) \to \infty$ as $y \to \infty$ and we have for sufficiently large x that

$$\mathbb{P}(\overline{\Phi}(L_{k+1} - L_{k+1}) > x, L_{k+1} - L_k > 0) = \mathbb{P}(|\overline{\Phi}(L_{k+1} - L_{k+1})| > x, L_{k+1} - L_k > 0).$$

Replacing $\overline{\Phi}$ with $|\overline{\Phi}|$ allows us to apply Lemma II.4.5 with $\psi(x) = \overline{\Phi}(x)$ and $\xi(x) = x^{1/(k-\alpha)}\kappa_+$, and we obtain from (II.3.64) that

$$\begin{aligned}\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q > x, L_{k+1} - L_k > 0) &= \lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}\big(\kappa_+^{k-\alpha}(L_{k+1} - L_k) > x^{k-\alpha}\big) \\ &= \tau_\beta \rho_L^\beta \kappa_+^{(k-\alpha)\beta}.\end{aligned}$$

The second identity follows from [26, Property 1.2.15], where we recall that $L_{k+1} - L_k \sim \mathrm{S}\beta\mathrm{S}$ with scale parameter ρ_L. If $\kappa_+ < 0$, it follows from (II.3.64) that $\limsup_{x\to\infty}\overline{\Phi}(x) \le 0$ and therefore that $\overline{\Phi}(x)$ is bounded for $x \ge 0$. We obtain

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q > x, L_{k+1} - L_k > 0) = 0.$$

The same identity holds for $\kappa_+ = 0$, as follows from Lemma II.4.5, (II.3.64), and the estimate

$$\mathbb{P}(\overline{\Phi}(L_{k+1} - L_{k+1}) > x, L_{k+1} - L_k > 0) \le \mathbb{P}(|\overline{\Phi}(L_{k+1} - L_{k+1})| > x, L_{k+1} - L_k > 0).$$

We conclude that

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q > x, L_{k+1} - L_k > 0) = \tau_\beta \rho_L \kappa_+^{k-\alpha}\mathbb{1}_{\{\kappa_+>0\}}.$$

By similar arguments, applying Lemma II.4.5 on the function $\psi(x) = \overline{\Phi}(-x)$, we deduce from (II.3.65) the convergence

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q > x, L_{k+1} - L_k < 0) = \tau_\beta \rho_L \kappa_-^{k-\alpha}\mathbb{1}_{\{\kappa_->0\}},$$

which completes the proof of (II.3.66). Arguing similarly for $\mathbb{P}(Q < -x)$ we derive that

$$\lim_{x\to\infty} x^{(k-\alpha)\beta}\mathbb{P}(Q < -x) = \tau_\beta \rho_L\big(|\kappa_+|^{k-\alpha}\mathbb{1}_{\{\kappa_+<0\}} + |\kappa_-|^{k-\alpha}\mathbb{1}_{\{\kappa_-<0\}}\big) := c_-. \tag{II.3.67}$$

This shows that Q is in the domain of attraction of a $(k-\alpha)\beta$-stable random variable with location parameter 0, and scale and skewness parameters as given in (II.3.62).

Now the proof of the theorem is completed by showing (II.3.61). To this end it is by Markov's inequality sufficient to show that $\mathbb{E}[|Z_k - Q|^r] < \infty$ for some $r > (k-\alpha)\beta$. Since $(k-\alpha)\beta > 1$ an application of Minkowski's inequality yields

$$\|Z_k - Q\|_r \leq \sum_{j=1}^{\infty} \left\| \Phi_{\rho_j^\infty}(U_{j+k,k}^\infty) - \Phi_{\rho_j^\infty}\big(\phi_j^\infty(0)(L_{k+1} - L_k)\big) \right\|_r. \tag{II.3.68}$$

We remark that by the mean value theorem there is a constant $C > 0$ such that for all $x \in [0,1]$ and $j \in \mathbb{N}$ it holds that

$$|\phi_{j+k}^\infty(x) - \phi_j^\infty(0)| = |h_k(j+k+x) - h_k(j)| \leq C j^{\alpha-k-1}.$$

Since $\{\rho_j^\infty\}_{j\in\mathbb{N}}$ is bounded away from 0, there is a $\delta > 0$ with $\delta < \rho_j^\infty$ for all j. Letting $r_\varepsilon = (k-\alpha)\beta + \varepsilon$ with $\varepsilon \in (0,\delta)$, an application of Lemma II.4.3 yields

$$\begin{aligned}&\left\| \Phi_{\rho_j^\infty}(U_{j+k,k}^\infty) - \Phi_{\rho_j^\infty}(\phi_j^\infty(0)(L_{k+1} - L_k)) \right\|_{r_\varepsilon} \\ &\leq C\big(\|\phi_{j+k}^\infty - \phi_j^\infty(0)\|_{L^\beta([0,1])}^{1-\varepsilon} + \|\phi_{j+k}^\infty - \phi_j^\infty(0)\|_{L^\beta([0,1])}^{\frac{1}{k-\alpha+\varepsilon/\beta}}\big) \leq C\big(j^{(\alpha-k-1)(1-\varepsilon)} + j^{\frac{\alpha-k-1}{k-\alpha+\varepsilon/\beta}}\big).\end{aligned}$$

For sufficiently small $\varepsilon > 0$, both powers are smaller than -1, which together with (II.3.68) implies $\|Z_k - Q\|_r < \infty$, and thus (II.3.61). Since Q is in the domain of attraction of a $(k-\alpha)\beta$-stable random variable with scale parameter ρ_S and skewness parameter η_S, and $r > (k-\alpha)\beta$, so is Z_k. This completes the proof of Theorem (II.1.2). □

Proof of Theorem II.1.2 (ii)

We recall the definition of $Y_t^n, Y_r^{n,m}, S_n$ and $S_{n,m}$ from the beginning of this section and define additionally, for $a < b$, $a, b \in [0,\infty]$

$$Y_r^{n,[a,b]} := \int_{r-b}^{r-a} \phi_r^n(s)dL_s,$$

and $Y_r^{n,m} = Y_r^{n,[0,m]}$. By [11, Theorem 3.2], the statement of the theorem follows if we derive the following three identities;

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}[n^{-1}(S_n - S_{m,n})^2] = 0, \tag{II.3.69}$$

$$\frac{1}{\sqrt{n}} S_{n,m} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \eta_m^2), \quad \text{for some } \eta_m^2 \in [0,\infty), \text{ and} \tag{II.3.70}$$

$$\eta_m^2 \to \eta^2, \quad \text{as } m \to \infty. \tag{II.3.71}$$

We show (II.3.70) and (II.3.71) first. The sequence $(Y_r^{n,m})_{r=1,\ldots}$ is stationary and we denote $\theta_j^{n,m} = \operatorname{cov}(f(Y_k^{n,m}), f(Y_{k+j}^{n,m}))$ for $n \in \mathbb{N} \cup \{\infty\}$. The variance of $S_{n,m}$ is then given by

$$n^{-1}\operatorname{var}(S_{n,m}) = n^{-1}\left\{(n-k+1)\theta_0^{n,m} + 2\sum_{j=1}^{m}(n-k-j)\theta_j^{n,m}\right\}.$$

The covariances $\theta_j^{n,m}$ converge to $\theta_j^{\infty,m}$ for all m, j, as $n \to \infty$, by the following argument. The random variables $Y_k^{n,m} - Y_k^{\infty,m}$ are symmetric β-stable distributed with scale parameter $\|\phi_1^n - \phi_1^\infty\|_{L^\beta([1-m,1])} \le \|\phi_1^n - \phi_1^\infty\|_{L^\beta(\mathbb{R})}$, which converges to 0 by Lemma II.4.4. Consequently, it holds that $\mathbb{E}[|Y_k^{n,m} - Y_k^{\infty,m}|^p] \to 0$ for all $p < \beta$, which by boundedness and continuity of f implies $\mathbb{E}[(f(Y_k^n) - f(Y_k^\infty))^2] \to 0$ and it follows that $\theta_j^{n,m} \to \theta_j^{\infty,m}$. Since the sequence $(Y_r^{n,m})_{r=k,\dots}$ is m-dependent, (II.3.70) follows now from the central limit theorem for m-dependent sequences, see [10], with the limiting variance

$$\eta_m^2 = \theta_0^{\infty,m} + 2\sum_{j=1}^m \theta_j^{\infty,m}. \tag{II.3.72}$$

Next we argue that η_m^2 is a Cauchy sequence, which then shows (II.3.71) with $\eta^2 := \lim_{m\to\infty} \eta_m^2$. This is indeed an immediate consequence of (II.3.69) and (II.3.70) since

$$\begin{aligned}
|\eta_m^2 - \eta_r^2| &= \left| \lim_{n\to\infty} \left\{ n^{-1}\big(\mathrm{var}(S_{n,m}) - \mathrm{var}(S_{n,r})\big)\right\}\right| \\
&\le \left| \limsup_{n\to\infty} \left\{ n^{-1}\big(\mathrm{var}(S_n - S_{n,m}) + \mathrm{var}(S_n - S_{n,r})\big)\right\}\right| \to 0,
\end{aligned}$$

as $m, r \to \infty$ by (II.3.69). The proof of (II.1.5) can now be completed by showing (II.3.69), which we do in the following.

As in the last section, we denote by $(\mathcal{F}_r)_{r\in\mathbb{R}}$ the filtration generated by the Lévy process, i.e. $\mathcal{G}_r = \sigma(L_s - L_u : s, u \le r)$. Our goal is to show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \big(n^{-1}\mathbb{E}[(S_n - S_{n,m})^2]\big) = 0.$$

We can express S_n and $S_{n,m}$ as the telescoping sums

$$\begin{aligned}
S_n &= \sum_{r=k}^n \sum_{j=1}^\infty (\mathbb{E}[f(Y_r^n)|\mathcal{G}_{r-j+1}] - \mathbb{E}[f(Y_r^n)|\mathcal{G}_{r-j}]), \\
S_{n,m} &= \sum_{r=k}^n \sum_{j=1}^m (\mathbb{E}[f(Y_r^{n,m})|\mathcal{G}_{r-j+1}] - \mathbb{E}[f(Y_r^{n,m})|\mathcal{G}_{r-j}]).
\end{aligned}$$

Indeed, the first telescoping sum coincides with S_n almost surely, since by the backwards martingale convergence theorem and Kolmogorov's 0-1 law it holds that $\mathbb{E}[f(Y_r^n)|\mathcal{G}_{r-j}] \xrightarrow{\text{a.s.}} \mathbb{E}[f(Y_r^n)]$, as $j \to \infty$. We denote for $n \ge 1$ and $m, r, j \ge 0$

$$\zeta_{r,j}^{n,m} = \mathbb{E}[f(Y_r^n) - f(Y_r^{n,m})|\mathcal{G}_{r-j+1}] - \mathbb{E}[f(Y_r^n) - f(Y_r^{n,m})|\mathcal{G}_{r-j}].$$

and obtain

$$S_n - S_{n,m} = \sum_{r=k}^n \sum_{j=1}^\infty \zeta_{r,j}^{n,m}.$$

Now we use the estimate

$$\begin{aligned}
&n^{-1}\mathbb{E}[(S_n - S_{n,m})^2] \\
&\le 3n^{-1}\mathbb{E}\left[\left(\sum_{r=k}^n \sum_{j=m+1}^\infty \zeta_{r,j}^{n,m}\right)^2\right] + 3n^{-1}\mathbb{E}\left[\left(\sum_{r=k}^n \sum_{j=2}^m \zeta_{r,j}^{n,m}\right)^2\right] + 3n^{-1}\mathbb{E}\left[\left(\sum_{r=k}^n \zeta_{r,1}^{n,m}\right)^2\right],
\end{aligned}$$

and show that each summand on the right hand side converges to 0. Observing that

$$\operatorname{cov}(\zeta_{r,j}^{n,m}, \zeta_{r',j'}^{n,m}) = 0, \quad \text{unless } r - j = r' - j',$$

an application of Cauchy-Schwarz' inequality and Fatou's lemma yields

$$n^{-1}\mathbb{E}[(S_n - S_{n,m})^2] \le 3n^{-1}Q_{n,1,m} + 3n^{-1}Q_{n,2,m} + 3n^{-1}Q_{n,3,m},$$

where

$$\begin{aligned} Q_{n,1,m} &= \sum_{r=k}^{n} \sum_{j=m+1}^{\infty} \sum_{j'=m+1}^{\infty} \mathbb{E}[(\zeta_{r,j}^{n,m})^2]^{1/2} \mathbb{E}[(\zeta_{r',j'}^{n,m})^2]^{1/2}, \\ Q_{n,2,m} &= \sum_{r=k}^{n} \sum_{j=2}^{m} \sum_{j'=2}^{m} \mathbb{E}[(\zeta_{r,j}^{n,m})^2]^{1/2} \mathbb{E}[(\zeta_{r',j'}^{n,m})^2]^{1/2}, \quad \text{and} \\ Q_{n,3,m} &= \sum_{r=k}^{n} \mathbb{E}[(\zeta_{r,1}^{n,m})^2], \end{aligned}$$

where we denoted $r' = r - j + j'$. For the proof of (II.1.5) it remains to show that

$$\limsup_{n\to\infty} \frac{1}{n} Q_{n,i,m} \to 0, \quad \text{as } m \to \infty, \text{ for } i = 1, 2, 3.$$

Estimation of $Q_{n,1,m}$: Throughout this argument the index $r \in \{k, \dots, n\}$ is arbitrary but fixed. We recall that $(Y_r^{n,j})_{r\ge 0}$ is a stationary sequence. We introduce the notation

$$\widetilde{f}_j^n(x) = \mathbb{E}\big(f(x + Y_r^{n,j})\big),$$

which allows us to write $\mathbb{E}[f(Y_r^n)|\mathcal{G}_{r-j}] = \widetilde{f}_j^n(Y_r^{n,[j,\infty]})$. In the sum $Q_{n,1,m}$ we have $j > m$, implying that $\mathbb{E}[f(Y_r^{n,m})|\mathcal{G}_{r-j+1}] = \mathbb{E}[f(Y_r^{n,m})|\mathcal{G}_{r-j}]$. Thus we can write

$$\zeta_{r,j}^{n,m} = \widetilde{f}_{j-1}^n\big(Y_r^{n,[j-1,\infty]}\big) - \widetilde{f}_j^n\big(Y_r^{n,[j,\infty]}\big), \quad \text{for } j > m.$$

Observe that $Y_r^{n,[j-1,\infty]} = Y_r^{n,[j-1,j]} + Y_r^{n,[j,\infty]}$ and denote by $F_{[j-1,j]}^n$ and $F_{[j,\infty]}^n$ the corresponding distribution functions. Then, it follows that

$$\mathbb{E}(\zeta_{r,j}^{n,m})^2 = \int_{\mathbb{R}} \int_{\mathbb{R}} \big(\widetilde{f}_{j-1}^n(u+v) - \widetilde{f}_j^n(u)\big)^2 dF_{[j-1,j]}(v) dF_{[j,\infty]}(u).$$

Using moreover that $\widetilde{f}_j^n(u) = \mathbb{E}\big(f\big(u + Y_r^{n,j-1} + Y_r^{n,[j-1,j]}\big)\big) = \int_{\mathbb{R}} \widetilde{f}_{j-1}^n(u+z) dF_{[j-1,j]}(z)$, we obtain

$$\begin{aligned} \mathbb{E}(\zeta_{r,j}^{n,m})^2 &= \int_{\mathbb{R}} \int_{\mathbb{R}} \bigg(\int_{\mathbb{R}} D(u,v,z) dF_{[j-1,j]}(z)\bigg)^2 dF_{[j-1,j]}(v) dF_{[j,\infty]}(u) \\ &\le \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} D^2(u,v,z) dF_{[j-1,j]}(z) dF_{[j-1,j]}(v) dF_{[j,\infty]}(u), \end{aligned}$$

where $D(u,v,z) = \widetilde{f}_{j-1}^n(u+v) - \widetilde{f}_{j-1}^n(u+z)$. Our goal is to use mean value theorem to derive an upper bound for this integral. By Lemma 3.2 of [23] the l-th derivative of $\widetilde{f}^n_{j-1}$ is bounded by $C_l(\rho^n_{j-1})^{-l}(1+|\log(\rho^n_{j-1})|+\log^2(\rho^n_{j-1}))$, where ρ^n_{j-1} is the scale parameter of the symmetric β-stable random variable $Y^{n,j-1}_r$, i.e. $\rho^n_{j-1}=\big(\int_{r-j+1}^r|\phi^n_r(s)|^\beta ds\big)^{1/\beta}$. We have for all $j\geq 2$ that

$$(\rho^n_{j-1})^\beta=\int_{r-j+1}^r|\phi^n_r(s)|^\beta ds\geq\int_{r-1}^r|\phi^n_r(s)|^\beta ds.$$

The right hand side is positive for all n and converges to $\int_0^1 s^{\alpha\beta}ds>0$ as $n\to\infty$, by the dominated convergence theorem, since by Assumption (A) there is a constant C such that $|\phi^n_r(s)|\leq C|r-s|^\alpha$ for all $s\in[r-1,r]$ and all $n\geq 1$. Consequently, the scale parameters ρ^n_{j-1} are bounded away from 0 for all $j\geq 2$, $n\geq 1$, and [23, Lemma 3.2] implies that for all $l\geq 0$ there is a constant C_l such that

$$|\widetilde{f}^{n,(l)}_{j-1}(x)|<C_l \tag{II.3.73}$$

for all $j\geq 2$, all $n\in\mathbb{N}$ and all $x\in\mathbb{R}$. Now, applying mean value theorem on $D(u,v,z)=\widetilde{f}^n_{j-1}(u+v)-\widetilde{f}^n_{j-1}(u+z)$, it follows that

$$(D(u,v,z))^2\leq C\min(1,(v-z)^2),$$

where the constant does not depend on j or n. Consequently, we obtain

$$\begin{aligned}\limsup_{n\to\infty}\mathbb{E}[(\zeta^{n,m}_{r,j})^2]\ \leq& C\int_{|v-z|\leq 1}(v-z)^2dF^n_{[j-1,j]}(z)dF^n_{[j-1,j]}(v)\\ &+C\int_{|v-z|>1}dF^n_{[j-1,j]}(z)dF^n_{[j-1,j]}(v)\\ &=C\mathbb{E}[(S^1_{n,j}-S^2_{n,j})^2\mathbb{1}_{\{|S^1_{n,j}-S^2_{n,j}|\leq 1\}}]+C\mathbb{P}(|S^1_{n,j}-S^2_{n,j}|>1),\end{aligned}$$

where $S^1_{n,j}$ and $S^2_{n,j}$ are independent symmetric β-stable random variables with scale parameter $\big(\int_{r-j}^{r-j+1}|\phi^n_r(s)|^\beta ds\big)^{1/\beta}=\|\phi^n_j\|_{L^\beta([0,1])}$, see (II.3.45). Consequently $S^1_{n,j}-S^2_{n,j}\overset{\mathrm{d}}{=}2^{1/\beta}\|\phi^n_j\|_{L^\beta([0,1])}S$ where S is symmetric β-stable with scale parameter 1. It follows now from (II.3.47) and (II.3.46) that there is a constant C such that

$$\mathbb{E}[(\zeta^{n,m}_{r,j})^2]\leq Cj^{\beta(\alpha-k)},$$

for all $n,m,r\in\mathbb{N}$ and $j>m$, and we obtain

$$\frac{1}{n}Q_{n,1,m}\leq C\frac{1}{n}\sum_{r=k}^n\Big(\sum_{j=m+1}j^{\beta(\alpha-k)/2}\Big)^2.$$

This shows that $\limsup_{n\to\infty}n^{-1}Q_{n,1,m}\to 0$ as $m\to\infty$ since $\beta(\alpha-k)/2<-1$.

Estimation of $Q_{n,2,m}$: For $j\leq m$ we obtain

$$\zeta^{n,m}_{r,j}=\widetilde{f}^n_{j-1}\big(Y^{n,[j-1,\infty]}_r\big)-\widetilde{f}^n_j\big(Y^{n,[j,\infty]}_r\big)-\big\{\widetilde{f}^n_{j-1}\big(Y^{n,[j-1,m]}_r\big)-\widetilde{f}^n_j\big(Y^{n,[j,m]}_r\big)\big\}.$$

The involved random variables can be decomposed into the sum of independent random variables as

$$\begin{aligned}Y^{n,[j-1,\infty]}_r&=Y^{n,[j-1,j]}_r+Y^{n,[j,m]}_r+Y^{n,[m,\infty]}_r\\ Y^{n,[j,\infty]}_r&=Y^{n,[j,m]}_r+Y^{n,[m,\infty]}_r\\ Y^{n,[j-1,m]}_r&=Y^{n,[j-1,j]}_r+Y^{n,[j,m]}_r.\end{aligned}$$

Denoting by $F^n_{[j-1,j]}, F^n_{[j,m]}$ and $F^n_{[m,\infty]}$ the corresponding distribution functions, we obtain

$$\mathbb{E}[(\zeta^{n,m}_{r,j})^2] = \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \big\{\widetilde{f}^n_{j-1}(u+v+w) - \widetilde{f}^n_j(v+w) \\ - \big(f^n_{j-1}(u+v) - \widetilde{f}^n_j(v)\big)\big\}^2 dF^n_{[j-1,j]}(u)dF^n_{[j,m]}(v)dF^n_{[m,\infty]}(w).$$

Using the relation $\widetilde{f}^n_j(x) = \mathbb{E}f\big(x + Y^{n,j-1}_r + Y^{n,[j-1,j]}_r\big) = \int_{\mathbb{R}} \widetilde{f}^n_{j-1}(x+z)dF^n_{[j-1,j]}(z)$, we obtain

$$\mathbb{E}[(\zeta^{n,m}_{r,j})^2] = \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\bigg(\int_{\mathbb{R}} D(u,v,w,z)dF^n_{[j-1,j]}(z)\bigg)^2 dF^n_{[j-1,j]}(u)dF_{[j,m]}(v)dF^n_{[m,\infty]}(w) \\ \leq \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} D^2(u,v,w,z)dF^n_{[j-1,j]}(z)dF^n_{[j-1,j]}(u)dF^n_{[j,m]}(v)dF^n_{[m,\infty]}(w),$$

where

$$D(u,v,w,z) = \widetilde{f}^n_{j-1}(u+v+w) - \widetilde{f}^n_{j-1}(v+w+z) - \big(\widetilde{f}^n_{j-1}(u+v) - \widetilde{f}^n_{j-1}(v+z)\big).$$

As we argued in (II.3.73), for $j \geq 2$ the first two derivatives of $\widetilde{f}^n_{j-1}$ are uniformly bounded with the bound not depending on j or n. Therefore, we obtain by the mean value theorem that

$$D^2(u,v,w,z) \leq C\min\{1, w^2, (u-z)^2, (u-z)^2w^2\}.$$

This leads to the estimate

$$\mathbb{E}[(\zeta^{n,m}_{r,j})^2] \leq C\bigg(\int_{|u-z|\leq 1}(u-z)^2 dF^n_{[j-1,j]}(u)dF^n_{[j-1,j]}(z) + \int_{|u-z|>1} dF^n_{[j-1,j]}(u)dF^n_{[j-1,j]}(z)\bigg) \\ \times\bigg(\int_{|w|\leq 1} w^2 dF^n_{[m,\infty]}(w) + \int_{|w|>1} dF^n_{[m,\infty]}(w)\bigg).$$

Similar as in the estimation of $Q_{n,1,m}$ we derive from (II.3.47) and (II.3.46) that $\mathbb{E}[(\zeta^{n,m}_{r,j})^2] \leq C(\rho^n_{[j-1,j]}\rho^n_{[m,\infty]})^\beta$, where $\rho^n_{[j-1,j]}$ and $\rho^n_{[m,\infty]}$ are the scale parameters of the stable distributions $F^n_{[j-1,j]}$ and $F^n_{[m,\infty]}$, respectively. By (II.3.45) and (II.3.47) the scale parameters satisfy $\rho^n_{[j-1,j]} = \|\phi^n_j\|_{L^\beta([0,1])} \leq Cj^{\alpha-k}$, and

$$(\rho^n_{[m,\infty]})^\beta = \int_{-\infty}^{r-m} |\phi^n_r(s)|^\beta ds = \sum_{l=m+1}^{\infty} \|\phi^n_l\|^\beta_{L^\beta([0,1])} < C\sum_{l=m+1}^{\infty} l^{\beta(\alpha-k)}. \tag{II.3.74}$$

It follows that

$$\mathbb{E}(\zeta^{n,m}_{r,j})^2 \leq Cj^{\beta(\alpha-k)}\sum_{l=m+1}^{\infty} l^{\beta(\alpha-k)},$$

for all $j \in \{2,...,m\}$ and we obtain

$$\limsup_{n\to\infty}\frac{1}{n}Q_{n,2,m} \leq C\bigg(\sum_{j=2}^{m} j^{\frac{\beta}{2}(\alpha-k)}\bigg)^2\bigg(\sum_{l=m+1}^{\infty} l^{\beta(\alpha-k)}\bigg),$$

which converges to 0, as $m \to \infty$ since $\beta(\alpha - k) < -2$.

