Functional weak convergence of stochastic integrals for moving averages and continuous-time random walks

Andreas Søjmark* and Fabrice Wunderlich†

Abstract

There is by now an extensive theory of weak convergence for moving averages and continuous-time random walks (CTRWs) with respect to Skorokhod's M1 and J1 topologies. Here we address the fundamental question of how this translates into functional limit theorems, in the M1 or J1 topology, for stochastic integrals driven by these processes. As an important application, we provide weak approximation results for general SDEs driven by time-changed Lévy processes. Such SDEs and their associated fractional Fokker–Planck–Kolmogorov equations are central to models of anomalous diffusion in statistical physics. Our results yield a rigorous functional characterisation of these as continuum limits of the underlying models driven by CTRWs. In regard to strictly M1 convergent moving averages and correlated CTRWs, it turns out that the convergence of stochastic integrals can fail markedly. Nevertheless, we are able to identify natural classes of integrand processes for which M1 convergence holds. We show that these results are general enough to yield functional limit theorems, in the M1 topology, for certain stochastic delay differential equations driven by moving averages.

1. Introduction

In view of the central limit theorem—and its functional extensions—the concept of diffusion is a remarkably robust modelling paradigm: with the appropriate square-root scaling, it gives the correct macroscopic description of any random walk whose i.i.d. jumps are of finite variance. Yet, many phenomena across the natural and social sciences involve heavy-tailed power laws that lead to different notions of anomalous—or fractional—diffusion for the macroscopic movement of particles (be they actual particles or some quantities identified as such). This deviation from classical diffusion lies at the heart of a fast growing field of research known as fractional calculus, sitting firmly at the interface of mathematical physics, probability theory, and the theory of partial differential equations.

Befittingly, anomalous diffusion, too, is underpinned by a robust class of limit theorems for random walks, now with jumps of infinite variance. In fact, not only jumps of infinite variance are relevant, but also infinite mean waiting times between the jumps, thus leading to the concept of continuous-time random walks (CTRWs). For details, see Section 3.1. As long as the jumps and waiting times are, in a certain sense, attracted to given stable laws, the appropriate scaling limit will be described by a fractional equivalent of the heat equation for classical diffusion. Here a fractional derivative in time captures an element of sub-diffusion due to infinite-mean power laws of the waiting times, while fractional derivatives in space capture elements of super-diffusion due to infinite-variance power laws of the jumps.

For an excellent introduction to anomalous diffusion and the field of fractional calculus from a probabilistic viewpoint, we refer to the monograph of Meerschaert & Sikorskii [45]. In particular, [45, Ch. 4] covers the underlying functional limit theory for CTRWs. The main definitions and results needed for the present paper are recalled in Section 3.1.

*London School of Economics, Department of Statistics, London, WC2A 2AE, UK. a.sojmark@lse.ac.uk
†University of Oxford, Mathematical Institute, Oxford, OX2 6GG, UK. wunderlich@maths.ox.ac.uk

As one might expect, stochastic calculus for CTRWs and their scaling limits play an important role in the study of anomalous diffusion; see e.g. the treatment of stochastic calculus for CTRWs in [19] and of stochastic calculus for time-changed Lévy processes in [36, 51]. Since the justification for continuum models of anomalous diffusion comes from CTRWs and their functional limit theorems, it is essential to understand the behaviour of stochastic integrals driven by CTRWs, when these integrators converge to their scaling limit. In this paper, we shall address this question in the sense of weak convergence on Skorokhod space with respect to Skorokhod's J1 and M1 topologies, depending on the structure of the CTRWs.

1.1. Anomalous diffusion and convergence of stochastic integrals

In the physics literature, models of anomalous diffusion are often formulated in terms of fractional Fokker–Planck–Kolmogorov equations. The intuition is similar to that of the fractional equivalents of the heat equation discussed above, only now the equations take the general form

$$D_t^{\beta} p(t,x) = \mathcal{A}_x p(t,x), \quad p(0,x) = p_0, \tag{1.1}$$

where D_t^{β} is a fractional Caputo derivative of order $\beta \in (0,1)$ and $\mathcal{A}_x = \mathcal{A}(x, D_x)$ is a pseudo-differential operator with a given symbol $\Psi(x,\xi)$. One can also have a distributed-order fractional derivative $D_t^{\mu} = \int_0^1 D_t^{\beta} \mu(\mathrm{d}\beta)$ instead of a standalone Caputo derivative.

In many applications, the Cauchy problem (1.1) is intimately linked to an underlying CTRW model and may be derived in that way, as in the pioneering work of Metzler, Barkai & Klafter [46]. See also Metzler & Klafter's influential survey papers [47, 48]. Recently, there has been a growing interest in stochastic representations for variants of (1.1), see e.g. [4, 21, 20, 22, 39, 40, 41, 42]. This provides a rigorous connection between the Fokker–Planck–Kolmogorov formulations and single-particle tracking for anomalous diffusion (as surveyed by Metzler et al. [49]), and it also open up for Monte Carlo methods in the numerical approximation of (1.1).

The stochastic representations of (1.1) generally involve SDEs driven by time-changed Lévy processes, where the symbol $\Psi(x,\xi)$ of $\mathcal{A}_x$ corresponds to the characteristic exponent of the parent Lévy process, while β is the index of stability for a stable subordinator whose generalised inverse yields the time-change. The analysis of such SDEs is of independent interest and there is a growing literature on both their qualitative properties, see e.g. [36, 51, 50], and numerical schemes, see e.g. [17, 29, 31]. Likewise, there is an interest in understanding time-changed Lévy processes on their own, as in, e.g., [38] motivated by the study of CTRWs, [11] related to applications in ruin theory, and [10] pertaining to applications in finance.