Estimation of $Q_{n,3,m}$: Using the inequality $\mathbb{E}\big\{\mathbb{E}[X|\mathcal{G}] - \mathbb{E}[Y|\mathcal{F}]\big\}^2 \le 2\mathbb{E}X^2 + 2\mathbb{E}Y^2$ we obtain

$$\frac{1}{n} Q_{n,3,m} \le \frac{4}{n} \sum_{r=k}^{n} \mathbb{E}[(f(Y_r^n) - f(Y_r^{n,m}))^2] = \frac{n-k+1}{n} \mathbb{E}[(f(Y_1^n) - f(Y_1^{n,m}))^2].$$

In order to argue that $\limsup_{n\to\infty} \mathbb{E}[(f(Y_1^n) - f(Y_1^{n,m}))^2] \to 0$ as $m \to \infty$, it is by boundedness and continuity of f sufficient to show that the family of random variables $\{Y_1^n, Y_1^{n,m}\}_{m\in\mathbb{N}, n\in\mathbb{N}\cup\{\infty\}}$ is tight and satisfies

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}[|Y_1^n - Y_1^{n,m}| > \varepsilon] = 0, \quad \text{for all } \varepsilon > 0.$$

The latter follows from (II.3.74), since $Y_1^n - Y_1^{n,m}$ is SβS distributed with scale parameter $\rho^n_{[m,\infty]}$. For the tightness we first recall that $\mathbb{E}[|Y_1^n - Y_1^\infty|^p] \to 0$ for all $p < \gamma$, which follows from Lemma II.4.4, since $Y_1^n - Y_1^\infty$ is SβS distributed with scale parameter $\|\phi_1^n - \phi_1^\infty\|_{L^\beta(\mathbb{R})}$. Consequently, given $\varepsilon > 0$, we can choose N sufficiently large such that

$$\begin{aligned}
\mathbb{P}(|Y_1^{n,m}| > N) &\le \mathbb{P}(|Y_1^{n,m} - Y_1^{\infty,m}| > N/2) + \mathbb{P}(|Y_1^{\infty,m}| > N/2) \\
&\le \mathbb{P}(|Y_1^n - Y_1^\infty| > N/2) + \mathbb{P}(|Y_1^\infty| > N/2) < \varepsilon \quad \text{for all } m, n \in \{1, ..., \infty\}.
\end{aligned}$$

In the second inequality we used that all random variables are SβS distributed and that the scale parameters of $Y_1^n - Y_1^\infty$ and Y_1^∞ are greater or equal than the scale parameters of $Y_1^{n,m} - Y_1^{\infty,m}$ and $Y_1^{\infty,m}$, respectively. This shows the tightness of $\{Y_1^n, Y_1^{n,m}\}_{m\in\mathbb{N}, n\in\mathbb{N}\cup\{\infty\}}$ and it follows that $\limsup_{n\to\infty} n^{-1} Q_{n,3,m} \to 0$ as $m \to \infty$.

□

II.4 Auxiliary results

Here we give some technical results used in the proof of Theorem II.1.2. First we argue that the various telescope sum expressions used throughout the proof converge almost surely to the limit claimed in the proof.

Remark 1. We argue first that the sum Q_r^n defined in (II.3.43) is absolutely convergent with probability 1. By Kolmogorov's three-series theorem and Markov's inequality it is sufficient to show that

$$\sum_{j=1}^{\infty} \mathbb{E}[|\mathbb{E}[V_r^n | \mathcal{F}^1_{r-j}]|] < \infty.$$

Recalling the representation (II.3.55), it holds by Lemma II.3.1, (II.3.47) and (II.3.46) that

$$\mathbb{E}[|\mathbb{E}[V_r^n|\mathcal{F}^1_{r-j}]|] \le C\mathbb{E}[|\Phi_{\rho_j^n}(U^n_{r,r-j})|] \le C\mathbb{E}[|(U^n_{r,r-j})^2 \wedge 1|] \le \|\phi_j^n\|^\beta_{L^\beta([0,1])} \le Cj^{\beta(\alpha-k)} < \infty,$$

since $\beta(\alpha - k) < -1$, showing that Q_r^n is indeed well-defined. For the sum R_r^n it is now sufficient to argue that $\mathbb{E}[V_r^n|\mathcal{F}_{r-j}] \xrightarrow{\text{a.s.}} 0$, as $j \to \infty$, which is a consequence of Kolmogorov's 0-1 law and the backward martingale convergence theorem. The convergence of the sum in the definition of Z_r in (II.3.49) follows by the same argument as given for Q_r^n where we remark that $Z_r = Q_r^\infty$. The convergence of the sum and identity in (II.3.53) follows from the backward martingale convergence theorem and the fact that $\mathbb{E}[\zeta_{r,j}^n|\mathcal{F}_{r-j}^1] = 0$.

Lemma II.4.1. *For any $\varepsilon > 0$ there exists a finite constant $C > 0$ such that for all $\rho \geq \varepsilon$ and all $a \in \mathbb{R}$ we have that*

$$F(a,x,y) := \left| \int_0^y \int_0^x \Phi_\rho''(a+u+v)\, du\, dv \right| \leq C(1 \wedge x)(1 \wedge y).$$

Proof. By [23, Lemma 3.2], $\Phi_\rho(x)$, $\Phi_\rho'(x)$ and $\Phi_\rho''(x)$ are uniformly bounded for $\rho \geq \varepsilon$ and $x \in \mathbb{R}$. Boundedness of Φ_ρ'' immediately implies $F(a,x,y) \leq Cxy$. Moreover, it holds that

$$\begin{aligned}\int_0^y \int_0^x \Phi_\rho''(a+u+v)\, du\, dv &= \int_0^y \Phi_\rho'(a+x+v) - \Phi_\rho'(a+v)\, dv \\ &= \big(\Phi_\rho(a+x+y) - \Phi_\rho(a+y)\big) - \big(\Phi_\rho(a+x) - \Phi_\rho(a)\big).\end{aligned}$$

The first equality and boundedness of Φ_ρ' implies $F(a,x,y) \leq Cy$, and similarly $F(a,x,y) \leq Cx$, whereas the second equality shows that $F(a,x,y) \leq C$. $\square$

Lemma II.4.2. *For all $\gamma > \beta$ there is a $C > 0$ such that for all $n \in \mathbb{N}$, $r \in \{k,\dots,n\}$, $j \in \mathbb{N}$ and $l \geq j$ it holds that*

$$\mathbb{E}[|\vartheta_{r,j,l}^n|^\gamma] \leq C j^{(\alpha-k)\beta} l^{(\alpha-k)\beta},$$

where $\vartheta_{r,j,l}^n$ is defined in (II.3.52).

Proof. It is sufficient to consider the case $r = 1$, since for fixed j,l,n the sequence $(\vartheta_{r,j,l}^n)_{r\in\mathbb{N}}$ is stationary. Without loss of generality we assume that $l \geq 2 \vee j$. By definition of ϑ it holds that

$$\begin{aligned}\vartheta_{1,j,l}^n = \mathbb{E}[f(Y_1^n)\,|\,\mathcal{G}_{1-j}^1 \vee \mathcal{G}_{1-l}] - \mathbb{E}[f(Y_1^n)\,|\,\mathcal{G}_{1-l}] \\ - \big\{\mathbb{E}[f(Y_1^n)\,|\,\mathcal{G}_{1-j}^1 \vee \mathcal{G}_{-l}] - \mathbb{E}[f(Y_1^n)\,|\,\mathcal{G}_{-l}]\big\},\end{aligned}$$

Define for $-\infty \leq a < b \leq 1$ the random variable

$$U_{[a,b]}^n = \int_a^b \phi_1^n(s)\, dL_s.$$

Let in the following $\widetilde{L}$ be an independent copy of L and define $\widetilde{U}_{[a,b]}^n$ accordingly, and denote by $\widetilde{\mathbb{E}}$ the expectation with respect to $\widetilde{L}$ only. Moreover we denote by $\rho_{j,l}^n = \|\phi_1^n\|_{L^\beta([1-l,1-j]\cup[2-j,1])}$, i.e. the scale parameter of $\int_{1-l}^{1-j} \phi_1^n\, dL_s + \int_{2-j}^1 \phi_1^n\, dL_s$. Then, decompsing $\int_{-\infty}^1 \phi_1^n\, dL_s$ into the independent integrals

$$\int_{-\infty}^1 \phi_1^n\, dL_s = \int_{-\infty}^{-l} \phi_1^n\, dL_s + \int_{-l}^{1-l} \phi_1^n\, dL_s + \int_{1-l}^{1-j} \phi_1^n\, dL_s + \int_{1-j}^{2-j} \phi_1^n\, dL_s + \int_{2-j}^{1} \phi_1^n\, dL_s$$

we obtain the expression

$$\begin{aligned}\vartheta^n_{1,j,l} &= \widetilde{\mathbb{E}}\big[\Phi_{\rho^n_{j,l}}(U^n_{[-\infty,-l]} + U^n_{[-l,1-l]} + U^n_{[1-j,2-j]}) \\ &\quad - \Phi_{\rho^n_{j,l}}(U^n_{[-\infty,-l]} + U^n_{[-l,1-l]} + \widetilde{U}^n_{[1-j,2-j]}) \\ &\quad - \Phi_{\rho^n_{j,l}}(U^n_{[-\infty,-l]} + \widetilde{U}^n_{[-l,1-l]} + U^n_{[1-j,2-j]}) \\ &\quad + \Phi_{\rho^n_{j,l}}(U^n_{[-\infty,-l]} + \widetilde{U}^n_{[-l,1-l]} + \widetilde{U}^n_{[1-j,2-j]})\big] \\ &= \widetilde{\mathbb{E}}\bigg[\int_{\widetilde{U}^n_{[-l,1-l]}}^{U^n_{[-l,1-l]}} \int_{\widetilde{U}^n_{[1-j,2-j]}}^{U^n_{[1-j,2-j]}} \Phi''_{\rho^n_{j,l}}(U^n_{[-\infty,-l]} + u + v)\, du\, dv\bigg],\end{aligned}$$

and by substitution there is a random variable $\widetilde{W}^n_{j,l}$ such that

$$|\vartheta^n_{1,j,l}| \le \widetilde{\mathbb{E}}\bigg[\bigg|\int_0^{|\widetilde{U}^n_{[-l,1-l]} - U^n_{[-l,1-l]}|} \int_0^{|\widetilde{U}^n_{[1-j,2-j]} - U^n_{[1-j,2-j]}|} \Phi''_{\rho^n_{j,l}}(\widetilde{W}^n_{j,l} + u + v)\, du\, dv\bigg|\bigg].$$

We obtain by Lemma II.4.1 and using that $|x-y| \wedge 1 \le |x| \wedge 1 + |y| \wedge 1$ that

$$\begin{aligned}\mathbb{E}[|\vartheta^n_{1,j,l}|^\gamma] &\le C\mathbb{E}[\widetilde{\mathbb{E}}[(1 \wedge |\widetilde{U}^n_{[-l,1-l]} - U^n_{[-l,1-l]}|^\gamma)(1 \wedge |\widetilde{U}^n_{[1-j,2-j]} - U^n_{[1-j,2-j]}|^\gamma)]] \\ &\le C\mathbb{E}[\widetilde{\mathbb{E}}[1 \wedge |\widetilde{U}^n_{[-l,1-l]}|^\gamma + 1 \wedge |U^n_{[-l,1-l]}|^\gamma]]\mathbb{E}[\widetilde{\mathbb{E}}[1 \wedge |\widetilde{U}^n_{[1-j,2-j]}|^\gamma + 1 \wedge |U^n_{[1-j,2-j]}|^\gamma]] \\ &\le C\|\phi^n_1\|^\beta_{L^\beta([-l,1-l])}\|\phi^n_1\|^\beta_{L^\beta([1-j,2-j])} \le Cl^{(\alpha-k)\beta} j^{(\alpha-k)\beta},\end{aligned}$$

where we used in the second inequality the independence of both factors which follows from $l \ge j$. The third inequality uses that for a SβS random variable S with scale parameter ρ it holds that $\mathbb{E}[|S|^\gamma \wedge 1] \le C\rho^\beta$ for any $\gamma > \beta$, see (II.3.46). The last inequality follows from (II.3.47). □

Lemma II.4.3. *([7, Lemma 5.4]) For any $q \ge 1$ with $q \ne \beta$ there exists $\delta > 0$ and a finite constant C such that for all $\varepsilon \in (0,\delta)$, $\rho > \delta$ and $\kappa, \tau \in L^\beta([0,1])$ with $\|\kappa\|_{L^\beta([0,1])}, \|\tau\|_{L^\beta([0,1])} \le 1$ we have*

$$\begin{aligned}&\bigg\|\Phi_\rho\bigg(\int_0^1 \kappa(s)dL_s\bigg) - \Phi_\rho\bigg(\int_0^1 \tau(s)dL_s\bigg)\bigg\|_q \\ &\le \begin{cases} \|\kappa-\tau\|^{\beta/q}_{L^\beta([0,1])} & \beta < q \\ \big(\|\kappa\|^{(\beta-q)/q-\varepsilon}_{L^\beta([0,1])} + \|\tau\|^{(\beta-q)/q-\varepsilon}_{L^\beta([0,1])}\big)\|\kappa-\tau\|^{1-\varepsilon}_{L^\beta([0,1])} + \|\kappa-\tau\|^{\beta/q}_{L^\beta([0,1])} & \beta > q. \end{cases}\end{aligned}$$

Proof. Denote $U = \int_0^1 \kappa(s)dL_s$ and $V = \int_0^1 \tau(s)dL_s$. By Lemma II.3.1 we obtain

$$\begin{aligned}&\|\Phi_\rho(U) - \Phi_\rho(V)\|_q \\ &\qquad \le C\big\|\big(|U| \wedge 1 + |V| \wedge 1\big)|U-V|\mathbb{1}_{\{|U-V|<1\}}\big\|_q + C\mathbb{P}(|U-V| \ge 1)^{1/q}.\end{aligned}$$

For the second summand, (II.3.46) yields

$$C\mathbb{P}(|U-V| \ge 1)^{1/q} \le C\|\kappa-\tau\|^{\beta/q}_{L^\beta([0,1])}.$$

The first summand can be estimated as in [7, Lemma 5.4]. □

Lemma II.4.4. *Let $(\alpha - k)\beta < -1$. There is a constant $C > 0$ such that*

$$\|\phi_1^n - \phi_1^\infty\|_{L^\beta(\mathbb{R})} \leq C(n^{\alpha-k+1/\beta} \vee n^{-1}).$$

Proof. The function $\zeta(s) = s^{-\alpha} g(s)$ is k-times continuously differentiable on $(0, \infty)$ and can by Assumption (A2) be extended to a k-times continuously differentiable function on $\mathbb{R}$, which we also will denote ζ. By substitution it holds that

$$\int_{\mathbb{R}} |\phi_1^n(x) - \phi_1^\infty(x)|^\beta dx = \int_0^\infty |D^k g_n(s) - h_k(s)|^\beta ds.$$

For $s \geq n$ we have $|h_k(s)| \leq C s^{\alpha-k}$ and $|D^k g_n(s)| \leq C s^{\alpha-k}$ by Lemma II.2.1 and Assumption (A2). This implies together with $(\alpha - k)\beta < -1$ that

$$\int_n^\infty |D^k g_n(s) - h_k(s)|^\beta ds \leq C \int_n^\infty s^{(\alpha-k)\beta} ds \leq C n^{(\alpha-k)\beta+1}. \tag{II.4.75}$$

Using the linearity of D^k and that $\zeta(0) = 1$ it follows from the mean value theorem that

$$\begin{aligned} \int_0^k |D^k g_n(s) - h_k(s)|^\beta ds &= \int_0^k |D^k\{(\zeta(s/n) - 1)s_+^\alpha\}|^\beta ds \qquad \text{(II.4.76)} \\ &\leq C n^{-\beta} \sup_{t \in [-k/n, k/n]} |\zeta'(t)| \int_0^k s_+^{\alpha\beta} ds \leq C n^{-\beta}. \end{aligned}$$

It remains to show that

$$\int_k^n |D^k g_n(s) - h_k(s)|^\beta ds \leq C(n^{(\alpha-k)\beta+1} \vee n^{-\beta}). \tag{II.4.77}$$

Since ζ is k-times continuously differentiable by assumption, it follows from Taylor expansion for ζ of order k that for $s \in [k, n]$

$$D^k(g_n(s)) = \zeta(s/n) h_k(s) + \sum_{l=1}^k \lambda_l^n(s) n^{-l} D^{k-l}(s+l)_+^\alpha,$$

where the coefficients $\lambda_l^n(s)$ are bounded uniformly in $n \in \mathbb{N}$ and $s \in [k, n]$, see [7, Lemma 5.3] for details. It follows that

$$\begin{aligned} &\int_k^n |D^k g_n(s) - h_k(s)|^\beta ds \\ &\qquad \leq C \int_k^n |(\zeta(s/n) - 1)h_k(s)|^\beta ds + C \sum_{l=1}^k n^{-l\beta} \int_k^n (s^{\alpha-k+l})^\beta ds \\ &\qquad \leq C \sup_{t \in [0,1]} |\zeta'(t)| \int_k^n |(s/n) s^{\alpha-k}|^\beta ds + C(n^{-\beta} \vee n^{(\alpha-k)\beta+1}) \\ &\qquad \leq C(n^{(\alpha-k)\beta+1} \vee n^{-\beta}) \end{aligned}$$

This shows (II.4.77), which together with (II.4.75) and (II.4.76) completes the proof of the lemma. □

Lemma II.4.5. *Let ψ, ξ be continuous functions on $\mathbb{R}$ with $\psi \sim \xi$ for $x \to \infty$. Let X be a random variable taking values in $\mathbb{R}_+$ and $\gamma \geq 0$ such that*

$$\lim_{x\to\infty} x^\gamma \mathbb{P}(|\psi(X)| > x) = \kappa$$

where $\kappa \in [0, \infty)$. Then it holds that

$$\lim_{x\to\infty} x^\gamma \mathbb{P}(|\xi(X)| > x) = \kappa.$$

Proof. Denote $\psi(x) = \xi(x)\varphi(x)$ with $\varphi(x) \to 1$ for $x \to \infty$. Let $\varepsilon > 0$. By continuity of ψ and ξ we can choose x sufficiently large such that $\varphi(y) \in (1-\varepsilon, 1+\varepsilon)$ whenever $\min(|\psi(y)|, |\xi(y)|) > x$ and $y \geq 0$. Since X takes values in $\mathbb{R}_+$, this implies that $\varphi(X) \in (1-\varepsilon, 1+\varepsilon)$ whenever $|\psi(X)| > x$ or $|\xi(X)| > x$. It follows that

$$\begin{aligned}
x^\gamma |\mathbb{P}(|\psi(X)| > x) - \mathbb{P}(|\xi(X)| > x)| &= \mathbb{E}\big[x^\gamma \big(\mathbb{1}_{\{|\psi(X)|>x>|\xi(X)|\}} + \mathbb{1}_{\{|\psi(X)|<x<|\xi(X)|\}}\big)\big] \\
&\leq 2\mathbb{E}\big[x^\gamma \mathbb{1}_{\{\frac{x}{1+\varepsilon}<|\psi(X)|<\frac{x}{1-\varepsilon}\}}\big] \\
&= 2\mathbb{E}\big[x^\gamma \mathbb{1}_{\{\frac{x}{1+\varepsilon}<|\psi(X)|\}} - x^\gamma \mathbb{1}_{\{\frac{x}{1-\varepsilon}\leq|\psi(X)|\}}\big] \\
&\to 2\kappa((1+\varepsilon)^\gamma - (1-\varepsilon)^\gamma), \quad \text{as } x \to \infty.
\end{aligned}$$

The lemma follows by letting $\varepsilon \to 0$. □

We conclude this section by showing a sufficient criterion for the condition $k_\rho^* > 1$ for all $\rho > 0$ that is used in Theorem II.1.2 (i).

Theorem II.4.6. *Denote by f_+ and f_- the positive and negative part of f, and denote by $k_+^*(\rho)$ and $k_-^*(\rho)$ the Appell rank of f_+ and f_- at ρ, respectively. Suppose there is a $\widetilde{\rho} > 0$ such that $k_+^*(\widetilde{\rho}) > 1$ and $k_-^*(\widetilde{\rho}) > 1$. Then $k_+^*(\rho_0) > 1$ and $k_-^*(\rho_0) > 1$ for all $\rho_0 > 0$ implying that $k^*(\rho_0) > 1$ for all $\rho_0 > 0$.*

Proof. We first assume that f is nonnegative. Let $\rho_0 > 0$ be arbitrary but fixed. Let $R_0 \subset (0, \infty)$ be a compact set containing $\widetilde{\rho}$ and ρ_0. Introducing the function

$$h_\rho(y) := \int_{\mathbb{R}} |t|^\beta \cos(ty) \exp(-\rho^\beta |t|^\beta)\, dt$$

we show first that there exists a $C > 0$ such that

$$|h_\rho(y)| \leq C|1 \wedge y^{-1-\beta}| \tag{II.4.78}$$

for all $\rho \in R_0$. For sufficiently large $C > 0$ we have

$$|h_\rho(y)| \leq \int_{\mathbb{R}} |t|^\beta \exp(-\rho^\beta |t|^\beta)\, dt \leq C, \quad \text{for all } \rho \in R_0.$$

Therefore it suffices to show $|h_\rho(y)| \leq C|y^{-1-\beta}|$, which can be done along the lines of Lemma 5.8 of [7], replacing $\nu - 1 - p$ by β. Denote by g_ρ the density of a symmetric β-stable distribution with scale parameter ρ and recall that $\lim_{y\to\infty} \rho^{-\beta} y^{1+\beta} g_\rho(y) = const$. Consequently, we can find a constant C such that

$$|1 \wedge y^{-\beta-1}| \leq C g_\rho(y), \quad \text{for all } \rho \in R_0, y \in \mathbb{R}.$$

Indeed, it is easy to see that such a constant C_ρ exists for any fixed $\rho \in R_0$ and can be chosen continuously in ρ, which allows us to set $C := \sup_{R_0} C_\rho$. By (II.4.78) it follows that there is a C such that

$$h_{\rho_1}(y) \le C g_{\rho_2}(y), \quad \text{for all } \rho_1, \rho_2 \in R_0,\ y \in \mathbb{R}.$$

It follows that for all $\rho \in R_0$

$$\begin{aligned} \left| \frac{\partial}{\partial \rho} \Phi_\rho(x) \right| &= \beta \rho^{\beta-1} \int_{\mathbb{R}} f(x+y) h_\rho(y) dy \\ &\le C \rho^{\beta-1} \int_{\mathbb{R}} f(x+y) g_\rho(y) dy \\ &= C \rho^{\beta-1} \Phi_\rho(x), \end{aligned} \tag{II.4.79}$$

where we used that f is nonnegative. An application of Gronwall's lemma yields that

$$\Phi_{\rho_0}(x) \le \Phi_{\widetilde{\rho}}(x) \exp(C|\rho_0^\beta - \widetilde{\rho}^\beta|).$$

Now, $k^*(\widetilde{\rho}) > 1$ implies that $|\Phi_{\widetilde{\rho}}(x)| \le C(1 \wedge x^2)$ for all x (see Lemma II.3.1), which implies that $|\Phi_{\rho_0}(x)| \le C(1 \wedge x^2)$, and consequently $k^*(\rho_0) > 1$.

For general f satisfying the conditions of the theorem the statement follows immediately from the decomposition $f = f_+ - f_-$. □

Corollary II.4.7. *Let $R_0 \subset (0, \infty)$ be compact, and assume that $k_+^*(\widetilde{\rho}) > 1$ and $k_-^*(\widetilde{\rho}) > 1$ for some $\rho_\infty \in R_0$. There is a constant C such that for all $\rho_1, \rho_2 \in R_0$*

$$|\Phi_{\rho_1}(x) - \Phi_{\rho_2}(x)| \le C|\rho_1^\beta - \rho_2^\beta|(1 \wedge x^2),$$

for all x.

Proof. This follows immediately from the estimate (II.4.79), Lemma II.3.1 and the fundamental theorem of analysis. □

Bibliography

[1] Ayache, A. and J. Hamonier (2012). Linear fractional stable motion: A wavelet estimator of the parameter. *Statistics & Probability Letters 82*(8), 1569 – 1575.

[2] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2009). Power variation for Gaussian processes with stationary increments. *Stochastic Process. Appl. 119*(6), 1845–1865.

[3] Barndorff-Nielsen, O., J. Corcuera, and M. Podolskij (2011). Multipower variation for Brownian semistationary processes. *Bernoulli 17*(4), 1159–1194.

[4] Barndorff-Nielsen, O., J. Corcuera, M. Podolskij, and J. Woerner (2009). Bipower variation for Gaussian processes with stationary increments. *J. Appl. Probab. 46*(1), 132–150.

[5] Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From stochastic calculus to mathematical finance*, pp. 33–68. Springer, Berlin.

[6] Basse-O'Connor, A., C. Heinrich, and M. Podolskij (2017). On limit theory for Lévy semi-stationary processes. available at arXiv:1604.02307.

[7] Basse-O'Connor, A., R. Lachiéze-Rey, and M. Podolskij (2016). Power variation for a class of stationary increments levy driven moving averages. *Annals of Probability*. To appear.

[8] Basse-O'Connor, A. and J. Rosiński (2016). On infinitely divisible semimartingales. *Probab. Theory Related Fields 164*(1-2), 133–163.

[9] Benassi, A., S. Cohen, and J. Istas (2004). On roughness indices for fractional fields. *Bernoulli 10*(2), 357–373.

[10] Berk, K. (1973). A central limit theorem for m-dependent random variables with unbounded m. *Ann. Probability 1*, 352–354.

[11] Billingsley, P. (1999). *Convergence of probability measures* (Second ed.). John Wiley & Sons, Inc., New York.

[12] Braverman, M. and G. Samorodnitsky (1998). Symmetric infinitely divisible processes with sample paths in Orlicz spaces and absolute continuity of infinitely divisible processes. *Stochastic Process. Appl. 78*(1), 1–26.

[13] Cambanis, S., C. Hardin, Jr., and A. Weron (1987). Ergodic properties of stationary stable processes. *Stochastic Process. Appl. 24*(1), 1–18.

[14] Chronopoulou, A., C. Tudor, and F. Viens (2011). Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes. *Commun. Stoch. Anal. 5*(1), 161–185.

[15] Chronopoulou, A., F. Viens, and C. Tudor (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electron. J. Stat. 3*, 1393–1435.

[16] Cohen, S., C. Lacaux, and M. Ledoux (2008). A general framework for simulation of fractional fields. *Stochastic Process. Appl. 118*(9), 1489–1517.

[17] Glaser, S. (2015). A law of large numbers for the power variation of fractional Lévy processes. *Stoch. Anal. Appl. 33*(1), 1–20.

[18] Häusler, E. and H. Luschgy (2015). *Stable convergence and stable limit theorems.* Springer, Cham.

[19] Ho, H. and T. Hsing (1997). Limit theorems for functionals of moving averages. *Ann. Probab. 25*(4), 1636–1669.

[20] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic processes and their applications 118*(4), 517–559.

[21] Jacod, J. and P. Protter (2012). *Discretization of processes.* Springer, Heidelberg.

[22] Kallenberg, O. (2002). *Foundations of modern probability* (Second ed.). Springer-Verlag, New York Berlin Heidelberg.

[23] Pipiras, V. and M. Taqqu (2003). Central limit theorems for partial sums of bounded functionals of infinite-variance moving averages. *Bernoulli 9*(5), 833–855.

[24] Rajput, B. and J. Rosiński (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields 82*(3), 451–487.

[25] Rényi, A. (1963). On stable sequences of events. *Sankhyā Ser. A 25*, 293 302.

[26] Samorodnitsky, G. and M. Taqqu (1994). *Stable non-Gaussian random processes.* Chapman & Hall, New York.

[27] Skorohod, A. (1956). Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen. 1*, 289–319.

[28] Surgailis, D. (2004). Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli 10*(2), 327–355.

[29] Taqqu, M. and R. Wolpert (1983). Infinite variance self-similar processes subordinate to a Poisson measure. *Z. Wahrsch. Verw. Gebiete 62*(1), 53–72.

[30] von Bahr, B. and C. Esseen (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist 36*, 299–303.

[31] Whitt, W. (2002). *Stochastic-process limits.* Springer-Verlag, New York.

Paper III

Hybrid simulation scheme for volatility modulated moving average fields

Claudio Heinrich[1], Mikko S. Pakkanen[2] and Almut E.D. Veraart[2]

[1] *Department of Mathematics, Aarhus University, Denmark*

[2] *Department of Mathematics, Imperial College London, Great Britain*

Abstract: We develop a simulation scheme for a a class of spatial stochastic processes called volatility modulated moving averages. A characteristic feature of this model is that the behaviour of the moving average kernel at zero governs the roughness of realisations, whereas its behaviour away from zero determines global properties of the process, such as long range dependence. Our simulation scheme takes this into account and approximates the moving average kernel by a power function around zero and by a step function else. For this type of approach the authors of [7], who considered a comparable model in one dimension, coined the expression hybrid simulation scheme. We derive the asymptotic mean square error of the simulation scheme and compare it in a simulation study with several other simulation techniques.

III.1 Introduction

In this article we develop a simulation scheme for real valued random fields that we call *volatility modulated moving average* (VMMA) fields. A VMMA is defined by the formula

$$X_{\mathbf{t}} = \int_{\mathbb{R}^2} g(\mathbf{t}-\mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}), \tag{III.1.1}$$

where W is Gaussian white noise, $g \in L^2(\mathbb{R}^2)$ is a deterministic kernel, and σ is a random volatility field. This model has been used for statistical modelling of spatial phenomena throughout a variety of sciences, examples being modelling of vegetation and nitrate deposition [20], of sea surface temperature [26] and of wheat yields [29]. It is known that any stationary Gaussian random field with a continuous and integrable covariance function has a moving average representation of the form (III.1.1) with σ constant, cf. [19, Proposition 6]. Introducing the stochastic volatility factor σ allows for modelling spatial heteroscedasticity and non-Gaussian marginal distributions. We are interested in the case when the moving average kernel g has a singularity at zero. In this situation, the order of the singularity governs the roughness of the random field, specified by its Hausdorff dimension or its index of Hölder continuity. Spatial stochastic models with Hausdorff dimension greater 2 (i.e. with non-smooth realisations) are for example used in surface modelling, where it is of high importance to model the roughness of the surface accurately. Specific examples are modelling of seafloor morphology [15] or surface modelling of celestial bodies [18]. The challenge in simulating volatility modulated moving averages therefore lies in recovering the roughness of the field accurately, while simultaneously capturing the global properties of the field, such as for example long range dependence. Our hybrid simulation scheme relies on approximating the kernel g by a power function in a small neighbourhood of zero, and by a step function away from zero. This approach allows us to reproduce the explosive behavior at the origin, while simultaneously approximating the integrand on a large subset of $\mathbb{R}^2$. This idea is motivated by the recent work [7], where the authors proposed a similar simulation scheme for the simulation of the one-dimensional model of Brownian semi-stationary processes.

This article is structured as follows. In Section III.2 we introduce our model in detail and discuss some of its properties. In Section III.3 we describe the hybrid simulation scheme and derive the exact asymptotic error of the scheme. Section III.4 contains a simulation study comparing the hybrid scheme to other simulation schemes. Proofs for our theoretical results are given in Section III.5. The appendix contains some technical details and calculations.

III.2 Volatility modulated moving average fields

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and W white noise on $\mathbb{R}^2$. That is, W is an independently scattered random measure satisfying $W(A) \sim \mathcal{N}(0, \lambda(A))$ for all sets $A \in \mathcal{B}_0 = \{A \in \mathcal{B}(\mathbb{R}^2) : \lambda(A) < \infty\}$, where λ denotes the Lebesgue measure. Recall that a collection of real valued random variables $\Lambda = \{\Lambda(A) : A \in \mathcal{B}_0\}$ is called independently scattered random measure if for every sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint sets with $\lambda(\bigcup_n A_n) < \infty$, the random variables $\Lambda(A_n), n = 1, 2, ...$ are independent and $\Lambda(\bigcup_n A_n) = \sum_n \Lambda(A_n)$, almost surely.

The kernel function $g : \mathbb{R}^2 \to \mathbb{R}$ is assumed to be of the form

$$g(\mathbf{t}) = \widetilde{g}(\|\mathbf{t}\|) := \|\mathbf{t}\|^\alpha L(\|\mathbf{t}\|)$$

for some $\alpha \in (-1, 0)$, and a function $L : (0, \infty) \to (0, \infty)$ that is slowly varying at 0. Here and in the following $\|\cdot\|$ always denotes the Euclidean norm on $\mathbb{R}^2$. Recall that L is said to be slowly varying at 0 if for any $\delta > 0$

$$\lim_{x\to 0} \frac{L(\delta x)}{L(x)} = 1,$$

and that then the function $\widetilde{g}(x) = x^\alpha L(x)$ is called regularly varying at 0 of index α. The explosive behavior of the kernel at 0 is a crucial feature of this model, as it governs the roughness of the field. Indeed, under week additional assumptions the Hausdorff dimension of a realisation of X is $2 - \alpha$ with probability 1, see [17] and Theorem III.2.1, meaning that for $\alpha \to -1$ the realisations of X become extremely rough. In Figure III.1 we present samples of realisations of VMMAs for different α.

The roughness of realisations poses a challenge for simulation of volatility modulated moving averages. Indeed, the maybe most intuitive way to simulate the model (III.1.1) is by freezing the integrand over small blocks and simulating the white noise over these blocks as independent centered normal random variables with variance equaling the block size. However, this method does not account for the explosive behavior of g at 0 and therefore does a poor job in reproducing the roughness of the original process correctly, in particular for values of α close to -1. We will demonstrate this phenomenon in a simulation study in section III.4. The hybrid scheme resolves this issue by approximating g around 0 by a power kernel, and approximating it by a step function away from 0.

The integral in (III.1.1) is well defined, when σ is measurable with respect to $\mathcal{B}(\mathbb{R}^2) \otimes \mathcal{F}$ and the process $\mathbf{s} \mapsto g(\mathbf{t} - \mathbf{s})\sigma_\mathbf{s}(\omega)$ takes almost surely values in $L^2(\mathbb{R}^2)$. In particular we do not require independence of σ and W or any notion of filtration or predictability for the definition of the integral, as is usually used in the theory of temporal stochastic processes. This general theory of stochastic integration dates back to Bichteler [8], see also [24]. A brief discussion can be found in Appendix III.A. When σ and W are independent, we can realise them on a product space and it is therefore sufficient to define integration with respect to W for deterministic functions, which has been done in [27].

The volatility field $(\sigma_\mathbf{s})_{\mathbf{s}\in\mathbb{R}^2}$ is assumed to satisfy $\mathbb{E}[\sigma_\mathbf{s}^2] < \infty$ for all $\mathbf{s}$. Moreover, we assume σ to be covariance stationary, meaning that $\mathbb{E}[\sigma_\mathbf{s}]$ does not depend on $\mathbf{s}$ and $\mathrm{cov}(\sigma_{\mathbf{s}+\mathbf{r}}, \sigma_\mathbf{s}) = \mathrm{cov}(\sigma_\mathbf{r}, \sigma_0)$ for all $\mathbf{s}, \mathbf{r} \in \mathbb{R}^2$. In particular $\mathbb{E}[\sigma_\mathbf{s}^2] = \mathbb{E}[\sigma_0^2]$ for all $\mathbf{s} \in \mathbb{R}^2$. For some of our theoretical results we will assume that σ and W are independent, however we show in Appendix III.A that this is not required for the convergence of the hybrid scheme. We make the assumption that σ is sufficiently smooth such that freezing σ over small blocks will cause an asymptotically negligible

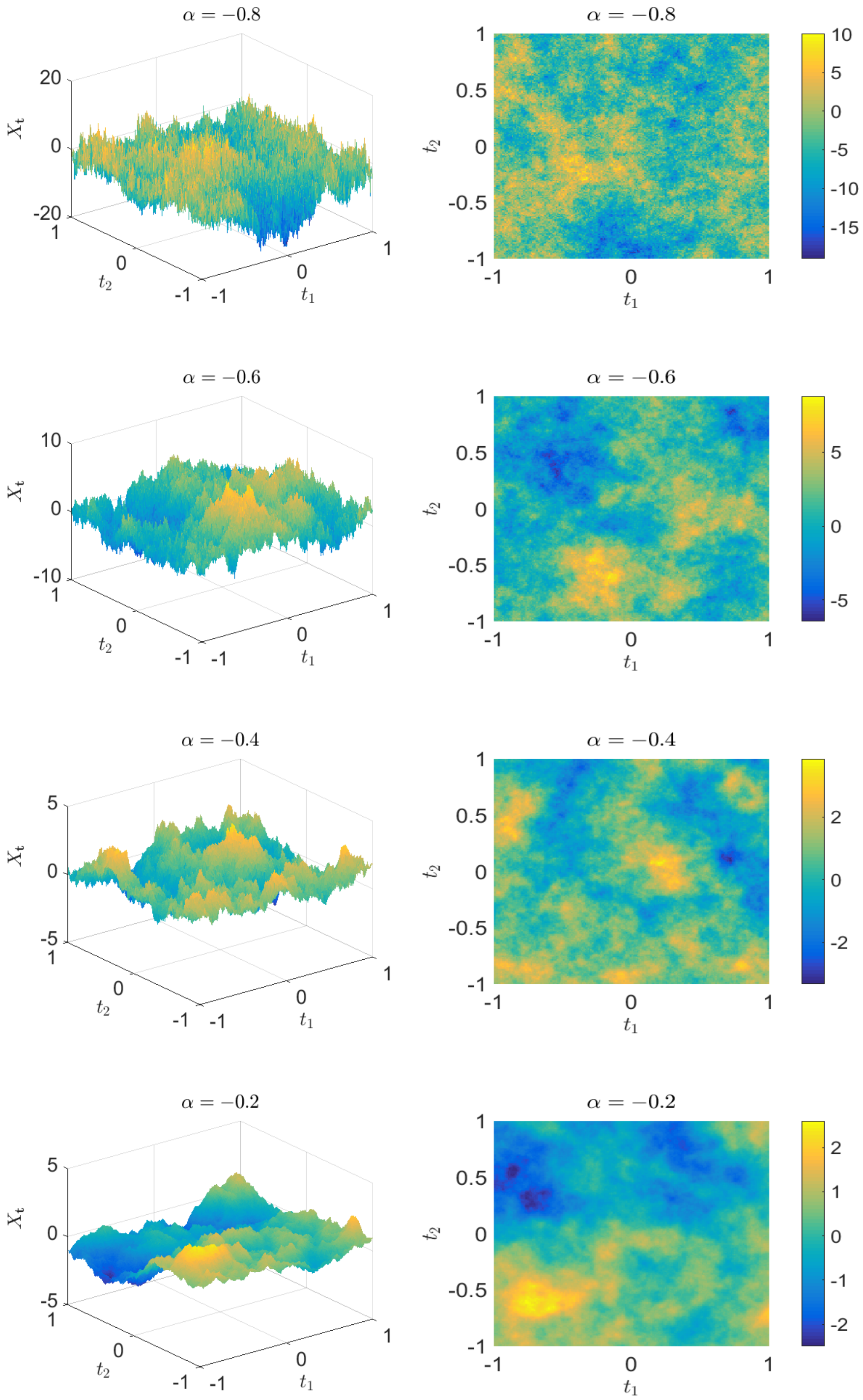

Figure III.1: Realisations of volatility modulated moving average fields for different α with Matérn covariance, see Example III.2.2. All plots range over $\mathbf{t} \in [-1,1]^2$ and are generated with constant volatility σ. In section III.4 we present examples of VMMAs with nontrivial volatility.

error in the simulation. It turns out that this is the case when σ satisfies

$$\mathbb{E}[|\sigma_0 - \sigma_{\mathbf{u}}|^2] = o(\|\mathbf{u}\|^{2\alpha+2}), \quad \text{for } \mathbf{u} \to 0. \tag{III.2.2}$$

When σ is independent of the Gaussian noise W, the covariance stationarity of σ implies that the process X is itself covariance stationary and covariance isotropic in the sense that $\mathbb{E}[(X_{\mathbf{t}+\mathbf{s}} - X_{\mathbf{t}})^2]$ depends only on $\|\mathbf{s}\|$. If σ is in fact stationary, X is stationary and isotropic.

Moreover, we pose the following assumptions on our kernel function g. They ensure in particular that g is square integrable, which together with covariance stationarity of σ ensures the existence of the integral in (III.1.1).

(A1) The slowly varying function L is continuously differentiable and bounded away from 0 on any interval $(u, \sqrt{2}]$ for $u > 0$.

(A2) It holds that $\widetilde{g}(x) = \mathcal{O}(x^{\beta})$, as $x \to \infty$, for some $\beta \in (-\infty, -1)$,

(A3) There is an $M > 0$ such that $|\widetilde{g}'|$ is decreasing on $[M, \infty)$ and satisfies

$$\int_1^\infty \widetilde{g}'(r)^2 r \; dr < \infty.$$

(A4) There is a $C > 0$ such that $|L'(x)| < C(1 + x^{-1})$ for all $x \in (0, 1]$.

An appealing feature of the VMMA model is its flexibility in modelling marginal distributions and covariance structure independently. Indeed, assuming that σ is stationary and independent of W, the covariance structure of X is entirely determined by the kernel g, whereas the marginal distribution of X is a centered Gaussian variance mixture with conditional variance $\int_{\mathbb{R}^2} g(-\mathbf{s})^2 \sigma_{\mathbf{s}}^2 \; d\mathbf{s}$, the distribution of which is governed by the distribution of σ.

The behavior of the kernel at 0 is determined by the power α, whereas its behavior away from 0, e.g. how quickly it decays at ∞, depends on the slowly varying function L. While the behavior of g at 0 determines local properties of the process X, like the roughness of realisations, the behavior of g away from 0 governs its global properties, e.g. whether it is long range dependent. Being able to independently choose α and L allows us therefore to model local and global properties of the VMMA independently, which underlines the flexibility of the model. This separation of local and global properties, and the desire to capture both of them correctly, is one of our main motivations to use a hybrid simulation scheme. We now formalise the statement that the roughness of X is determined by the power α.

Theorem III.2.1. *(i) Assume independence of σ and W. The variogram of X defined as $V(h) := \mathbb{E}[(X_0 - X_{\mathbf{t}})^2]$, where $h = \|\mathbf{t}\|$, satisfies*

$$h^{-2-2\alpha} L(h)^2 V(h) \to 2\pi \mathbb{E}[\sigma_0^2] \int_{\mathbb{R}^2} \left(\|\mathbf{x} + \mathbf{e}/2\|^{\alpha} - \|\mathbf{x} - \mathbf{e}/2\|^{\alpha}\right)^2 d\mathbf{x} \quad \text{as } h \to 0,$$

where $\mathbf{e}$ is any vector with $\|\mathbf{e}\| = 1$.

(ii) Assume additionally that the volatility is locally bounded in the sense that it satisfies $\sup_{\|\mathbf{s}\| \le M+1} \{\sigma_{\mathbf{s}}^2\} < \infty$ *almost surely, where M is as in assumption (A3). Then, for all $\varepsilon > 0$, the process X has a version with locally $\alpha + 1 - \varepsilon$-Hölder continuous realisations.*

The proof can be found in Section III.5. In [17] the authors analyse the variogram of a closely related model and derive similar results.

We conclude this section by discussing examples of possible choices for kernel functions g and volatility fields σ.

Example III.2.2 (Matérn)**.** Assume that σ is independent of W. Denote for $\nu \in (0,1)$ by K_ν the modified Bessel function of the second kind. Letting $\lambda > 0$ and

$$g(\mathbf{t}) = \|\mathbf{t}\|^{\frac{\nu-1}{2}} K_{\frac{\nu-1}{2}}(\lambda \|\mathbf{t}\|),$$

it has been argued in [22] that then the model (III.1.1) has correlation function

$$C(\|\mathbf{r}\|) = \mathbb{E}[(X_{\mathbf{r}} - X_0)^2]/\mathbb{E}[X_0^2] = \frac{(\lambda\|\mathbf{r}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu(\lambda\|\mathbf{r}\|), \quad \mathbf{r} \in \mathbb{R}^2,$$

and consequently belongs to the Matérn covariance family, cf. [25], see also [16] and references therein. We argue now that g as above satisfies our model assumptions with $\alpha = \nu - 1$. The function

$$L(x) = x^{\frac{1-\nu}{2}} K_{\frac{\nu-1}{2}}(\lambda x)$$

is continuously differentiable on $(0,\infty)$. It holds that $\lim_{x\downarrow 0} L(x) = 2^{-\frac{\nu+1}{2}} \Gamma\big(\frac{\nu-1}{2}\big)$, see [1, Eq. (9.6.9), p.375], which implies that L is slowly varying at 0 and satisfies condition (A4). Moreover, since $K_{\frac{\nu-1}{2}}(\lambda x)$ decays exponentially as $x \to \infty$, cf. [1, p.378], condition (A2) is satisfied for all $\beta < -1$. Condition (A3) follows as well from the exponential decay together with the identity

$$\frac{d}{dx}(x^{\alpha/2} K_{\alpha/2}(x)) = x^{\frac{\alpha}{2}-1} K_{\frac{\alpha}{2}-1}(x).$$

Example III.2.3 (ambit fields)**.** In a series of papers [5, 6] the authors proposed to model velocities of particles in turbulent flows by a class of spatio-temporal stochastic processes called *ambit fields*. Over the last years this model found manifold applications throughout various sciences, examples being [3, 21]. The VMMA model is a purely spatial analogue of an ambit field driven by white noise and can therefore be interpreted as a realisation of an ambit field at a fixed time t. In the framework of turbulence modeling, the squared volatility $\sigma_{\mathbf{s}}^2$ has the physical interpretation of local energy dissipation and it has been argued in [4] that it is natural to model $\sigma_{\mathbf{s}}^2$ as (exponential of an) ambit field itself. A possible model for the volatility is therefore $\sigma_{\mathbf{t}}^2 = \exp(X'_{\mathbf{t}})$ where X' is a volatility modulated moving average, independent of W. By Theorem III.2.1 (i) it is not difficult to see that this model satisfies assumption (III.2.2) when the roughness parameter α' of X' satisfies $\alpha' > \alpha$. In its core, an ambit field is a stochastic integral driven by a Lévy basis, which does not need to be Gaussian. A simulation of such integrals in the non-Gaussian case typically relies on a shot noise decomposition of the integral, as demonstrated in [28], see also [11].

III.3 The Hybrid Scheme

In this section we present the hybrid simulation scheme using the following notation. For $r>0$ and $\mathbf{t}=(t_1,t_2)\in\mathbb{R}^2$ we introduce the notation $\square_r\mathbf{t}$ for a square with side length $1/r$ centered at $\mathbf{t}$, that is $\square_r\mathbf{t}=\left[t_1-\frac{1}{2r},t_1+\frac{1}{2r}\right]\times\left[t_2-\frac{1}{2r},t_2+\frac{1}{2r}\right]$. We will suppress the index r if it is 1, and will denote $\square_r$ instead of $\square_r 0$. We simulate the process $X_{\mathbf{t}}$ for $\mathbf{t}\in[-1,1]^2$ on the square grid $\Gamma_n:=\left\{\frac{1}{n}(i,j),\ i,j\in\{-n,...,n\}\right\}$.

A first necessary step for approximating the integral (III.1.1) is to truncate the range of integration, i.e.

$$X_{\mathbf{t}}\approx\int_{\square_{1/C}\mathbf{t}}g(\mathbf{t}-\mathbf{s})\sigma_{\mathbf{s}}W(d\mathbf{s}),$$

for some large $C>0$. To ensure convergence of the simulated process as $n\to\infty$, we increase the range of integration simultaneously with increasing the grid resolution n. We let therefore $C=C_n\approx n^\gamma$ for some $\gamma>0$. More precisely, it proves to be convenient to choose $C_n=\frac{N_n+1/2}{n}$ with $N_n=[n^{1+\gamma}]$, where $[x]$ denotes the integer part of x.

An intuitive approach to simulating the model (III.1.1) is approximating the integrand on $\square_{C_n^{-1}}\mathbf{t}$ by freezing it over squares with side length $1/n$, i.e.

$$X^{R,n}_{\mathbf{t}}=\sum_{\mathbf{j}\in\mathbf{t}+\{-N_n,\dots,N_n\}^2}g(\mathbf{t}-\mathbf{b_j}/n)\sigma_{\mathbf{j}/n}\int_{\square_n\mathbf{j}}W(d\mathbf{s}),\tag{III.3.3}$$

where $\mathbf{b_j}\in\square\mathbf{j}$ are evaluation points chosen such that $\mathbf{t}-\mathbf{b_j}/n\neq 0$ for all $\mathbf{t}\in\Gamma_n$ and $\mathbf{j}\in\mathbb{Z}^2$. Indeed, $X^{R,n}_{\mathbf{t}}$ can be simulated, assuming that the volatility σ can be simulated on the square grid $\left\{\frac{1}{n}(i,j),\ i,j\in\mathbb{Z}\right\}$, since $\left\{\int_{\square_n\mathbf{j}}W(d\mathbf{s})\right\}_{\mathbf{j}\in\mathbb{Z}^2}\overset{\text{i.i.d}}{\sim}\mathcal{N}\left(0,\frac{1}{n^2}\right)$. We will refer to this simulation method as Riemann-sum scheme. The authors of [26] use this technique to simulate volatility moving averages with bounded moving average kernel and demonstrate that it performs well in this setting. In our framework, however, a crucial weakness of this approach is the inaccurate approximation of the kernel function g around its singularity at 0, which results in a poor recovery of the roughness of X.

This weakness can be overcome by choosing a small $\kappa\in\mathbb{N}_0$ (typically, $\kappa\in\{0,1,2\}$) and approximating g by a power kernel on $\frac{1}{n}[-\kappa-1/2,\kappa+1/2]^2$. More specifically, denoting $K_\kappa=\{-\kappa,\dots,\kappa\}^2$ and $\overline{K}_\kappa=\{-N_n,\dots,N_n\}^2\setminus K_\kappa$, the hybrid scheme approximates $X_{\mathbf{t}}$ by

$$\begin{aligned}X^n_{\mathbf{t}}:=&\sum_{\mathbf{j}\in K_\kappa}\sigma_{\mathbf{t}-\mathbf{j}/n}L(\|\mathbf{b_j}\|/n)\int_{\square_n(\mathbf{t}-\mathbf{j}/n)}\|\mathbf{t}-\mathbf{s}\|^\alpha W(d\mathbf{s})\\&+\sum_{\mathbf{j}\in\overline{K}_\kappa}\sigma_{\mathbf{t}-\mathbf{j}/n}g(\mathbf{b_j}/n)\int_{\square_n(\mathbf{t}-\mathbf{j}/n)}W(d\mathbf{s}).\end{aligned}\tag{III.3.4}$$

In order to simulate $X_{\mathbf{t}}$ on the grid $\mathbf{t}\in\Gamma_n$, we simulate the families of centered Gaussian random variables $\mathcal{W}^1_n$ and $\mathcal{W}^2_n$, defined as

$$\mathcal{W}^1_n:=\left\{W^n_{\mathbf{i},\mathbf{j}}=\int_{\square_n\mathbf{i}/n}\|(\mathbf{i}+\mathbf{j})/n-\mathbf{s}\|^\alpha W(d\mathbf{s}),W^n_{\mathbf{i}}=\int_{\square_n\mathbf{i}/n}W(d\mathbf{s})\right\},$$

for $\mathbf{i} \in \{-n-\kappa, \dots, n+\kappa\}^2$ and $\mathbf{j} \in K_\kappa$, and

$$\mathcal{W}_n^2 := \left\{ W_{\mathbf{i}}^n = \int_{\square_n \mathbf{i}/n} W(d\mathbf{s}) \right\},$$

for $\mathbf{i} \in \{-N_n - n, \dots, N_n + n\}^2 \setminus \{-n-\kappa, \dots, n+\kappa\}^2$. Indeed, replacing $\mathbf{t}$ by $\mathbf{i}/n$ in (III.3.4) yields

$$\begin{aligned} X_{\mathbf{i}/n}^n &= \sum_{\mathbf{j} \in K_\kappa} L(\|\mathbf{b_j}\|)\sigma_{\frac{\mathbf{i}-\mathbf{j}}{n}} W_{\mathbf{i}-\mathbf{j},\mathbf{j}}^n + \sum_{\mathbf{j} \in \overline{K}_\kappa} g(\mathbf{b_j}/n)\sigma_{\frac{\mathbf{i}-\mathbf{j}}{n}} W_{\mathbf{i}-\mathbf{j}}^n \\ &:= \tilde{X}(\mathbf{i}/n) + \hat{X}(\mathbf{i}/n), \qquad \text{for } \mathbf{i} \in \{-n, \dots, n\}. \end{aligned}$$

By definition the random vectors $(W_{\mathbf{i},\mathbf{j}}^n, W_{\mathbf{i}}^n)$ are independent and identically distributed for varying $\mathbf{i}$. As a consequence, $\mathcal{W}_n^1$ and $\mathcal{W}_n^2$ are independent and $\mathcal{W}_n^2$ is composed of i.i.d. $\mathcal{N}(0, 1/n^2)$-distributed random variables. In order to simulate $\mathcal{W}_n^1$ we need to compute the covariance matrix of $(W_{0,\mathbf{j}}^n, W_0^n)_{\mathbf{j} \in K_\kappa}$, which is of size $(|K_\kappa| + 1)^2$ with $|K_\kappa| = (2\kappa + 1)^2$. In contrast to the purely temporal model considered in [7], computing the covariance structure becomes much more involved in our spatial setting. It relies partially on explicit expressions derived in appendix III.B, and partially on numeric integration.

Note that the complexity of computing $\tilde{X}(\frac{\mathbf{i}}{n})$ for all $\mathbf{i} \in \{-n, ..., n\}^2$ is $\mathcal{O}(n^2)$, as the number of summands does not increase with n. The sum $\hat{X}(\frac{\mathbf{i}}{n})$ can be written as the two dimensional discrete convolution of the matrices A and B defined by

$$A_{\mathbf{k}} := \begin{cases} 0 & \mathbf{k} \in K_\kappa \\ g(\mathbf{b_k}/n) & \mathbf{k} \in \overline{K}_\kappa \end{cases}, \qquad B_{\mathbf{k}} := \sigma_{\mathbf{k}/n} W_{\mathbf{k}}^n, \quad \text{for } \mathbf{k} \in \{-N-n, ..., N+n\}^2.$$

We remark that this expression as convolution is the main motivation that in (III.3.3) and (III.3.4) we chose to evaluate σ at the midpoints $\mathbf{t}-\mathbf{j}/n$ of $\square_n(\mathbf{t}-\mathbf{j}/n)$. Using FFT to carry out the convolution leads to a computational complexity of $\mathcal{O}(N^2 \log N) = \mathcal{O}(n^{2+2\gamma} \log n)$ for computing $\{\hat{X}(\frac{\mathbf{i}}{n})\}_{\mathbf{i} \in \{-n,...,n\}^2}$. Consequently, the computational complexity of the hybrid scheme is $\mathcal{O}(n^{2+2\gamma} \log n)$, provided the computational complexity of simulating $\{\sigma_{\mathbf{i}/n}\}_{\mathbf{i} \in \{-N-n,...,N+n\}^2}$ does not exceed $\mathcal{O}(n^{2+2\gamma} \log n)$. For a comparison we recall that the exact simulation of an isotropic Gaussian field using circulant embeddings is of complexity $\mathcal{O}(n^2 \log n)$, see [14].

Next we derive the asymptotics for the mean square error of the hybrid simulation scheme.

Theorem III.3.1. *Let $\alpha \in (-1, 0)$. Assume that σ is independent of W and satisfies* (III.2.2)*. If $\gamma > -(1+\alpha)/(1+\beta)$, we have for all $\mathbf{t} \in \mathbb{R}^2$ that*

$$n^{2(\alpha+1)} L(1/n)^{-2} \mathbb{E}[|X_{\mathbf{t}} - X_{\mathbf{t}}^n|^2] \to \mathbb{E}[\sigma_0^2] J(\alpha, \kappa, \mathbf{b}), \quad \text{as } n \to \infty.$$

Here the constant $J(\alpha, \kappa, \mathbf{b})$ is defined as

$$J(\alpha, \kappa, \mathbf{b}) = \sum_{\mathbf{j} \in \mathbb{Z}^2 \setminus \{-\kappa,...,\kappa\}^2} \int_{\square_{\mathbf{j}}} (\|\mathbf{x}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2 \, d\mathbf{x},$$

which is finite for $\alpha < 0$.

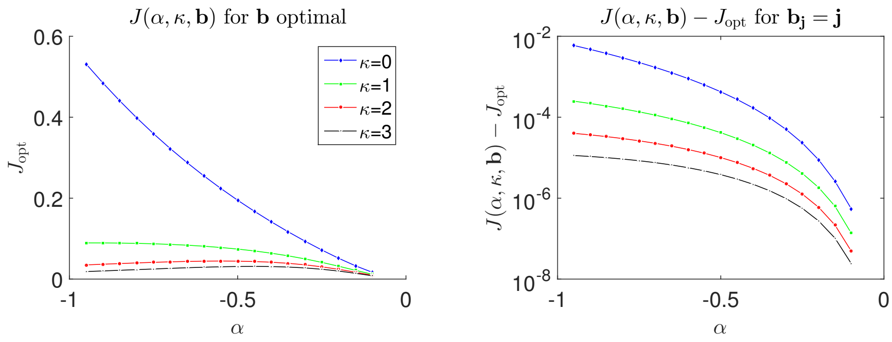

Figure III.2: The first figure shows the value of $J(\alpha, \kappa, \mathbf{b}) = J_{\text{opt}}$ for different values of α and κ for $\mathbf{b}$ chosen optimal, as in (III.3.5). The second figure shows the absolute error $J(\alpha, \kappa, \mathbf{b}) - J_{\text{opt}}$ for $\mathbf{b}$ chosen as midpoints, i.e. $\mathbf{b_j} = \mathbf{j}$, demonstrating that this choice leads to close to optimal results.

The proof is given in Section III.5. The sequence of evaluation points $\mathbf{b} = (\mathbf{b_j})_{\mathbf{j}\in\mathbb{Z}^2}$ can be chosen optimally, such that it minimises the limiting constant $J(\alpha, \kappa, \mathbf{b})$ and thus the asymptotic mean square error of the hybrid scheme. To this end $\mathbf{b_j}$ needs to be chosen in such a way that it minimises

$$\int_{\square_\mathbf{j}} (\|\mathbf{x}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2 d\mathbf{x},$$

for all $\mathbf{j} \in \mathbb{Z}^2$. By standard L^2 theory, $c \in \mathbb{R}$ minimises $\int_{\square_\mathbf{j}} (\|\mathbf{x}\|^\alpha - c)^2 d\mathbf{x}$ if and only if the function $\mathbf{x} \mapsto \|\mathbf{x}\|^\alpha - c$ is orthogonal to constant functions, that is, if it satisfies

$$\int_{\square_\mathbf{j}} (\|\mathbf{x}\|^\alpha - c) d\mathbf{x} = 0.$$

It follows then that $J(\alpha, \kappa, \mathbf{b})$ becomes minimal if we choose $\mathbf{b}$ such that

$$\|\mathbf{b_j}\| = \left(\int_{\square_\mathbf{j}} \|\mathbf{x}\|^\alpha d\mathbf{x} \right)^{1/\alpha}. \tag{III.3.5}$$

In Appendix III.B, we derive an explicit expression for this integral involving the Gauß hyperbolic function ${}_2F_1$. However, in our numerical experiments computing these integrals explicitly for all $\mathbf{j} \in \overline{K}_\kappa$ slowed the hybrid scheme down considerably, and we recommend choosing the midpoints $\mathbf{b_j} = \mathbf{j}$ instead. Figure III.2 shows the constant $J(\alpha, \kappa, \mathbf{b}) = J_{\text{opt}}$ for optimally chosen $\mathbf{b}$ and the error caused by choosing midpoints $\mathbf{b_j} = \mathbf{j}$ instead, giving evidence that choosing midpoints leads to a close to optimal result.

For $\mathbf{j} \in K_\kappa \setminus \{0\}$, the evaluation points $\mathbf{b_j}$ do not appear in the limiting expression in Theorem III.3.1, and we will simply choose $\mathbf{b_j} = \mathbf{j}$. However, for $\mathbf{j} = 0$ the expression $L(\|\mathbf{j}\|)$ is not necessarily defined. Indeed, the slowly varying function L might have a singularity at 0, which shows that particular attention should be paid to the choice of $\mathbf{b}_0$. The choice of $\mathbf{b}_0 \in \square_n \setminus \{0\}$ is optimal if it minimises the L^2 error of the central cell, i.e.,

$$\mathbf{b}_0 = \underset{\mathbf{b}\in\square_n\setminus\{0\}}{\arg\min}\ \mathbb{E}\bigg(\int_{\square_n} g(\mathbf{s})W(d\mathbf{s}) - L(\|\mathbf{b}\|)\int_{\square_n}\|\mathbf{s}\|^{\alpha}W(d\mathbf{s})\bigg)^2.$$

By straightforward calculation it can be shown that this is equivalent to

$$\begin{aligned} L(\|\mathbf{b}_0\|) &= \bigg(\int_{\square_n}\|\mathbf{s}\|^{2\alpha}L(\|\mathbf{s}\|)\,d\mathbf{s}\bigg)\bigg(\int_{\square_n}\|\mathbf{s}\|^{2\alpha}\,d\mathbf{s}\bigg)^{-1} \\ &= 8C_{0,0}^{-1}\int_0^{1/\sqrt{2}} r^{2\alpha+1}L(r/n)\big(\pi/4 - \arccos(\sqrt{2}r)\mathbb{1}_{\{r>1/2\}}\big)\,dr, \end{aligned}$$

where $C_{0,0}$ is defined in Appendix III.B. The integral on the right hand side is finite for $\alpha > -1$, which follows from the Potter bound (III.5.6), and can be evaluated numerically.

III.4 Numerical results

In this section we demonstrate in a simulation study that the hybrid scheme is capable of capturing the roughness of the process correctly, and compare it in that aspect to other simulation schemes. Before doing so, we present in Figure III.3 samples of VMMAs highlighting the effect of volatility. The volatility is modelled as $\sigma_{\mathbf{t}}^2 = \exp(X'_{\mathbf{t}})$, where X' is again a volatility modulated moving average, compare Example III.2.3. For X' we choose the roughness parameter $\alpha = -0.2$ and the slowly varying function $L(x) = e^{-x}$. For the first realisation we chose $\alpha = -0.3$ and $L(x) = e^{-x}$. For the second we chose $\alpha = -0.7$ and L such that the model has Matérn covariance, see Example III.2.2.

For our simulation study we first recall the definition of fractal or Hausdorff dimension. For a set $S \subset \mathbb{R}^d$ and $\varepsilon > 0$, an ε-cover of S is a countable collection of balls $\{B_i\}_{i\in\mathbb{N}}$ with diameter $|B_i| \le \varepsilon$ such that $S \subset \bigcup_i B_i$. The δ-dimensional Hausdorff measure of S is then defined as

$$H^{\delta}(S) = \lim_{\varepsilon\to 0}\inf\bigg\{\sum_{i=1}^{\infty}|B_i|^{\delta}\ :\ \{B_i\}_{i\in\mathbb{N}} \text{ is } \varepsilon\text{-cover of } S\bigg\},$$

and the fractal or Hausdorff dimension of S is $\mathrm{HD}(S) := \inf\{\delta > 0\ :\ H^{\delta}(S) = 0\}$. The Hausdorff dimension of a spatial stochastic process $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{R}^2}$ is the (random) Hausdorff dimension of its graph $\mathrm{HD}(\{(\mathbf{t}, X_{\mathbf{t}}),\ \mathbf{t}\in\mathbb{R}^2\})$, and takes consequently values in $[2,3]$. For the model (III.1.1) with constant volatility $\sigma \equiv 1$ it follows easily from a standard result [2, Theorem 8.4.1] and Theorem III.2.1 that $\mathrm{HD}(X) = 2 - \alpha$, see also [17]. In [13], the authors give an overview over existing methods for estimating the Hausdorff dimension of both time series data and spatial data, and provide implementations for various estimators in form of the `R` package `fractaldim`, which we rely on.

We estimate the Hausdorff dimension from simulations of X generated by the hybrid scheme, and compare to other simulation methods. We consider the model (III.1.1) with constant volatility σ and Matérn covariance, see example III.2.2. In

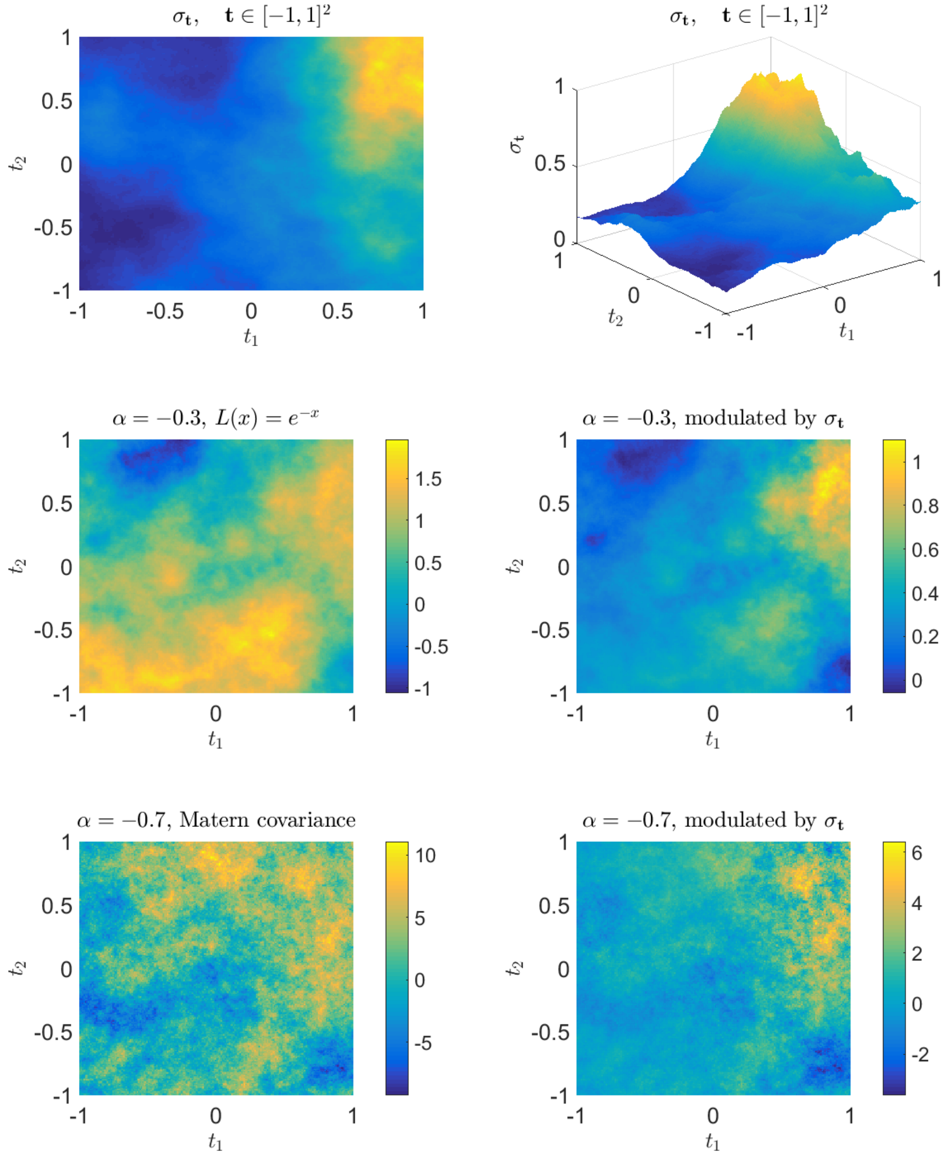

Figure III.3: Examples for moving average fields modulated by volatility. The first row shows the volatility $(\sigma_{\mathbf{t}})_{\mathbf{t}\in\mathbb{R}^2}$ modelled as $\sigma_{\mathbf{t}}^2 = \exp(X'_{\mathbf{t}})$, where X' is again a VMMA field. The second and third row show realisations of VMMAs. On the left hand side the field is simulated with constant volatility, the right hand side is generated by the same Gaussian noise and with the same model parameters, but is modulated by $(\sigma_{\mathbf{t}})_{\mathbf{t}\in\mathbb{R}^2}$. For the second row we chose $\alpha = -0.3$ and the slowly varying function $L(x) = e^{-x}$. The third row is generated with $\alpha = -0.7$ and Matérn covariance.

this case the process X can be simulated exactly using circulant embeddings of the covariance matrix. For this we use the `R` package `RandomFields`. For a discussion and many properties of the circulant embedding method in the context of simulating spatial Gaussian fields we refer to [14]. As this technique is restricted to Gaussian processes, and cannot be applied for general VMMAs, we compare additionally to the Riemann-sum scheme introduced in (III.3.3). These simulation techniques are compared to the hybrid scheme for $\kappa = 0, 1, 2$. With each technique we simulate 100 i.i.d. Monte-Carlo samples of the process $(X_{\mathbf{t}})_{\mathbf{t}\in[-1,1]^2}$ for every $\alpha \in \{-0.8, -0.7, ..., -0.1\}$. As grid resolution we chose $n = 100$ and, for the hybrid scheme and the Riemann-sum scheme, $N_n = [n^{1+\gamma}]$ with $\gamma = 0.3$, i.e. $N_n = 398$. Thereafter we estimate the roughness of X using the isotropic estimator $\hat{\nu}_I$ that was introduced in [12], see also [13], and average the estimates over the Monte-Carlo samples. Figure III.4 shows the results and compares them to the theoretical value of the Hausdorff dimension $2 - \alpha$, plotted as dashed line. Let us remark that there is a variety of methods to estimate roughness of spatial stochastic processes, a detailed comparison can be found in [13]. All estimators discussed there lead to similar results when applied to our simulations.

III.5 Proofs

This section is dedicated to the proofs of our theoretical results. We begin by recalling the Potter bound which follows from [9, Theorem 1.5.6]. For any $\delta > 0$ there exists a constant $C_\delta > 0$ such that

$$L(x)/L(y) \leq C_\delta \max\left\{\left(\frac{x}{y}\right)^{\delta}, \left(\frac{x}{y}\right)^{-\delta}\right\}, \quad x, y \in (0, 1]. \tag{III.5.6}$$

This bound will play an important role throughout all the proofs in this section.

Proof of Theorem III.2.1 (i). The proof is similar to the proof of [7, Proposition 2.1] We have for $h > 0$ by covariance stationarity of σ that

$$V(h) = \mathbb{E}[\sigma_0^2] \int_{\mathbb{R}^2} \big(g(\mathbf{s} + h\mathbf{e}) - g(\mathbf{s})\big)^2 \, d\mathbf{s},$$

where $\mathbf{e}$ is any unit vector and we used transformation into polar coordinates. We obtain

$$\begin{aligned} V(h) &= \mathbb{E}[\sigma_0^2](A_h + A'_h), \quad \text{where} \\ A_h &= \int_{\{\|\mathbf{s}\|\leq 1\}} \big(g(\mathbf{s} + h\mathbf{e}/2) - g(\mathbf{s} - h\mathbf{e}/2)\big)^2 \, d\mathbf{s}, \qquad \text{and} \\ A'_h &= \int_{\{\|\mathbf{s}\|>1\}} \big(g(\mathbf{s} + h\mathbf{e}/2) - g(\mathbf{s} - h\mathbf{e}/2)\big)^2 \, d\mathbf{s}. \end{aligned}$$

Since the function $\widetilde{g}$ is continuous differentiable on $(0, \infty)$, we obtain by mean value

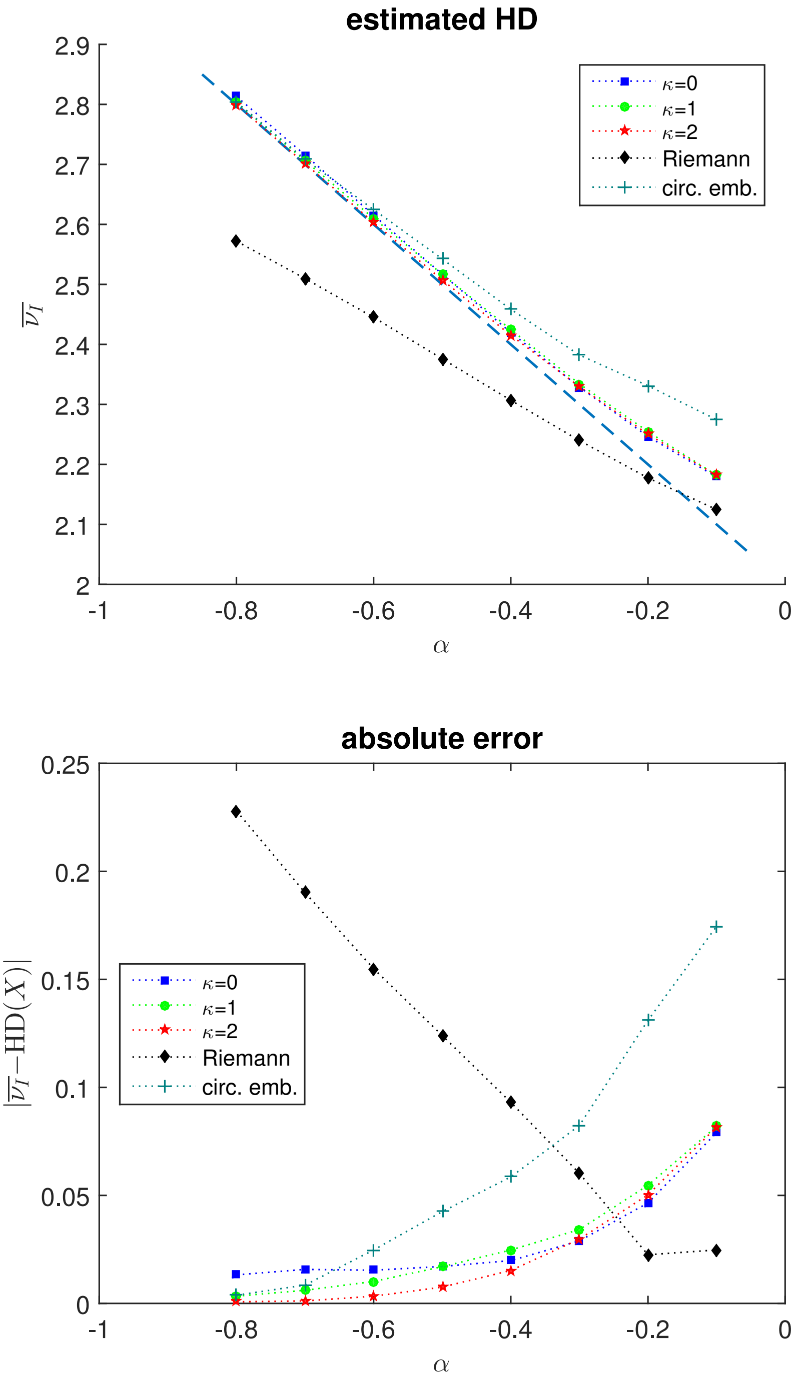

Figure III.4: Roughness estimated from samples generated by the hybrid scheme, the Riemann-sum approximation method and by exact simulation using the circulant embedding method for Gaussian fields. The roughness is estimated by the isotropic estimator ν_I introduced in [12], averaged over 100 i.i.d. samples. The second plot shows in more detail the absolute error between the estimation and the theoretical value, which is marked by the dashed line in the first plot.

theorem the following estimate for A_h'.

$$A_h' \le h^2\bigg\{ \int_{\{1<\|\mathbf{s}\|<M+1\}} \sup_{\{\xi\,:\,|\xi-\|\mathbf{s}\||\le h/\sqrt{2}\}} (\widetilde{g}'(\xi))^2 \; d\mathbf{s} + 2\pi \int_M^\infty \widetilde{g}'(r)^2 r \, dr \bigg\},$$

where we used that $|\widetilde{g}'|$ is decreasing on $[M,\infty)$. The term in curly brackets is finite by Assumption (A3), and we obtain that $A_h' = \mathcal{O}(h^2)$, as $h \to 0$. For A_h we make the substitution $\mathbf{x} = \mathbf{s}/h$ and obtain

$$\begin{aligned} A_h &= h^2 \int_{\|\mathbf{x}\|\le 1/h} \big(g(h(\mathbf{x}+\mathbf{e}/2)) - g(h(\mathbf{x}-\mathbf{e}/2))\big)^2 \, dbx \\ &= h^{2+2\alpha} L^2(h) \int_{\|\mathbf{x}\|\le 1/h} G_h(\mathbf{x}) \, d\mathbf{x}, \end{aligned}$$

where

$$G_h(\mathbf{x}) = \left(\|\mathbf{x}+\mathbf{e}/2\|^\alpha \frac{L(h\|\mathbf{x}+\mathbf{e}/2\|)}{L(h)} - \|\mathbf{x}-\mathbf{e}/2\|^\alpha \frac{L(h\|\mathbf{x}-\mathbf{e}/2\|)}{L(h)} \right)^2.$$

Note that $G_h(\mathbf{x}) \to \big(\|\mathbf{x}+\mathbf{e}/2\|^\alpha - \|\mathbf{x}-\mathbf{e}/2\|^\alpha\big)^2$, as $h \to 0$. Therefore the first statement of the theorem follows by the dominated convergence theorem if there is an integrable function G satisfying $G(\mathbf{x}) \ge |G_h(\mathbf{x})|$ for all $\mathbf{x}$ for sufficiently small h. The existence of such a function follows since L is bounded away from 0 on $(0,1]$ and by Assumption (A4). For details we refer to the proof of [7, Proposition 2.1]. $\square$

Proof of Theorem III.2.1 (ii). The proof relies on the Kolmogorov-Chentsov theorem (cf. [23, Theorem 3.23]), which requires localisation of the process, as σ does not necessarily have sufficiently high moments. We therefore first show the existence of a Hölder continuous version under the assumption that there is an $m > 0$ such that

$$|\sigma_\mathbf{s}|^2 \le m, \qquad \text{for all } \mathbf{s} \text{ with } \|\mathbf{s}\| \le M+1,\ \omega \in \Omega, \text{ and} \tag{III.5.7}$$

$$\int_{\{\|\mathbf{s}\|\ge M+1\}} (g(\mathbf{t}-\mathbf{s}) - g(-\mathbf{s}))^2 \sigma_\mathbf{s}^2 \, d\mathbf{s} \le m\|\mathbf{t}\|^2, \qquad \text{for all } \mathbf{t} \text{ with } \|\mathbf{t}\| \le 1,\ \omega \in \Omega, \tag{III.5.8}$$

where M is as in (A3). Thereafter we argue that the theorem remains valid if we relax these assumptions to $\mathbb{E}[\sup_{\|\mathbf{s}\|\le M} \sigma_\mathbf{s}^2] < \infty$.

For $\|\mathbf{t}\| \le 1$ we have for all $p > 0$ that

$$\begin{aligned} \mathbb{E}[(X_\mathbf{t} - X_0)^p] &\le C_p \mathbb{E}\bigg[\bigg(\int_{\mathbb{R}^2} \big(g(\mathbf{t}-\mathbf{s}) - g(-\mathbf{s})\big)^2 \sigma_\mathbf{s}^2 \, d\mathbf{s} \bigg)^{p/2} \bigg] \\ &\le C_p m^{p/2} \bigg(\int_{\{\|\mathbf{s}\|\le M+1\}} \big(g(\mathbf{t}-\mathbf{s}) - g(-\mathbf{s})\big)^2 \, d\mathbf{s} + \|\mathbf{t}\|^2 \bigg)^{p/2} \\ &\le C_p m^{p/2} \Big(V_0(\|\mathbf{t}\|) + \|\mathbf{t}\|^2 \Big)^{p/2}, \end{aligned}$$

where V_0 denotes the variogram of the process $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{R}^2}$ with $\sigma \equiv 1$. In the first inequality we used that σ and W are independent and therefore $X_{\mathbf{t}} - X_0$ has a Gaussian mixture distribution with the integral on the right hand side being the conditional variance. Applying the first part of the theorem and the Potter bound (III.5.6) we obtain that for any $\delta > 0$ a constant $C_{p,m,\delta}$ such that for all $\mathbf{t}$ with $\|\mathbf{t}\| \leq 1$

$$\mathbb{E}[(X_{\mathbf{t}} - X_0)^p] \leq C_{p,m,\delta}\|\mathbf{t}\|^{p+p\alpha-\delta}.$$

Therefore, the Kolmogorov-Chentsov Theorem [23, Theorem 2.23] implies that X has a continuous version that is Hölder continuous of any order $\gamma < 1 + \alpha - \frac{\delta}{p} - \frac{2}{p}$, and the result follows for any $\gamma \in (0, 1+\alpha)$ by letting $p \to \infty$.

We will now complete the proof of the theorem by extending it to processes not satisfying assumptions (III.5.7) and (III.5.8). By mean value theorem we obtain that for all $\mathbf{t}$ with $\|\mathbf{t}\| \leq 1$

$$\begin{aligned}
&\|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\|\geq M+1\}} (g(\mathbf{t}-\mathbf{s}) - g(-\mathbf{s}))^2 \sigma_{\mathbf{s}}^2 \, d\mathbf{s} \\
&\qquad \leq \|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\|\geq M+1\}} |\|\mathbf{t}-\mathbf{s}\| - \|\mathbf{s}\||^2 \sup_{r\in[\|\mathbf{s}\|,\|\mathbf{t}-\mathbf{s}\|]} \big(\widetilde{g}'(r)^2\big)\sigma_{\mathbf{s}}^2 \, d\mathbf{s} \\
&\qquad \leq \int_{\{\|\mathbf{s}\|\geq M+1\}} \widetilde{g}'(\|\mathbf{s}\| - 1)^2 \sigma_{\mathbf{s}}^2 \, d\mathbf{s}
\end{aligned}$$

where we used that $|\widetilde{g}'|$ is decreasing on $[M, \infty)$. By taking expectation and transformation into polar coordinates it follows from assumption (A3) that the right hand side is almost surely finite. Consequently, the random variable

$$Z := \max\left\{ \sup_{\|\mathbf{s}\|\leq M+1} (\sigma_{\mathbf{s}}^2), \sup_{\|\mathbf{t}\|\leq 1} \left(\|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\|\geq M+1\}} (g(\mathbf{t}-\mathbf{s}) - g(-\mathbf{s}))^2 \sigma_{\mathbf{s}}^2 \, d\mathbf{s} \right) \right\}$$

is almost surely finite. The process $(X_{\mathbf{t}}\mathbb{1}_{\{Z\leq m\}})_{\mathbf{t}\in\mathbb{R}^2}$ satisfies conditions (III.5.7) and (III.5.8) and coincides with X on $\{Z \leq m\}$. Therefore, the existence of a version of X with $\alpha + 1 - \varepsilon$-Hölder continuous sample paths follows by letting $m \to \infty$. $\square$

For the proof of Theorem III.3.1 we need the following auxiliary result. The proof is similar to the proof of [7, Lemma 4.2] and not repeated.

Lemma III.5.1. *Let $\alpha \in \mathbb{R}$ and $\mathbf{j} \in \mathbb{Z}^2 \setminus \{(0,0)\}$. If $\mathbf{b}_{\mathbf{j}} \in \square\mathbf{j}$, it holds that*

$$(i) \qquad \lim_{n\to\infty} \int_{\square\mathbf{j}} \left(\|\mathbf{x}\|^{\alpha} \frac{L(\|\mathbf{x}\|/n)}{L(1/n)} - \|\mathbf{b}_{\mathbf{j}}\| \frac{L(\|\mathbf{b}_{\mathbf{j}}\|/n)}{L(1/n)} \right)^2 d\mathbf{x} = \int_{\square\mathbf{j}} (\|\mathbf{x}\|^{\alpha} - \|\mathbf{b}\|^{\alpha})^2 \, d\mathbf{x},$$

$$(ii) \qquad \lim_{n\to\infty} \int_{\square\mathbf{j}} \|\mathbf{x}\|^{2\alpha} \left(\frac{L(\|\mathbf{x}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{b}_{\mathbf{j}}\|/n)}{L(1/n)} \right)^2 d\mathbf{x} = 0.$$

The same holds for $\mathbf{j} = (0,0)$ if $\mathbf{b}_{(0,0)} \neq (0,0)$ and $\alpha > -1$.

Proof of Theorem III.3.1. Recall the definition

$$\begin{aligned} X^n_{\mathbf{t}} := &\sum_{\mathbf{j}\in K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \|\mathbf{t}-\mathbf{s}\|^\alpha L(\|\mathbf{b_j}\|)\sigma_{\mathbf{t}-\mathbf{j}/n} W(d\mathbf{s}) \\ &+ \sum_{\mathbf{j}\in \overline{K}_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} g(\mathbf{b_j}/n)\sigma_{\mathbf{t}-\mathbf{j}/n} W(d\mathbf{s}). \end{aligned}$$

We introduce the auxiliary object X'^n defined as

$$\begin{aligned} X'^n_{\mathbf{t}} := &\sum_{\mathbf{j}\in K_\kappa \cup \overline{K}_\kappa} \sigma_{\mathbf{t}-\mathbf{j}/n} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} g(\mathbf{t}-\mathbf{s}) W(d\mathbf{s}) \\ &+ \int_{\mathbb{R}^2\setminus \square_{N_n/n}\mathbf{t}} g(\mathbf{t}-\mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}). \end{aligned}$$

Denoting $E_n := \mathbb{E}[|X^n_{\mathbf{t}} - X'^n_{\mathbf{t}}|^2]$ and $E'_n := \mathbb{E}[|X_{\mathbf{t}} - X'^n_{\mathbf{t}}|^2]$, Minkowski's inequality yields

$$E_n(1-\sqrt{E'_n/E_n})^2 \leq \mathbb{E}[|X^n_{\mathbf{t}} - X_{\mathbf{t}}|^2] \leq E_n(1+\sqrt{E'_n/E_n})^2. \tag{III.5.9}$$

We will show later that $E'_n/E_n \to 0$ as $n\to\infty$, and it is thus sufficient to analyse the asymptotic behavior of E_n.

We have that

$$\begin{aligned} E_n = &\sum_{\mathbf{j}\in K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \big(\|\mathbf{t}-\mathbf{s}\|^\alpha L(\|\mathbf{b_j}\|/n) - g(\mathbf{t}-\mathbf{s})\big)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]\, d\mathbf{s} \\ &+ \sum_{\mathbf{j}\in \{-n,\dots,n\}^2\setminus K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \big(g(\mathbf{t}-\mathbf{s}) - g(\mathbf{b_j}/n)\big)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]\, d\mathbf{s} \\ &+ \sum_{\mathbf{j}\in \overline{K}_\kappa\setminus \{-n,\dots,n\}^2} \int_{\square_n(\mathbf{t}-\mathbf{j})} \big(g(\mathbf{t}-\mathbf{s}) - g(\mathbf{b_j}/n)\big)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]\, d\mathbf{s} \\ &+ \int_{\mathbb{R}^2\setminus\square_{(2N_n+1)/n}\mathbf{t}} g(\mathbf{t}-\mathbf{s})^2 \mathbb{E}[\sigma^2_{\mathbf{s}}]\, d\mathbf{s} \\ =&\mathbb{E}[\sigma_0^2](D_1+D_2+D_3+D_4). \end{aligned} \tag{III.5.10}$$

For D_4 we obtain, recalling assumption (A2) and $N_n = n^{\gamma+1}$ that

$$D_4 \leq \int_{\|\mathbf{s}\|>N_n/n} g(\mathbf{s})^2 d\mathbf{s} = \mathcal{O}((N_n/n)^{2\beta+2}) = \mathcal{O}(n^{2\gamma(1+\beta)}).$$

Therefore, we have

$$n^{2(1+\alpha)} D_4 \to 0. \tag{III.5.11}$$

For D_3 we obtain

$$D_3 = \sum_{\mathbf{j}\in \overline{K}_\kappa\setminus \{-n,\dots,n\}^2} \int_{\square_n \mathbf{j}/n} \big(g(\mathbf{s}) - g(\mathbf{b_j}/n)\big)^2\, d\mathbf{s}.$$

Recalling the notation $\tilde{g}(\|\mathbf{s}\|) = g(\mathbf{s})$ we have for $\mathbf{s} \in \Box\mathbf{j}$ with $\mathbf{j} \in \overline{K_\kappa} \setminus \{-n, \dots, n\}^2$ by the mean value theorem $\xi \in [\|\mathbf{s}\| \wedge \|\mathbf{b_j}/n\|, \|\mathbf{s}\| \vee \|\mathbf{b_j}/n\|]$. Since $\tilde{g}'$ is decreasing on $[M, \infty)$ by assumption (A3) it follows that

$$\begin{aligned}|g(\mathbf{s}) - g(\mathbf{b_j}/n)| &= |\tilde{g}'(\xi)(\|\mathbf{s}\| - \|\mathbf{b_j}\|/n)| \\ &\leq \begin{cases} \frac{1}{n} \sup_{y \in [1-1/(\sqrt{2}n), M+1/(\sqrt{2}n)]} |\tilde{g}'(y)|, & (\|\mathbf{j}\| - \sqrt{2})/n < M, \\ \frac{1}{n} |\tilde{g}'((\|\mathbf{j}\| - \sqrt{2})/n)|, & (\|\mathbf{j}\| - \sqrt{2})/n \geq M. \end{cases}\end{aligned}$$

Consequently, we obtain with transformation into polar coordinates

$$\limsup_{n\to\infty} n^2 D_3 \leq \big(\pi(M+1)^2 \sup_{z\in[1/2, M+1/2]} |\tilde{g}'(z)| + C \int_M^\infty r|\tilde{g}'(r)|^2\, dr\big) < \infty. \quad \text{(III.5.12)}$$

For D_1 we have that

$$\begin{aligned} D_1 &= \frac{1}{n^2} \sum_{\mathbf{j}\in K_\kappa} \int_{\Box\mathbf{j}} \big(\|\mathbf{s}/n\|^\alpha L(\|\mathbf{b_j}\|/n) - g(\mathbf{s}/n)\big)^2\, d\mathbf{s} \\ &= \frac{L(1/n)}{n^{2+2\alpha}} \sum_{\mathbf{j}\in K_\kappa} \int_{\Box\mathbf{j}} \|\mathbf{s}\|^{2\alpha} \bigg(\frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{s}\|/n)}{L(1/n)}\bigg)^2\, d\mathbf{s}. \end{aligned}$$

Since the number of elements of K_κ does not depend on n, we have by Lemma III.5.1

$$\lim_{n\to\infty} \frac{n^{2+2\alpha} D_1}{L(1/n)} = 0. \quad \text{(III.5.13)}$$

For the asymptotic of D_2 it holds that

$$\begin{aligned} D_2 &= \frac{1}{n^2} \sum_{\mathbf{j}\in\{-n,\dots,n\}^2\setminus K_\kappa} \int_{\Box\mathbf{j}} \big(g(\mathbf{s}/n) - g(\mathbf{b_j}/n)\big)^2\, d\mathbf{s} \\ &= \frac{L(1/n)^2}{n^{2+2\alpha}} \sum_{\mathbf{j}\in\{-n,\dots,n\}^2\setminus K_\kappa} \underbrace{\int_{\Box\mathbf{j}} \bigg(\|\mathbf{s}\|^\alpha \frac{L(\|\mathbf{s}\|/n)}{L(1/n)} - \|\mathbf{b_j}\|^\alpha \frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)}\bigg)^2 d\mathbf{s}}_{:=A_{\mathbf{j},n}}. \end{aligned}$$

From Lemma III.5.1 we know that $\lim_{n\to\infty} A_{\mathbf{j},n} = \int_{\Box\mathbf{j}} (\|\mathbf{s}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2\, d\mathbf{s}$. Consequently, if we find a dominating sequence $A_\mathbf{j}$ such that $A_\mathbf{j} \geq A_{\mathbf{j},n}$ for all n and $\sum_{\mathbf{j}\in\mathbb{Z}^2\setminus K_\kappa} A_\mathbf{j} < \infty$, it follows from dominated convergence theorem that

$$\lim_{n\to\infty} \frac{D_2 n^{2\alpha+2}}{L(1/n)^2} = \sum_{\mathbf{j}\in\mathbb{Z}^2\setminus K_\kappa} \int_{\Box\mathbf{j}} (\|\mathbf{s}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2\, d\mathbf{s}, \quad \text{for } \alpha \in (-1, 0). \quad \text{(III.5.14)}$$

It holds that

$$\begin{aligned} A_{\mathbf{j},n} &= \int_{\Box\mathbf{j}} \bigg\{ (\|\mathbf{s}\|^\alpha - \|\mathbf{b_j}\|^\alpha) \frac{L(\|\mathbf{s}\|/n)}{L(1/n)} + \|\mathbf{b_j}\|^\alpha \bigg(\frac{L(\|\mathbf{s}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)}\bigg) \bigg\}^2 d\mathbf{s} \\ &\leq 2 \int_{\Box\mathbf{j}} (\|\mathbf{s}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2 \bigg(\frac{L(\|\mathbf{s}\|/n)}{L(1/n)}\bigg)^2 d\mathbf{s} \\ &\quad + 2 \int_{\Box\mathbf{j}} \|\mathbf{b_j}\|^{2\alpha} \bigg(\frac{L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|/n)}{L(1/n)}\bigg)^2 d\mathbf{s} \\ &:= I_{\mathbf{j},n} + I'_{\mathbf{j},n}. \end{aligned}$$

For $I'_{\mathbf{j},n}$ we note that $\|\mathbf{b_j}\|^{2\alpha} \le (\|\mathbf{j}\| - 1/\sqrt{2})^{2\alpha}$ for $\alpha < 0$. By the mean value theorem we have a $\xi \in [\|\mathbf{s}\|/n \wedge \|\mathbf{b_j}\|/n, \|\mathbf{s}\|/n \vee \|\mathbf{b_j}\|/n]$ such that

$$|L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|/n)| = L'(\xi)|\|\mathbf{s}\|/n - \|\mathbf{b_j}\|/n| \le \frac{C}{n} + \frac{C}{\|\mathbf{j}\| - 1/\sqrt{2}} \le \frac{2C}{\|\mathbf{j}\| - 1/\sqrt{2}},$$

where we used (A4) and that $\|\mathbf{j}\| \le n$. Consequently, we obtain

$$\begin{aligned} I'_{\mathbf{j},n} &\le \frac{C}{\inf_{x\in(0,1]} L(x)} (\|\mathbf{j}\| - 1/\sqrt{2})^{2\alpha} \int_{\square \mathbf{j}} (L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|))^2 \, d\mathbf{s} \\ &\le C(\|\mathbf{j}\| - 1/\sqrt{2})^{2(\alpha-1)}. \end{aligned}$$

For the term $I_{\mathbf{j},n}$ we obtain by the Potter bound and the mean value theorem that

$$I_{\mathbf{j},n} \le C_\delta \int_{\square \mathbf{j}} \min(\|\mathbf{s}\|, b_{\mathbf{j}})^{2\alpha-2} \|\mathbf{s}\|^{2\delta} \, d\mathbf{s} \le C_\delta (\|\mathbf{j}\| - 1/\sqrt{2})^{2(\alpha-1+\delta)},$$

where we choose $\delta \in (0, -\alpha)$. Consequently, we obtain $I_{\mathbf{j},n} + I'_{\mathbf{j},n} \le C(\|\mathbf{j}\| - 1/\sqrt{2})^{-2}$ for all $n > 0$, and since

$$\sum_{\mathbf{j}\in\mathbb{Z}^2 \setminus K_\kappa} C(\|\mathbf{j}\| - 1/\sqrt{2})^{-2} < \infty,$$

(III.5.14) follows from dominated convergence theorem and Lemma III.5.1. Now (III.5.10) together with (III.5.11), (III.5.12), (III.5.13) and (III.5.14) show that

$$E_n \sim \mathbb{E}[\sigma_0^2] J(\alpha, \kappa, \mathbf{b}) n^{-2(\alpha+1)} L(1/n)^2, \quad n \to \infty.$$

Therefore, recalling (III.5.9), the proof of statement (i) of the Theorem can be completed by showing that $E'_n/E_n \to 0$ as $n \to \infty$.

Since σ is covariance stationary, we obtain for E'_n

$$\begin{aligned} E'_n &= \sum_{\mathbf{j}\in K_\kappa \cup \overline{K}_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \mathbb{E}[(\sigma_{\mathbf{t}-\mathbf{j}/n} - \sigma_{\mathbf{s}})^2] g(\mathbf{t}-\mathbf{s})^2 d\mathbf{s} \\ &= \sup_{\mathbf{u}\in\square_n} \mathbb{E}[|\sigma_{\mathbf{u}} - \sigma_0|^2] \int_{\mathbb{R}^2} g(\mathbf{s})^2 d\mathbf{s}, \end{aligned}$$

and $E'_n/E_n \to 0$ follows by the assumption (III.2.2) $\square$

Appendix III.A On general stochastic integrals

Here we recall the definition of general stochastic integrals of the form $\int_{\mathbb{R}^2} H_{\mathbf{s}} W(d\mathbf{s})$ where H is a real valued stochastic process, not necessarily independent of W. The construction of such integrals dates back to Bichteler [8]. In a recent publication [10], this theory is revisited in a spatio-temporal setting and the authors derive a general integrability criterion for stochastic integrals driven by a random measure that is easy to check. In the context of integrals of the form (III.1.1), this criterion yields the following statement.

Proposition III.A.1. *Let $(H_{\mathbf{s}})_{\mathbf{s}\in\mathbb{R}^2}$ be a real valued stochastic process, measurable with respect to $\mathcal{B}(\mathbb{R}^2)\otimes\mathcal{F}$, such that $H\in L^2(\mathbb{R}^2)$, almost surely. Then, the stochastic integral $\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})$ exists in the sense of [8].*

Proof. We apply the integrability criterion [10, Theorem 4.1] that is formulated in a spatio-temporal framework. To this end, we introduce an artificial time component and lift the white noise $W(d\mathbf{s})$ to a space time white noise $\widetilde{W}(dt;d\mathbf{s})$ such that $W(A)=\widetilde{W}([0,1]\times A)$ for all $A\in\mathcal{B}(\mathbb{R}^2)$. Equipping $(\Omega,\mathcal{F},\mathbb{P})$ with the maximal filtration $\mathcal{F}_t=\mathcal{F}$ for all $t\in[0,1]$, the spatio-temporal process defined as $H_{\mathbf{s}}(t):=H_{\mathbf{s}}$ for all $t\in[0,1]$ is predictable and it holds that

$$\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})=\int_{[0,1]\times\mathbb{R}^2} H_{\mathbf{s}}(t)\widetilde{W}(dt;d\mathbf{s})$$

if the latter exists. The random measure $\widetilde{W}$ satisfies the conditions of [10, Theorem 4.1] with characteristics $B=\mu=\nu=0$ and $C(A;B)=\lambda(A\cap B)$ for all $A,B\in\mathcal{B}([0,1]\times\mathbb{R}^2)$, where λ denotes the Lebesgue measure. The theorem then implies that H is integrable with respect to W if and only if it satisfies almost surely $\int_{\mathbb{R}^2} H_{\mathbf{s}}^2 d\mathbf{s}<\infty$. □

Note that the proofs for some of our theoretical results rely on the isometry

$$\mathbb{E}\Big[\Big(\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})\Big)^2\Big]=\mathbb{E}\Big[\int_{\mathbb{R}^2} H_{\mathbf{s}}^2 d\mathbf{s}\Big],$$

which does not necessarily hold when H and W are dependent. In particular, we cannot rely on Theorem III.3.1 in this more general framework. We argue next that the hybrid scheme converges for dependent σ and W, when σ admits a continuous version, without specifying the speed of convergence.

Proposition III.A.2. *Assume that $(\sigma_{\mathbf{s}})_{\mathbf{s}\in\mathbb{R}^2}$ has a continuous version. Then, $X_{\mathbf{t}}^n\xrightarrow{\mathbb{P}}X_{\mathbf{t}}$ for all $\mathbf{t}\in\mathbb{R}^2$, i.e. the hybrid scheme converges.*

Proof. Using the notation of Section III.3, we consider the auxiliary integrals

$$\widetilde{X}_{\mathbf{t}}^n:=\sum_{\mathbf{k}\in K_\kappa\cup\overline{K}_\kappa}\sigma_{\mathbf{t}-\mathbf{k}/n}\int_{\square_n(\mathbf{t}-\mathbf{k}/n)} g(\mathbf{t}-\mathbf{s})W(d\mathbf{s})=\int_{\mathbb{R}^2}\widetilde{\sigma}_{\mathbf{s}}^n g(\mathbf{t}-\mathbf{s})W(d\mathbf{s}),$$

where

$$\widetilde{\sigma}_{\mathbf{s}}^n:=\sum_{\mathbf{k}\in K_\kappa\cup\overline{K}_\kappa}\sigma_{\mathbf{t}-\mathbf{k}/n}\mathbb{1}_{\square_n(\mathbf{t}-\mathbf{k}/n)}(\mathbf{s}).$$

By arguing as in the proof of Theorem III.3.1, it follows that $\mathbb{E}[(\widetilde{X}_{\mathbf{t}}^n-X_{\mathbf{t}}^n)^2]\to 0$ as $n\to\infty$, and it is therefore sufficient to argue that $\widetilde{X}_{\mathbf{t}}^n\xrightarrow{\mathbb{P}}X_{\mathbf{t}}$. It holds that

$$X_{\mathbf{t}}=\int_{\mathbb{R}^2} g(\mathbf{t}-\mathbf{s})\sigma_{\mathbf{s}}W(d\mathbf{s})=\int_{\mathbb{R}^2}\sigma_{\mathbf{s}}M_{g,\mathbf{t}}(d\mathbf{s}),$$

where the random measure $M_{g,\mathbf{t}}$ is defined as $M_{g,\mathbf{t}}(A) = \int_A g(\mathbf{t}-\mathbf{s})W(d\mathbf{s})$. Since $(\sigma_\mathbf{s})_{\mathbf{s}\in\mathbb{R}^2}$ is continuous, the sequence of simple integrands $\widetilde{\sigma}^n$ converges pointwise to σ, and it follows that

$$X_\mathbf{t} = \int_{\mathbb{R}^2} \sigma_\mathbf{s} M_{g,\mathbf{t}}(d\mathbf{s}) = \lim_{n\to\infty} \int_{\mathbb{R}^2} \widetilde{\sigma}_\mathbf{s}^n M_{g,\mathbf{t}}(d\mathbf{s}) = \lim_{n\to\infty} \widetilde{X}_\mathbf{t}^n, \quad \text{in probability,}$$

by integrability of σ with respect to $M_{g,\mathbf{t}}$. □

Appendix III.B The covariance of $\mathcal{W}_n^1$

In this section we analyse the covariance structure of the Gaussian family $\mathcal{W}_n^1$ introduced in Section III.3. For a wide range of covariances we are able to derive closed expressions, whereas the remaining covariances are computed by numeric integration. Let us remark that in addition to the symmetry of the covariance matrix the isotropy of the process adds 8 more spatial symmetries (corresponding to the linear transformations in $SO(2)$ that map $\mathbb{Z}^2$ onto itself), which reduces the number of necessary computations drastically. Since the random variables in $\mathcal{W}_n^1$ are i.i.d. along $\mathbf{i}$, it is sufficient to derive the covariance matrix for

$$\left\{W_{0,\mathbf{j}}^n,\ W_0^n\right\}_{\mathbf{j}\in K_\kappa}.$$

For $\mathbf{j}_1, \mathbf{j}_2 \in \{-\kappa,\dots,\kappa\}^2$ it holds that

$$\begin{aligned}
C_{1,1} &:= \operatorname{var}(W_0^n) = \frac{1}{n^2}\\
C_{1,\mathbf{j}_1} &:= \operatorname{cov}(W_0^n, W_{0,\mathbf{j}_1}^n) = \frac{1}{n^{2+\alpha}} \int_\Box \|\mathbf{j}_1 - \mathbf{s}\|^\alpha \, d\mathbf{s}\\
C_{\mathbf{j}_1,\mathbf{j}_2} &:= \operatorname{cov}(W_{0,\mathbf{j}_1}^n, W_{0,\mathbf{j}_2}^n) = \frac{1}{n^{2+2\alpha}} \int_\Box \|\mathbf{j}_1 - \mathbf{s}\|^\alpha \|\mathbf{j}_2 - \mathbf{s}\|^\alpha \, d\mathbf{s}.
\end{aligned}$$

We now derive explicit expressions for $C_{\mathbf{j},\mathbf{j}}$ using the Gauss hypergeometric function ${}_2F_1$. Clearly, these expressions can be applied to compute $C_{1,\mathbf{j}}$ by replacing α with $\alpha/2$. Using symmetries we may assume without loss of generality that $\mathbf{j} = (j_1, j_2)$ with $j_1 \geq j_2 \geq 0$. We introduce the notation $\lhd\mathbf{j}$ for the area $\{(x_1,x_2) : j_2 \leq x_1 \leq j_1,\ j_2 \leq x_2 \leq x_1\}$, that is a right triangle with lower right point (j_1, j_2) and hypotenuse lying on the diagonal $\{(x_1,x_2) : x_1 = x_2\}$. In order to obtain explicit expressions for $C_{\mathbf{j},\mathbf{j}}$, we first derive explicit expressions for

$$\int_{\lhd\mathbf{j}} \|\mathbf{x}\|^{2\alpha} \, d\mathbf{x}, \quad \text{for all } \mathbf{j} = (j_1,j_2) \in \mathbb{R}^2, 0 \leq j_2 < j_1. \tag{III.B.15}$$

Thereafter we give for all $\mathbf{j} = (j_1, j_2) \in \mathbb{Z}^2$ with $0 \leq j_2 \leq j_2$ an explicit formula to write $C_{\mathbf{j},\mathbf{j}}$ as linear combination of such integrals.

Transforming into polar coordinates we obtain that

$$\begin{aligned}
\int_{\lhd\mathbf{j}} \|\mathbf{x}\|^{2\alpha} \, d\mathbf{x} &= \int_{\arctan(j_2/j_1)}^{\pi/4} \int_{j_2/\sin(\theta)}^{j_1/\cos(\theta)} r^{2\alpha+1} \, dr \, d\theta\\
&= \frac{1}{2\alpha+2} \int_{\arctan(j_2/j_1)}^{\pi/4} \left(\frac{j_1}{\cos(\theta)}\right)^{2\alpha+2} - \left(\frac{j_2}{\sin(\theta)}\right)^{2\alpha+2} d\theta.
\end{aligned} \tag{III.B.16}$$

It holds that $\arctan(j_2/j_1) = \arccos(\frac{j_1}{\|\mathbf{j}\|})$, and consequently we obtain by substituting $\cos(\theta) = z$ the following expression for the first summand:

$$\begin{aligned}
&\frac{j_1^{2\alpha+2}}{2\alpha+2}\int_{\arctan(j_2/j_1)}^{\pi/4} \cos(\theta)^{-2\alpha-2}\, d\theta \\
&\qquad = -\frac{j_1^{2\alpha+2}}{2\alpha+2}\int_{j_1/\|\mathbf{j}\|}^{\cos(\pi/4)} z^{-2\alpha-2}(1-z^2)^{-1/2}\, dz \\
&\qquad = \frac{j_1^{2\alpha+2}}{4(\alpha+1)}\int_{1/2}^{j_1^2/\|\mathbf{j}\|^2} z^{-\alpha-\frac{3}{2}}(1-z)^{-1/2}\, dz \\
&\qquad = \frac{j_1^{2\alpha+2}}{4(\alpha+1)}\int_{j_2^2/\|\mathbf{j}\|^2}^{1/2} (1-z)^{-\alpha-\frac{3}{2}} z^{-1/2}\, dz \\
&\qquad = \frac{j_1^{2\alpha+2}}{4(\alpha+1)}(B(1/2;1/2,-\alpha-1/2) - B(j_2^2/\|\mathbf{j}\|^2;1/2,-\alpha-1/2)) \\
&\qquad = \frac{j_1^{2\alpha+2}}{2^{3/2}(\alpha+1)}\,{}_2F_1(1/2,3/2+\alpha;3/2;1/2) \\
&\qquad\quad - \frac{j_1^{2\alpha+2}j_2}{2\|\mathbf{j}\|(\alpha+1)}\,{}_2F_1(1/2,3/2+\alpha;3/2;j_2^2/\|\mathbf{j}\|^2).
\end{aligned}$$

Here, $B(x;p,q)$ denotes the incomplete beta function, satisfying $B(x;p,q) = \frac{x^p}{p}\,{}_2F_1(p,1-q;p+1;x)$. For the first equality we used that $d/dz(\arccos(z)) = -(1-z^2)^{-1/2}$ For the second summand in (III.B.16) we argue similarly, using that $\arctan(j_2/j_1) = \arcsin(\frac{j_2}{\|\mathbf{j}\|})$,

$$\begin{aligned}
&-\frac{j_2^{2\alpha+2}}{2\alpha+2}\int_{\arctan(j_2/j_1)}^{\pi/4} \sin(\theta)^{-2\alpha-2}\, d\theta \\
&\qquad = -\frac{j_2^{2\alpha+2}}{2\alpha+2}\int_{j_2/\|\mathbf{j}\|}^{\sin(\pi/4)} z^{-2\alpha-2}(1-z^2)^{-1/2}\, dz \\
&\qquad = -\frac{j_2^{2\alpha+2}}{4(\alpha+1)}\int_{j_2^2/\|\mathbf{j}\|^2}^{1/2} z^{-\alpha-\frac{3}{2}}(1-z)^{-1/2}\, dz \\
&\qquad = -\frac{j_2^{2\alpha+2}}{4(\alpha+1)}\int_{1/2}^{j_1^2/\|\mathbf{j}\|^2} (1-z)^{-\alpha-\frac{3}{2}} z^{-1/2}\, dz \\
&\qquad = -\frac{j_2^{2\alpha+2}}{4(\alpha+1)}(B(j_1^2/\|\mathbf{j}\|^2;1/2,-\alpha-1/2) - B(1/2;1/2,-\alpha-1/2)) \\
&\qquad = \frac{j_2^{2\alpha+2}}{2^{3/2}(\alpha+1)}\,{}_2F_1(1/2,3/2+\alpha;3/2;1/2) \\
&\qquad\quad - \frac{j_2^{2\alpha+2}j_1}{2\|\mathbf{j}\|(\alpha+1)}\,{}_2F_1(1/2,3/2+\alpha;3/2;j_1^2/\|\mathbf{j}\|^2).
\end{aligned}$$

This leads to

$$\begin{aligned}\int_{\lhd\,\mathbf{j}} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} = {} & \frac{j_2^{2\alpha+2} + j_1^{2\alpha+2}}{2^{3/2}(\alpha+1)}\, {}_2F_1(1/2, 3/2+\alpha; 3/2; 1/2) \\ & - \frac{j_1 j_2^{2\alpha+2}}{2\|\mathbf{j}\|(\alpha+1)}\, {}_2F_1(1/2, 3/2+\alpha; 3/2; j_1^2/\|\mathbf{j}\|^2) \\ & - \frac{j_1^{2\alpha+2} j_2}{2\|\mathbf{j}\|(\alpha+1)}\, {}_2F_1(1/2, 3/2+\alpha; 3/2; j_2^2/\|\mathbf{j}\|^2),\end{aligned}$$

for all $0 \le j_2 < j_1$. For implementation we remark that in the case $j_2 = 0$ the hypergeometric function in the second line is not defined since in this case $j_1^2/\|\mathbf{j}\|^2 = 1$, and we use

$$\int_{\lhd\,(j_1,0)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} = \frac{\sqrt{2} j_1^{2\alpha+2}}{4(\alpha+1)}\, {}_2F_1(1/2, 3/2+\alpha; 3/2; 1/2).$$

Thus, we have explicit expressions for integrals of the form (III.B.15) and all that remains to do is to argue that for $0 \le j_2 < j_1$ we can write $C_{\mathbf{j},\mathbf{j}}$ as linear combinations of such integrals. By symmetry we obtain that

$$C_{(0,0),(0,0)} = \int_{\Box} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} = 8 \int_{\lhd\,(1/2,0)'} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x}.$$

For $j > 0$ we obtain

$$\begin{aligned} C_{(j,j),(j,j)} = {} & 2\int_{\lhd\,(j+1/2,j-1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x}, \quad \text{and} \\ C_{(j,0),(j,0)} = {} & 2\Bigg(\int_{\lhd\,(j+1/2,0)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} - \int_{\lhd\,(j-1/2,0)'} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} \\ & - \int_{\lhd\,(j+1/2,1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} + \int_{\lhd\,(j-1/2,1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x}\Bigg). \end{aligned}$$

For $0 < j_2 < j_1$ we obtain

$$\begin{aligned} C_{(j_1,j_2),(j_1,j_2)} = {} & \int_{\lhd\,(j_1+1/2,j_2-1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} - \int_{\lhd\,(j_1-1/2,j_2-1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} \\ & - \int_{\lhd\,(j_1+1/2,j_2+1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x} + \int_{\lhd\,(j_1-1/2,j_2+1/2)} \|\mathbf{x}\|^{2\alpha}\, d\mathbf{x}. \end{aligned}$$

This covers all possible choices for $0 \le j_2 < j_1$ and consequently we have explicit expressions for $C_{\mathbf{j},\mathbf{j}}$ and $C_{\mathbf{j},1}$ for all $\mathbf{j}$.

Bibliography

[1] Abramowitz, M. and I. Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* U.S. Government Printing Office, Washington, D.C.

[2] Adler, R. J. (1981). *The geometry of random fields.* John Wiley & Sons, Ltd., Chichester. Wiley Series in Probability and Mathematical Statistics.

[3] Barndorff-Nielsen, O., F. Benth, and A. Veraart (2011). Modelling electricity forward markets by ambit fields. *Available at https://ssrn.com/abstract=1938704*.

[4] Barndorff-Nielsen, O., E. Hedevang, J. Schmiegel, and B. Szozda (2016). Some recent developments in ambit stochastics. In *Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015*, Volume 138, pp. 3–25. Springer, Cham.

[5] Barndorff-Nielsen, O. and J. Schmiegel (2007). Ambit processes; with applications to turbulence and tumour growth. In *Stochastic analysis and applications*, pp. 93–124. Springer.

[6] Barndorff-Nielsen, O. and J. Schmiegel (2008). Time change, volatility, and turbulence. In *Mathematical control theory and finance*, pp. 29–53. Springer, Berlin.

[7] Bennedsen, M., A. Lunde, and M. Pakkanen (2016). Hybrid scheme for brownian semistationary processes. available at arXiv:1507.03004.

[8] Bichteler, K. (2002). *Stochastic integration with jumps.* Cambridge University Press, Cambridge.

[9] Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989). *Regular variation.* Cambridge University Press, Cambridge.

[10] Chong, C. and C. Klüppelberg (2015). Integrability conditions for space-time stochastic integrals: theory and applications. *Bernoulli 21*(4), 2190–2216.

[11] Cohen, S., C. Lacaux, and M. Ledoux (2008). A general framework for simulation of fractional fields. *Stochastic Process. Appl. 118*(9), 1489–1517.

[12] Davies, S. and P. Hall (1999). Fractal analysis of surface roughness by using spatial data. *J. R. Stat. Soc. Ser. B Stat. Methodol. 61*(1), 3–37.

[13] Gneiting, T., H. Ševčíková, and D. Percival (2012). Estimators of fractal dimension: assessing the roughness of time series and spatial data. *Statist. Sci. 27*(2), 247–277.

[14] Gneiting, T., H. Ševčíková, D. Percival, M. Schlather, and Y. Jiang (2006). Fast and exact simulation of large Gaussian lattice systems in $\mathbf{R}^2$: exploring the limits. *J. Comput. Graph. Statist. 15*(3), 483–501.

[15] Goff, J. and T. Jordan (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research: Solid Earth 93*(B11), 13589–13608.

[16] Guttorp, P. and T. Gneiting (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika 93*(4), 989–995.

[17] Hansen, L. and T. Thorarinsdottir (2013). A note on moving average models for Gaussian random fields. *Statist. Probab. Lett. 83*(3), 850–855.

[18] Hansen, L., T. Thorarinsdottir, E. Ovcharov, T. Gneiting, and D. Richards (2015). Gaussian random particles with flexible Hausdorff dimension. *Adv. in Appl. Probab. 47*(2), 307–327.

[19] Hellmund, G., M. Prokešová, and E. Jensen (2008). Lévy-based Cox point processes. *Adv. in Appl. Probab. 40*(3), 603–629.

[20] Huang, W., K. Wang, F. Breidt, and R. Davis (2011). A class of stochastic volatility models for environmental applications. *J. Time Series Anal. 32*(4), 364–377.

[21] Jensen, E., K. Jonsdottir, J. Schmiegel, and O. Barndorff-Nielsen (2006). Spatio-temporal modelling-with a view to biological growth. *Monographs on statistics and applied probability 107*, 47.

[22] Jónsdóttir, K., A. Rønn-Nielsen, K. Mouridsen, and E. Jensen (2013). Lévy-based modelling in brain imaging. *Scand. J. Stat. 40*(3), 511–529.

[23] Kallenberg, O. (2002). *Foundations of modern probability* (Second ed.). Springer-Verlag, New York Berlin Heidelberg.

[24] Kwapień, S. and W. Woyczyński (1992). *Random series and stochastic integrals: single and multiple*. Birkhäuser Boston, Inc., Boston, MA.

[25] Matérn, B. (1986). *Spatial variation* (Second ed.). Springer-Verlag, Berlin.

[26] Nguyen, M. and A. Veraart (2017). Modelling spatial heteroskedasticity by volatility modulated moving averages. *Spatial Statistics*.

[27] Rajput, B. and J. Rosiński (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields 82*(3), 451–487.

[28] Rosiński, J. (1990). On series representations of infinitely divisible random vectors. *Ann. Probab. 18*(1), 405–430.

[29] Yan, J. (2007). Spatial stochastic volatility for lattice data. *J. Agric. Biol. Environ. Stat. 12*(1), 25–40.

Appendices

Appendix A

Technical supplement

In this appendix we provide some results that are of more technical nature, including several proofs for results stated in Chapter 1. Literature references are with respect to the bibliography of Chapter 1.

A.1 The Skorokhod topologies

In his original work [79], Skorokhod introduced four different topologies on the space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of cádlág functions, usually denoted by J_1, M_1, J_2 and M_2.

In this section we argue that the functional stable convergence in Theorem I.1.1 (i) does hold with respect to the M_1 and M_2 topology, but not with respect to the J_1 and J_2 topology, providing a complete picture.

We quickly recall the definition and some properties of the J_1, J_2 and M_2 topology, since especially the J_2 and M_2 topology are not widely used. An excellent analysis of the four Skorokhod topologies in the context of stochastic processes can be found in the monograph [87]. For simplicity we restrict ourselves in this section to the finite time horizon $t_\infty = 1$ and work on the space $\mathbb{D} = \mathbb{D}([0,1); \mathbb{R})$, since including the endpoint of the interval requires more technical notation to account for functions that jump at the endpoint. First of all we recall that the J_1 topology on $\mathbb{D}$ is induced by the metric

$$d_{J_1}(f_1, f_2) = \inf_{\lambda \in \Lambda} \{\|f_1 \circ \lambda - f_2\| \vee \|\lambda - e\|\},$$

where $\|\cdot\|$ denotes the uniform norm on $[0,1)$, e is the identity on $[0,1)$ and Λ is the set of all strictly increasing continuous bijections $[0,1) \to [0,1)$.

Similarly, the J_2 topology is induced by the metric

$$d_{J_2}(f_1, f_2) = \inf_{\lambda \in \Lambda'} \{\|f_1 \circ \lambda - f_2\| \vee \|\lambda - e\|\},$$

where Λ' is the class of all bijections of $[0,1)$, not requiring that they are increasing and continuous. As a consequence, a single jump can in the J_2 topology approximated by a function that jumps multiple times up and down near the jump, see Figure A.1.

The M_2 metric is defined as the Hausdorff distance of the completed graphs, introduced in Section 1.2, i.e. $d_{M_2}(f_1, f_2) = d_{\mathrm{HD}}(\Gamma_{f_1}, \Gamma_{f_2})$, where we recall that the

Hausdorff distance between compact sets A, B of $\mathbb{R}^2$ is defined as

$$d_{\mathrm{HD}}(A,B) := \sup_{x\in A}\inf_{y\in B}\|x-y\| \vee \sup_{x\in B}\inf_{y\in A}\|x-y\|,$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^2$. To gain some intuition for the M_2-topology we follow up on our picture of ants walking on the completed graphs that we gave for the M_1-metric in Section 1.2. Given two functions f and g we imagine two ants positioned at the starting points $(0, f(0))$ and $(0, g(0))$ in $\mathbb{R}^2$. For $\varepsilon > 0$ it holds $d_{M_2}(f_1, f_2) < \varepsilon$ if the ants can find any way to walk the completed graphs Γ_f and Γ_g from start to finish without ever being further apart than ε. The crucial difference to the M_1-distance is that they are now allowed to change directions in between. In Figure A.1 we show examples for convergence in J_2 and M_2. Indeed we see in the figure that the ant walking on the green graph needs to walk back and forth on the vertical line in order to stay close to the ant walking the blue line, which can walk forward the entire time.

The four Skorokhod topologies are ordered by

$$J_1 > J_2 > M_2 \quad \text{and} \quad J_1 > M_1 > M_2,$$

where $>$ means stronger than. The J_2 and M_1 topology are not comparable. The first example in Figure A.1 converges in J_2 but not in M_1, the second example in Figure 1.2 converges in M_1 but not in J_2.

The main result of this section is the following theorem.

Theorem A.1.1. *In the setting of Section I.4, the sequence $V^{n,\varepsilon}$ defined in* (I.4.3) *does not converge stably in law in $\mathbb{D}$ equipped with the J_2 topology.*

By the ordering of the 4 topologies this implies that the sequence neither converges in J_1, whereas M_2 convergence follows from Theorem I.1.1 (i).

We remark that this result is in fact quite intuitive, as by the definition of $V^{n,\varepsilon}$ a jump of the limiting process Z at time T is indeed approximated by jumps of $V^{n,\varepsilon}$ at times $([nT]+1)/n, ..., [n(T+\varepsilon)]/n$ forming a monotone staircase. This type of monotone staircase scenario is a popular model example for convergence in M_1 but not J_2. The formal proof of Theorem A.1.1 relies on the following Lemma.

Lemma A.1.2. *Let $m \geq 1$ and let*

$$A := \{f \in \mathbb{D} \ : \ f \textit{ is piecewise constant and has at most } m \textit{ jumps}\}$$

Let $g \in \mathbb{D}$ be increasing with at least $m+1$ jumps of size greater or equal $\delta > 0$. Then $d_{J_2}(f, g) \geq \delta/2$ for all $f \in A$.

Proof. The function g attains $m+1$ values $g_1, ..., g_{m+1}$ satisfying $\min_{i,j\in\{1,...,m+1\}}\{|g_i - g_j|\} > \delta$. For any $\lambda \in \Lambda'$, $g \circ \lambda$ attains the same values, and since f attains at most m different values, we have $\|g\circ\lambda - f\| \geq \delta/2$ for all $\lambda \in \Lambda'$ and the result follows. $\square$

For the proof of Theorem A.1.1 we recall the definition of the Prokhorov metric defined on the space $\mathcal{P}(S)$ of probability measures on a metric space (S, d). Denote

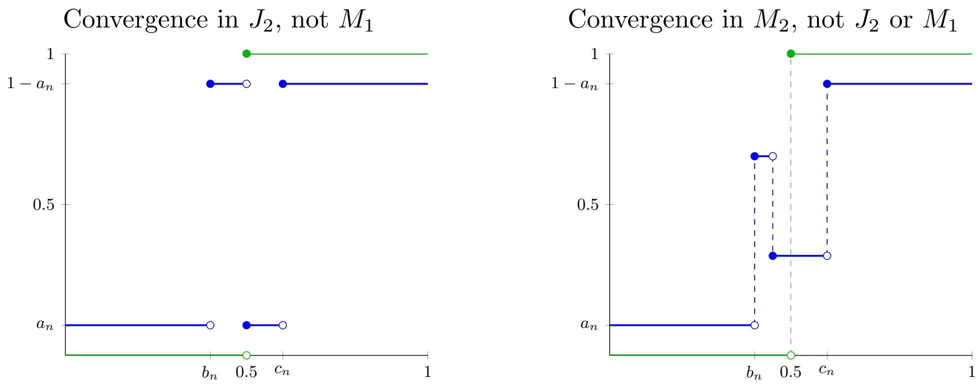

Figure A.1: Examples for convergence in J_2 and M_2. If $a_n \to 0$ and $b_n, c_n \to 0.5$, the function plotted in blue converges to the function $\mathbb{1}_{[1/2,1)}$ plotted in green. For J_2-convergence the approximating function is allowed to jump multiple times up and down when the limiting function jumps. For M_2-convergence the completed graphs (plotted as dashed lines) converge in the Hausdorff metric.

by A_d^ε the open ε-neighbourhood of A, i.e.

$$A_d^\varepsilon = \{y \in S \,:\, d(x,y) < \varepsilon \text{ for some } x \in A\}.$$

The Prohorov distance of two probability measures $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}(S)$ is then defined as

$$\pi_d(\mathbb{P}_1, \mathbb{P}_2) = \inf\{\varepsilon > 0 \,:\, \mathbb{P}_1(A) \le \mathbb{P}_2(A_d^\varepsilon) + \varepsilon \text{ for all } A \in \mathcal{B}(S)\}.$$

It has been shown in [87, Theorem 3.2.1] that weak convergence is equivalent to convergence in the Prohorov metric whenever the metric space (S,d) is separable.

Proof of Theorem A.1.1. Since it is sufficient to show that the convergence does not hold in a specific example, we can choose $\sigma \equiv 1$, $g(x) = x_+^\alpha$ and $k = 1$ and $\nu = \delta_{\{-1\}} + \delta_{\{1\}}$. Recall the definition of $\Omega_\varepsilon, V^{n,\varepsilon}$ and Z introduced in Subsection I.4. For the proof we introduce the set

$$\Omega_0 = \{\omega \,:\, L(\omega) \text{ has exactly one jump in } (0,1)\} \cap \Omega_\varepsilon.$$

It is then sufficient to show that $V^{n,\varepsilon}\mathbb{1}_{\Omega_0}$ does not converge stably in law. Moreover, the results of Subsection I.4 imply $V^{n,\varepsilon}\mathbb{1}_{\Omega_0} \xrightarrow{\mathcal{L}_{M_1}-s} Z\mathbb{1}_{\Omega_0}$ and it is sufficient to show that $V^{n,\varepsilon}\mathbb{1}_{\Omega_0} \xrightarrow{\mathcal{L}-s} Z\mathbb{1}_{\Omega_0}$ does not hold in $(\mathbb{D}, d_{J_2})$. Since $(\mathbb{D}, d_{J_1})$ is Polish and $d_{J_2} \le d_{J_1}$, the metric space $(\mathbb{D}, d_{J_2})$ is separable, and by Theorem 3.2.1 of [87] the proof is complete if we find a $\delta > 0$ such that $\pi_{J_2}(V^{n,\varepsilon}\mathbb{1}_{\Omega_0}, Z\mathbb{1}_{\Omega_0}) > \delta$ for infinitely many n. Let

$$\begin{aligned} A := \{f \in \mathbb{D} \,:\, & f \text{ is piecewise constant}, f(0) = 0, \\ & f \text{ has exactly 1 jump of absolute size} \ge \alpha^p\} \end{aligned}$$

Let $\omega \in \Omega_0$ and denote by $T_1(\omega)$ the jump time of the Lévy process in $(0,1)$. The process $Z\mathbb{1}_{\Omega_0}$ is piecewise constant, 0 at 0, increasing, and has one jump at time $T_1(\omega)$

of size V_1 defined in Theorem I.1.1. It is straightforward to show that $V_1 \geq \alpha^p$. In particular, we can choose δ sufficiently small such that $\mathbb{P}(Z\mathbb{1}_{\Omega_0} \in A) = \mathbb{P}(\Omega_0) > 2\delta$. It is then sufficient to show that

$$\mathbb{P}(V^{n,\varepsilon}\mathbb{1}_{\Omega_0} \in A^{\delta}_{d_{J_2}}) \leq \delta, \quad \text{for infinitely many } n,$$

since this implies $\pi_{J_2}(V^{n,\varepsilon}\mathbb{1}_{\Omega_0}, Z^{\varepsilon}\mathbb{1}_{\Omega_0}) \geq \delta$. We choose $\delta < \alpha^p 3^{p(\alpha-1)}/2$ and show that the set $\{\omega : V^{n,\varepsilon}(\omega)\mathbb{1}_{\Omega_0}(\omega) \in A^{\delta}_{d_{J_2}}\}$ is in fact empty for all n. First note that $\delta < \alpha^p 3^{p(\alpha-1)}/2$ implies $\delta < \alpha^p$ and thus $0 \notin A^{\delta}_{d_{J_2}}$. It is therefore sufficient to show that for all $\omega \in \Omega_0$ and all n it holds that

$$d_{J_2}(f, V^{n,\varepsilon}(\omega)) > \delta, \quad \text{for all } f \in A.$$

We denote by i_1 the index such that $T_1 \in ((i_1-1)/n, i_1/n]$. It is straight forward to show that

$$\begin{aligned}
&\Delta_{\frac{i_1+1}{n}} V^{n,\varepsilon} = |(1+i_1-nT_1)^{\alpha} - (i_1-nT_1)^{\alpha}|^p \geq \alpha^p 2^{p(\alpha-1)}, \\
&\Delta_{\frac{i_1+2}{n}} V^{n,\varepsilon} \geq \alpha^p 3^{p(\alpha-1)}.
\end{aligned}$$

Therefore $V^{n,\varepsilon}$ is increasing and has at least 2 jumps $\geq \alpha^p 3^{p(\alpha-1)}$. Since $f \in A$ has only one jump, an application of Lemma A.1.2 with $m = 1$ shows that $d_{J_2}(V^{n,\varepsilon}, f) \geq \alpha^p 3^{p(\alpha-1)}/2 > \delta$ for all $f \in A$. This completes the proof. □

A.2 Details on modulars

In this subsection we provide proofs and supplementary details to Section 1.3. We begin by showing the following proposition.

Proposition A.2.1. *For all $p \geq 1$, the modular $\Phi_{p,L}$ introduced in Section 1.3 is both of moderate growth and 0-convex.*

Proof. We first derive the following estimate for the function ϕ_p. For all $x \in \mathbb{R}$ it holds that

$$(\lambda^2 \wedge \lambda^p)\phi_p(x) \leq \phi_p(\lambda x) \leq (\lambda^2 \vee \lambda^p)\phi_p(x) \quad \text{for all } \lambda \geq 0. \tag{A.2.1}$$

We show the second inequality, the first one follows directly by an application of the second one with $x' = \lambda x$ and $\lambda' = \lambda^{-1}$. Assume w.l.o.g. that $x \geq 0$. For $x \in [0, 1 \wedge \lambda^{-1}]$ we have $\phi_p(\lambda x) = \lambda^2 \phi_p(x)$. For $\lambda > 1$ and $x \in (\lambda^{-1}, 1]$ it holds that

$$\phi_p(\lambda x) = \lambda^p x^p \begin{cases} \leq \lambda^p x^2 = \lambda^p \phi(x) & \text{if } p \geq 2 \\ = \lambda^2 (\lambda x)^{p-2} x^2 \leq \lambda^2 \phi(x) & \text{if } p \leq 2. \end{cases}$$

Similarly, for $\lambda < 1$ and $x \in [1, \lambda^{-1})$

$$\phi_p(\lambda x) = \lambda^2 x^2 \begin{cases} \leq \lambda^2 x^p = \lambda^2 \phi(x) & \text{if } p \geq 2 \\ = \lambda^p (\lambda x)^{2-p} x^p \leq \lambda^p \phi(x) & \text{if } p \leq 2. \end{cases}$$

Finally, when $x \geq \lambda^{-1} \vee 1$ it holds that $\phi_p(\lambda x) = \lambda^p \phi_p(x)$, which completes the proof of (A.2.1).

This estimate implies in particular that the modular $\Phi_{p,L}$ is of moderate growth, i.e. satisfies condition ((iii)) of Definition 1.3.1. Indeed, for $x, y \in \mathbb{R}$ we obtain by (A.2.1) that $\phi_p(x+y) \leq \phi_p(2(|x| \vee |y|)) \leq 2^{2\vee p}\phi_p(|x| \vee |y|) \leq 2^{2\vee p}(\phi_p(x) + \phi_p(y))$, which immediately implies

$$\Phi_{p,L}(f+g) \leq 2^{p\vee 2}(\Phi_p(f) + \Phi_p(g)) \quad \text{for all } f, g \in \mathbf{L}^p_{\mathrm{nr}}(dL).$$

Moreover, $\Phi_{p,L}$ is 0-convex, since $\phi_p(\alpha x + \beta y) \leq \phi_p(|x| \vee |y|) \leq \phi_p(x) + \phi_p(y)$ for all $x, y \in \mathbb{R}$ and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$. □

Next, we prove Theorem 1.3.2 and Proposition 1.3.4.

Proof of Theorem 1.3.2. (i) follows immediately from [71, Theorem 2.7] and [71, Theorem 3.3] where we remark that the quantity $\sigma(s)$ introduced in [71, (2.4)] vanishes since L has no Brownian part, and the functional U introduced in [71, Theorem 2.7] vanishes since L is symmetric.

Next we prove (ii). By [71, Theorem 3.4] the integration mapping $\Lambda : \mathbf{L}^p_{\mathrm{nr}}(dL) \to L^p(\Omega)$ is a linear homeomorphism onto its image. We recall that any linear homeomorphism between normed vector spaces is a quasi isometry which follows by considering the operator norm. Since in our case the vector spaces are not equipped with a norm but with homogeneous modulars, we need to generalise this idea slightly in the following way. Denote by B_1 the unit ball in $\mathbf{L}^p_{\mathrm{nr}}(dL)$, which coincides for $\Phi_{p,L}, \|\cdot\|_{p,L}$ and $|\!|\!|\cdot|\!|\!|_{p,L}$, since for any f the function $t \mapsto \Phi_{p,L}(tf)$ is strictly increasing on $[0,\infty)$. The set B_1 is bounded in the linear metric space $(\mathbf{L}^p_{\mathrm{nr}}(dL), |\!|\!|\cdot|\!|\!|_{p,L})$ and consequently its image under the continuous linear operator Λ is bounded as well by a standard result, see for example [73]. Thus we can define

$$\|\Lambda\|_{\mathrm{op}} := \sup_{f \in B_1} \|\Lambda(f)\|_p < \infty,$$

which defines a norm if $p \geq 1$, and a homogeneous modular for $p < 1$. It follows then by homogeneity of $\|\cdot\|_{p,L}$ and $\|\cdot\|_p$ that

$$\|\Lambda(f)\|_p = \|f\|_{p,L} \big\|\Lambda(f/\|f\|_{p,L})\big\|_p \leq \|\Lambda\|_{\mathrm{op}} \|f\|_{p,L}.$$

The same argument applied on the inverse mapping Λ^{-1} yields $\|f\|_{p,L} \leq C\|\Lambda(f)\|_p$. □

Proof of Proposition 1.3.4. It follows from [64, Theorem 1.10] and the comment thereafter that the homogeneous modular $\|\cdot\|_\Phi$ defines a norm if Φ is convex. We replace ϕ_p in (1.5) by the convex function

$$\widetilde{\phi}_p(x) := (2/p|x|^p + 1 - 2/p)\mathbb{1}_{\{|x|>1\}} + x^2 \mathbb{1}_{\{|x| \leq 1\}},$$

and define $\widetilde{\Phi}_{p,L}$ accordingly. Then, the convexity of $\widetilde{\Phi}_{p,L}$ implies that $\|\cdot\|'_{p,L} := \|\cdot\|_{\widetilde{\Phi}_{p,L}}$ defines a norm on $\mathbf{L}^p_{\mathrm{nr}}(dL)$. We show that it is equivalent to $\|\cdot\|_{p,L}$.

Clearly, there exist $c, C > 0$ such that $c\widetilde{\phi}_p(x) \leq \phi_p(x) \leq C\widetilde{\phi}_p(x)$ for all $x \in \mathbb{R}$, which implies $c\widetilde{\Phi}_{p,L}(x) \leq \Phi_{p,L}(x) \leq C\widetilde{\Phi}_{p,L}(x)$. Moreover, the modular $\Phi_{p,L}$ satisfies the moderate growth condition (iii) which implies the existence of a $C' > 0$ such that $c^{-1}\Phi_{p,L}(f) \leq \Phi_{p,L}(C'f)$ for all f. It follows for all $f \in \mathbf{L}^p_{\mathrm{nr}}(dL)$ that

$$\begin{aligned}
\|f\|'_{p,L} &= \inf\{t > 0 \,:\, \widetilde{\Phi}_{p,L}(t^{-1}f) \leq 1\} \\
&\leq \inf\{t > 0 \,:\, c^{-1}\,\Phi_{p,L}(t^{-1}f) \leq 1\} \\
&\leq \inf\{t > 0 \,:\, \Phi_{p,L}(C't^{-1}f) \leq 1\} \\
&= \|C'f\|_{p,L} = C'\|f\|_{p,L}.
\end{aligned}$$

Similarly it follows that there is a c' such that $c'\|f\|_{p,L} \leq \|f\|'_{p,L}$. This shows the equivalence of $\|\cdot\|_{p,L}$ and $\|\cdot\|'_{p,L}$.

The modular $\|\cdot\|_{p,L}$ satisfies ((i)) by definition. Property ((ii)) follows from equivalence to $\|\cdot\|'_{L,p}$. The estimate ((iii)) follows from (A.2.1). $\square$

A.3 Existence of Lévy semi-stationary processes

In this section we discuss conditions that ensure the existence of the integral (1.2).

We first argue that assumption (A) implies the following important estimate. For all $\varepsilon > 0$ there is a constant $C > 0$ such that

$$\int_{\mathbb{R}} (|yx|^2 \wedge 1)\nu(dx) \leq C(|y|^\theta \mathbb{1}_{\{|y|\leq 1\}} + |y|^{\beta+\varepsilon}\mathbb{1}_{\{|y|>1\}}). \tag{A.3.2}$$

Recall that the condition $\limsup_{t\to\infty} \nu([t,\infty))t^\theta < \infty$ implies that there is a C, such that for all nondecreasing functions f

$$\int_1^\infty f(x)\,\nu(dx) \leq C\int_1^\infty f(x)x^{-\theta-1}\,dx, \tag{A.3.3}$$

as we argued in Lemma I.4.4. First, consider the case $y > 1$. Choosing $\varepsilon > 0$ such that $\beta + \varepsilon \leq 2$, it holds that

$$\begin{aligned}
\int_{\mathbb{R}} (|yx|^2 \wedge 1)\nu(dx) &= 2\bigg(\int_0^{y^{-1}} (yx)^2\nu(dx) + \nu([y^{-1},\infty))\bigg) \\
&\leq 2\bigg(y^{\beta+\varepsilon}\int_0^1 x^{\beta+\varepsilon}\nu(dx) + \nu([1,\infty))\bigg) \\
&\leq Cy^{\beta+\varepsilon}.
\end{aligned}$$

For $|y| \leq 1$ we have

$$\begin{aligned}
\int_{\mathbb{R}} (|yx|^2 \wedge 1)\nu(dx) &\leq C\bigg(y^2 + \int_1^\infty ((yx)^2 \wedge 1)\nu(dx)\bigg) \\
&\leq C\bigg(y^2 + \int_1^\infty ((yx)^2 \wedge 1)x^{-\theta-1}dx\bigg) \\
&= Cy^2\bigg(1 + \int_1^{|y^{-1}|} x^{1-\theta}dx\bigg) + C\int_{|y^{-1}|}^\infty x^{-\theta-1}dx \\
&\leq C|y|^\theta,
\end{aligned}$$

where we used (A.3.3) in the second inequality, and $\theta \leq 2$ in the last. This shows the estimate (A.3.2).

This estimate ensures the existence of the integral (1.2) if $\sigma = 1$ by the following argument, see also [20]. By Theorem 1.3.2 it is sufficient to argue that for all $t \geq 0$ the function f_t defined as $f_t(s) = g(t-s) - g_0(-s)$ satisfies $\Phi_{0,L}(f_t) < \infty$, where $\Phi_{0,L}$ is as in Section 1.3. Applying the estimate (A.3.2) it holds that

$$\Phi_{0,L}(f_t) \leq C \int_{\mathbb{R}} |f_t(s)|^\theta \mathbb{1}_{\{|f_t(s)| \leq 1\}} + |f_t(s)|^{\beta+\varepsilon} \mathbb{1}_{\{|f_t(s)|>1\}} ds.$$

By the mean value theorem and the conditions on g, there is a $\xi_t \in [0,t]$ such that $|f_t(s)| \leq |f_0(s)| + |g'(\xi_t - s)| \mathbb{1}_{\{s>t+\delta\}} + C_t(t-s)_+^\alpha \mathbb{1}_{\{s \leq t+\delta\}}$, and the conditions of Assumption (A) ensure that $\Phi_{0,L}(f_t)$ is finite.

When the volatility factor σ is nontrivial, the following condition is sufficient for the integral (1.2) to exist, see also Remark (I.3.3). This follows easily from Theorem 1.3.3 and the estimate (A.3.2).

Assumption (B): Suppose that Assumption (A) is satisfied and define the two processes $F^{(1)}$ and $F^{(2)}$ by $F_s^{(1)} = (g(-s) - g_0(-s))\sigma_s$ and $F_s^{(2)} = g'(-s)\sigma_s$ for $s < 0$. Then the process X given by (1.2) is well-defined if there exists a $\beta' > \beta$ such that

$$\int_{-\infty}^{-\delta} \left(|F_s^{(i)}|^\theta \mathbb{1}_{\{|F_s^{(i)}| \leq 1\}} + |F_s^{(i)}|^{\beta'} \mathbb{1}_{\{|F_s^{(i)}|>1\}} \right) ds < \infty$$

almost surely for $i = 1, 2$.

Appendix B

MATLAB code for the hybrid scheme

Here we list the MATLAB code for simulating volatility modulated moving averages by the hybrid scheme presented in Paper III. The comments of the code follow the notation of the paper, denoting bold letters (i.e. variables representing vectors in $\mathbb{R}^2$) by `\b`, for example we denote $\mathbf{i}$ by `\bi`.

The code is written in MATLAB R2014b, version 8.4.0.150421. For fast 2d-convolution it relies on the function `conv2fft` by Luigi Rosa, available at `http://se.mathworks.com/matlabcentral/fileexchange/4334`

```

%%%%% The Hybrid scheme 2d %%%%%

% Claudio Heinrich, August 2016

clear all;
close all;

%% Simulates and plots a VMMA over [-1,1]^2 with grid
%  resultion 1/n. The values of the process are saved
%  in the (2n+1) x (2n+1) matrix X

kappa   =    2;              %depth of the Hybrid scheme
a       = -0.3;              %roughness parameter alpha
n       =  100;              %grid resolution is 1/n
g       =  0.2;              %parameter gamma
N       = floor(n^(1+g));    %the integral range is N/n


%% The volatility factor sigma

% the function vol returns the volatility process

sigma=vol(n,N);
```

```
%% The matrix containing the evaluation points \|b_k\|

bMat= bMatSimple(N);


%% Auxiliary objects:
%   LgMat contains the values L(\|\bk\|/n) for \bk\in
%   {-kappa,...,kappa}^2, and the values g(\bk/n) for
%   \bk\in {-N,...,N}^2 \setminus \{-kappa,...,kappa\}^2.
%   Choose 'LgMatMatern' for Matern covariance and
%   'LgMatexponential' for the slowly varying function
%   L(x)=exp(-x)


LgMat    =LgMatMatern(n,N,a,kappa,bMat);
%LgMat   =LgMatExponential(n,N,a,kappa,bMat);


%% Simulate Gaussian RVs

C=Cov3(kappa,a,n);  % returns the covariance matrix

W0=mvnrnd(zeros((2*kappa+1)^2+1,1),C,(2*n+2*kappa+1)^2).';


% The following array stores the random variables
% W^n_{\bi} for \bi in {-n-kappa,...,n+kappa}^2:

W01=reshape(W0(1,:),[2*n+2*kappa+1,2*n+2*kappa+1]);


% The following array supplements W0 and contains
% the random variables W^n_{\bi} for all \bi in
% {-N-n,...,N+n}^2:

We1=normrnd(0,1/n^2,[2*n+2*N+1,2*n+2*N+1]);
We1(N-kappa+2:N+2*n+kappa+2,N-kappa+2:N+2*n+kappa+2)=W01;


% The following array stores the random variables
% W^n_{\bi,\bj} for \bi in {-n-kappa,...,n+kappa}^2,
% \bj in {-kappa,...,kappa}^2:

W02=reshape(W0(2:end,:),[2*kappa+1,2*kappa+1,2*n+2*kappa+1,2*n+2*
    kappa+1]);
```

```
% The following auxiliary 4d array contains the same
% data as W02 as We2:

We2=zeros(2*kappa+1,2*kappa+1,2*n+2*N+1,2*n+2*N+1);
We2(:,:,N-kappa+2:N+2*n+kappa+2,N-kappa+2:N+2*n+kappa+2)=W02;




%% Simulation of \tilde X, i.e. of the integral around 0

% Wshift contains sigma_{\bi-\bk}W_{\bi-\bk,\bk}
% at position (k1+kappa+1,k2+kappa+1,i1+n+1,i2+n+1),

 Wshift=nan(2*kappa+1,2*kappa+1,2*n+1,2*n+1);
 for k1=-kappa:kappa
     for k2=-kappa:kappa
         for i1=-n:n
             for i2=-n:n
                 Wshift(k1+kappa+1,k2+kappa+1,i1+n+1,i2+n+1)=sigma(i1-k1+N+n+1,i2-k2+N+n+1)*We2(k1+kappa+1,k2+kappa+1,i1-k1+N+n+1,i2-k2+N+n+1);
             end
         end
     end
end



X1=nan(2*n+1);  %temporary, stores values of \tilde X

for i1= -n:n
    for i2 = -n:n
        B=LgMat(N-kappa+1:N+kappa+1,N-kappa+1:N+kappa+1).*Wshift(:,:,i1+n+1,i2+n+1);
        X1(i1+n+1,i2+n+1)= sum(B(:));
    end
end



%% Simulation of \hat X, that is the integral away from 0

% gMat contains the values g(\bk/n) for
% \bk\in\{-N,...,N\}^2\setminus\{-kappa,...,kappa\}^2,
% and 0 at the positions corresponding to \{-kappa,...,kappa\}^2

gMat=LgMat;
```

```
gMat(N+1-kappa:N+1+kappa,N+1-kappa:N+1+kappa)=zeros(2*kappa+1);

X2=conv2fft(sigma.*We1,gMat,'valid'); %stores \hat X


%% plotting

X=X1+X2;
surf(-1:1/n:1,-1:1/n:1,X,'EdgeColor','none');

set(gca,'FontSize',12)
xlabel('$t_1$','Interpreter','latex')
ylabel('$t_2$','Interpreter','latex')
zlabel('$X_{\bf{t}}$','Interpreter','latex')
title(['$\alpha=-$' num2str(a) ],'interpreter','latex','FontSize',14)
```

B.1 Auxiliary functions for the hybrid scheme

In this section we list all functions (and subfunctions) called by the hybrid scheme in alphabetical order.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [ b ] = bMatSimple( N )

% Contains the evaluation points \bb_{\bj}
% for \bj in {-N,...,N}^2.

b=zeros(2*N+1);

for i=0:N
    for j=0:i
        b(i+N+1,j+N+1)=norm([i,j]);
    end
end

b(N+1,N+1)=0; %b_(0,0)

for i=0:N-1
    for j=i+1:N
        b(i+N+1,j+N+1)=b(j+N+1,i+N+1);
    end
end

for i=-N:-1
    for j=-N:-1
        b(i+N+1,j+N+1)=b(-i+N+1,-j+N+1);
```

```
    end
    for j=0:N
        b(i+N+1,j+N+1)=b(-i+N+1,j+N+1);
    end
end

for i=0:N
    for j=-N:-1
        b(i+N+1,j+N+1)=b(i+N+1,-j+N+1);
    end
end

end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function C1 = Cov1(kappa,a,n)

% returns the covariances C_{1,\bj}.
% The output matrix C1 is defined as
% C1(j,k)=C_{1,(j-kappa-1,k-kappa-1)}
% See Appendix III.B for details.

TriMa=TriIntMat(kappa,a/2);

C1=nan(2*kappa+1,2*kappa+1);

% C_{(0,0),(0,0)}

C1(kappa+1,kappa+1)=8*TriMa(1,1);


% C_{(j,j),(j,j)}, for j > 0

coor=nan(1,4);  % stores coordinates of entries
                % of the Covariance matrix that
                % contain the same value by
                % symmetry arguments

for j=1:kappa
    value=2*TriMa(j+1,j+1);

    coor(1)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,j+kappa+1);
    coor(2)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,j+kappa+1);
    coor(3)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,-j+kappa+1);
    coor(4)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,-j+kappa+1);

    C1(coor)=value;

```

```
end


% C_{1,(1,0)}

if kappa>0
    value=2*(TriMa(2,1)-TriMa(2,2)-TriMa(1,1));

    coor(1)=sub2ind([2*kappa+1 2*kappa+1],kappa+2,kappa+1);
    coor(2)=sub2ind([2*kappa+1 2*kappa+1],kappa+1,kappa+2);
    coor(3)=sub2ind([2*kappa+1 2*kappa+1],kappa,kappa+1);
    coor(4)=sub2ind([2*kappa+1 2*kappa+1],kappa+1,kappa);

    C1(coor)=value;
end


% C_{1,(j,0)}, j > 1

if kappa > 1
    for j=2:kappa
        value=2*(TriMa(j+1,1)-TriMa(j+1,2)-TriMa(j,1)+TriMa(j,2));

        coor(1)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,kappa+1);
        coor(2)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,kappa+1);
        coor(3)=sub2ind([2*kappa+1 2*kappa+1],kappa+1,j+kappa+1);
        coor(4)=sub2ind([2*kappa+1 2*kappa+1],kappa+1,-j+kappa+1);

        C1(coor)=value;
    end
end


% C_{1,(j,k)},  0 < k = j+-1

coor=nan(8,1);

if kappa>1
    for j=2:kappa
        value=TriMa(j+1,j)-TriMa(j+1,j+1)-TriMa(j,j);

        coor(1)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,j+kappa);
        coor(2)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,j+kappa)
            ;
        coor(3)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,-j+kappa
            +2);
        coor(4)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,-j+kappa
            +2);
```

```
            coor(5)=sub2ind([2*kappa+1 2*kappa+1],j+kappa,j+kappa+1);
            coor(6)=sub2ind([2*kappa+1 2*kappa+1],j+kappa,-j+kappa+1)
                ;
            coor(7)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+2,j+kappa
                +1);
            coor(8)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+2,-j+kappa
                +1);

            C1(coor)=value;
        end
end


% C_{1,(j,k)},   0 < k < j-1

if kappa>2
     for j=3:kappa
          for k=1:j-2
               value=TriMa(j+1,k+1)-TriMa(j+1,k+2)-TriMa(j,k+1)+TriMa
                    (j,k+2);
               coor(1)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,k+
                    kappa+1);
               coor(2)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,k+
                    kappa+1);
               coor(3)=sub2ind([2*kappa+1 2*kappa+1],j+kappa+1,-k+
                    kappa+1);
               coor(4)=sub2ind([2*kappa+1 2*kappa+1],-j+kappa+1,-k+
                    kappa+1);
               coor(5)=sub2ind([2*kappa+1 2*kappa+1],k+kappa+1,j+
                    kappa+1);
               coor(6)=sub2ind([2*kappa+1 2*kappa+1],k+kappa+1,-j+
                    kappa+1);
               coor(7)=sub2ind([2*kappa+1 2*kappa+1],-k+kappa+1,j+
                    kappa+1);
               coor(8)=sub2ind([2*kappa+1 2*kappa+1],-k+kappa+1,-j+
                    kappa+1);

               C1(coor)=value;
          end
     end
end

C1=n^(-2-a)*C1;

end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function covM = Cov2(kappa,a,n)
```

```
% Returns the (2kappa+1)^4 array covM with entries
% covM(j1,j2,k1,k2)=
% C_{(j_1-kappa-1,j_2-kappa-1),(k_1-kappa-1,k_2-kappa-1)}.

format long;
TriMa=TriIntMat(kappa,a);
covM = nan(2*kappa+1,2*kappa+1,2*kappa+1,2*kappa+1);


% C_{(0,0),(0,0)}

covM(kappa+1,kappa+1,kappa+1,kappa+1)=8*TriMa(1,1);


% C_{(j,j),(j,j)}, for j > 0:

for j=1:kappa
    coor=symind([j+kappa+1; j+kappa+1; j+kappa+1; j+kappa+1],kappa);
    covM(coor)=2*TriMa(j+1,j+1);
end


% C_{(1,0),(1,0)}:

if kappa>0
    value=2*(TriMa(2,1)-TriMa(2,2)-TriMa(1,1));
    coor=symind([kappa+2; kappa+1; kappa+2; kappa+1],kappa);
    covM(coor)=value;
end


% C_{(j,0),(j,0)}, j > 1:

if kappa > 1
    for j=2:kappa
        value=2*(TriMa(j+1,1)-TriMa(j+1,2)-TriMa(j,1)+TriMa(j,2));
        coor=symind([j+kappa+1; kappa+1; j+kappa+1; kappa+1],kappa);
        covM(coor)=value;
    end
end


% C_{(j,k),(j,k)}, 0 < k = j-1:

if kappa>1
```

```
    for j=2:kappa
        value=TriMa(j+1,j)-TriMa(j+1,j+1)-TriMa(j,j);
        coor=symind([j+kappa+1; j+kappa; j+kappa+1; j+kappa],
            kappa);
        covM(coor)=value;
    end
end


% C_{(j,k),(j,k)},  0 < k < j-1:

if kappa>2
    for j=3:kappa
        for k=1:j-2
            value=TriMa(j+1,k+1)-TriMa(j+1,k+2)-TriMa(j,k+1)+TriMa
                (j,k+2);
            coor=symind([j+kappa+1; k+kappa+1; j+kappa+1; k+kappa
                +1],kappa);
            covM(coor)=value;
        end
    end
end


% The remaining entries are filled by numeric integration.
% The following loop computes the corresponding integrals
% for all slots of C that still contain a NaN.

for j1=1:kappa
    for j2=0:j1
        for k1=-kappa:kappa
            for k2=-kappa:kappa
                if isnan(covM(j1+kappa+1,j2+kappa+1,k1+kappa+1,k2+
                    kappa+1))
                    fun=@(x,y)(((j1-x).^2+(j2-y).^2).^(a/2).*((k1-
                        x).^2+(k2-y).^2).^(a/2));
                    value=integral2(fun,-0.5,0.5,-0.5,0.5,'AbsTol'
                        ,1e-20,'RelTol',0);
                    coor=symind([j1+kappa+1; j2+kappa+1; k1+kappa
                        +1; k2+kappa+1],kappa);
                    covM(coor)=value;
                end
            end
        end
    end
end

covM=n^(-2-2*a)*covM;
```

```

end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function  C = Cov3(kappa,a,n)

% returns the covariance matrix of the vector
% (W_0^n,W^1',...,W^(2kappa+1)'),
% where
% W^k'=(W_{(0,0),(-kappa,k-kappa-1)},...,W_{(0,0),(kappa,k-kappa
    -1)}).

A=Cov2(kappa,a,n);
A=reshape(A,[(2*kappa+1)^2,(2*kappa+1)^2]);

B=Cov1(kappa,a,n);
B=reshape(B,[], 1);

C=nan((2*kappa+1)^2+1);
C(1,1)=1/(n^2);
C(2:(2*kappa+1)^2+1,1)=B;
C(1,2:(2*kappa+1)^2+1)=B.';
C(2:(2*kappa+1)^2+1,2:(2*kappa+1)^2+1)=A;
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [  Mat ] = LgMatMatern(n,N,a,kappa,b)

% Mat contains the values L(\bk/n) for
% \bk\in {-kappa,...,kappa}^2, and the
% values g(\bk/n) for
% \bk\in {-N,...,N}^2 \setminus {-kappa,...,kappa}^2
% for the matern covariance case. In order
% to minimise function calls, we compute LgMat
% only on half a quadrant and exploit symmetries.

% Matern covariance kernel:
 lambda=1;
 Lfct = @(x)( norm(x)^(-a/2)*besselk(a/2,lambda*norm(x))) ;
 Lfct1d = @(x)( abs(x).^(-a/2).*besselk(a/2,lambda*abs(x))) ;


Mat=nan(2*N+1);
for i=0:N
    for j=0:i
        if abs(i)>kappa | abs(j)>kappa
```

```
                Mat(i+N+1,j+N+1)=Lfct1d(b(i+N+1,j+N+1)/n)*(b(i+N+1,j+N
                    +1)/n)^a;
            else
                Mat(i+N+1,j+N+1)=Lfct1d(b(i+N+1,j+N+1)/n);
            end
    end
end


% For the central square [-1/n,1/n]^2 the value
% of L at the optimal discretisation point is obtained
% as follows. The function TriInt0 is listed below
intfct = @(x)(Lfct1d(x/n).*(x.^(2*a+1)).*(pi/4-(x>=1/(2)).*acos(
    sqrt(2)*x)));
Mat(N+1,N+1) =integral(intfct,0,1/sqrt(2))./TriInt0(1/2,a);

% The rest of the matrix is filled by using symmetries

for i=0:N-1
    for j=i+1:N
        Mat(i+N+1,j+N+1)=Mat(j+N+1,i+N+1);
    end
end

for i=-N:-1
    for j=-N:-1
         Mat(i+N+1,j+N+1)=Mat(-i+N+1,-j+N+1);
    end
    for j=0:N
         Mat(i+N+1,j+N+1)=Mat(-i+N+1,j+N+1);
    end
end

for i=0:N
    for j=-N:-1
         Mat(i+N+1,j+N+1)=Mat(i+N+1,-j+N+1);
    end
end

end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function x = TriangleIntegral(j1,j2,a)

% Takes input (j1,j2) with 0<j2<j1 and computes the
% integral of \|x\|^2a over the set
% {(x,y) : 0 < x < i-1/2, j-1.5 < y < x}.
% see Appendix III.B for details.
```

```
x=((j1.^(2*a+2)+j2.^(2*a+2))/(2^(1.5)*(a+1))).*hypergeom([0.5,1.5+
    a],1.5,0.5);
x=x-(j1.*j2.^(2*a+2)).*hypergeom([0.5,1.5+a],1.5,j1.^2/(j1.^2+j2
    .^2))./(2*(a+1)*sqrt(j1.^2+j2.^2));
x=x-(j1.^(2*a+2).*j2).*hypergeom([0.5,1.5+a],1.5,j2.^2/(j1.^2+j2
    .^2))/(2*(a+1)*sqrt(j1.^2+j2.^2));
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function x = TriInt0(j1,a)

% Returns the integral of \|x\|^{2a} over the
%  triangle {(x,y) : 0 < x < 0.5, 0 < y < x}.

x=sqrt(2)*j1^(2*a+2)*hypergeom([0.5,1.5+a],1.5,0.5)/(4*(a+1));

end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function A = TriIntMat(kappa,a)

% TriIntMat contains integrals of \|x\|^a over
% triangular sets in the following structure.
% TriIntMat is a symmetric matrix. Its first
% column contains the entries
% TriIntMat(i,1) = integral of \|x\|^a over the set
% {(x,y) : 0 < x < i-1/2, 0 < y < x}.
% For all other columns, i.e. with j>1
% we have
% TriIntMat(i,j) = integral of \|x\|^a over the set
% {(x,y) : 0 < x < i-1/2, j-1.5 < y < x}
% See Appendix III.B for details.

A=nan(kappa+1,kappa+1);

A(1,1)= TriInt0(0.5,a);

for j1=2:kappa+1
A(1,j1)=TriInt0(j1-0.5,a);
A(j1,1)=A(1,j1);
    for j2=2:j1
        A(j1,j2)=TriangleIntegral(j1-0.5,j2-1.5,a);
        A(j2,j1)=A(j1,j2);
    end
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [ sigma ] = vol( n,N )

% Contains the values of the volatility field sigma.

sigma=ones(2*N+2*n+1);

end
```