Following on from the above, Hahn, Kobayashi & Umarov set forth a unifying paradigm for the study of anomalous diffusion in their recent monograph [63]. In short, they treat the driving process X—here a time-changed Lévy process—as the central object and call for an exhaustive exploration of the interconnections between the following three pillars surrounding it: (i) the underlying limit theory for CTRWs, (ii) the analysis of SDEs driven by X, and (iii) the analysis of the corresponding fractional Fokker–Planck–Kolmogorov equations.

Regarding (i) and (ii), Section 6.5 in [63, Ch. 6] stresses that an important but less studied question is that of functional weak convergence for stochastic integrals driven by CTRWs. Moreover, noting that this lies beyond the reach of existing results, Remark 6.4 of [63, Ch. 6] discusses the interest in applying such machinery to establish functional limit theorems for SDEs driven by CTRWs, conceivably leading to limiting SDEs driven by time-changed Lévy-processes in line with the analysis of such SDEs in [63, Chs. 6-7].

In this paper, we provide a systematic treatment of the functional weak convergence of stochastic integrals driven by CTRWs, both when the innovations are i.i.d. and when they display a linear correlation structure. The former leads to universal results on J1 convergence at a level of generality that the approaches of earlier works did not allow for (see [9, 52, 58] and [63, Sect. 6.5]), while the latter leads to tailored results on M1 convergence which, as far as the authors are aware, have not previously been addressed in the literature and which require entirely new ideas. As an important first application of our results, Section 5.1 derives the desired functional limit theorems for SDEs driven by CTRWs called for in [63].

A related area of application could be functional limit theorems for stochastic partial differential equations subject to Lévy noise, e.g. stochastic generalisations of (1.1) as in [34, 35], provided the Lévy noise is seen as a scaling limit of CTRWs. Moreover, as discussed in [15, 14], for many physical models one is directly interested in stochastic integrals of some kernel against a stable Lévy noise justified by stable limit theorems, thus immediately raising the issue of the convergence of the corresponding stochastic integrals driven by CTRWs.

1.2. Applications in the social sciences

As already alluded to above, CTRWs and time-changed Lévy processes are important objects in finance, insurance, and economics. A powerful option pricing theory for time-changed Lévy processes with tempering has recently been proposed in [26] motivated by the tick-by-tick CTRW models introduced in [57]. Another recent paper [1] studies the link between tick-by-tick and continuum asset price models, defining them to be compatible if there is weak M1 convergence. This naturally entails the question of what can be said about the functional weak convergence of the financial gains (i.e., the stochastic integrals of trading strategies against price processes), as the driving CTRWs converge to their scaling limit, noting that the latter takes place in either the M1 or J1 topology depending on whether the CTRWs have correlated innovations. The classical case of limiting price processes given by geometric Brownian motion was treated in [12].

In insurance mathematics, certain SDEs driven by Lévy processes are the central objects in ruin theory with risky investments [54]. Brownian dynamics are justified by suitable scaling limits of SDEs involving compound Poisson processes [55]. More generally, we can consider J1 or M1 convergent CTRW models for the claims and the risky investments. Our results then give convergence to the corresponding SDEs driven by time-changed Lévy processes.

Finally, as covered by [52, 53], functional limit theorems for stochastic integrals driven by random walks and moving averages, converging to stable Lévy processes, play an important role in statistical inference for cointegrated processes in econometric theory. [52] gives a rigorous treatment for J1 convergent random walks, while [53] pinpoints some imprecisions in the literature and highlights the lack of a systematic treatment for M1 convergent moving averages. Concerning the latter, we note that also [27, Example 2] brought attention to the relevance of exploring functional limit theorems for stochastic integrals driven by moving averages. As we shall see, existing results fall short and it turns out to be far from trivial what one can say.

Rather than going into details about the precise applications, for brevity we instead refer the reader to a companion paper [61]. There, we apply the results of the present paper to handle concrete problems emerging from the three strands of literature discussed above.

1.3. Key contributions and overview of the paper

The overall aim of this paper is to give a systematic treatment of weak M1 and J1 convergence for stochastic integrals driven by moving averages or CTRWs. Our starting point will be the general framework for weak convergence of stochastic integrals presented in [62], building on the seminal works [27, 28, 37]. Here, however, we present a more tailored analysis, exploiting the structure of the particular classes of integrators to obtain results where the general theory does not apply. Section 2 briefly recalls the central concepts from [62], while Section 3.1 covers the precise definitions of moving averages and CTRWs along with the associated limit theory.

In Section 3.2, our first contribution is a proof that uncorrelated, possibly coupled, CTRWs have good decompositions in the sense of Definition 2.1 (Theorem 3.3). Armed with these good decompositions, we then derive a universal result on the weak J1 convergence of stochastic integrals driven by uncorrelated CTRWs (Theorem 3.6). This provides a positive answer to the open question discussed by Hahn, Kobayashi & Umarov in [63, Rem. 6.4]. Furthermore, Section 3.3 derives an analogous result on weak M1 convergence in the case of moving averages and correlated CTRWs with a finite variation scaling limit (Theorem 3.9).

When the scaling limits display infinite variation, it was shown in [62, Prop. 4.6] that there can be a severe failure of weak convergence for stochastic integrals driven by strictly M1 convergent moving averages or correlated CTRWs (something which comes down to a lack of good decompositions). While the work in Section 3 follows closely the framework of [62], we thus need to develop a more flexible—but less universal—theory in order to handle general moving averages and correlated CTRWs. This is the topic of Section 4.

Compared to [62], the overarching idea in Section 4 is to allow for a more general class of integrators that only admit good decompositions after introducing a suitable remainder term, which has to be well-behaved with respect to the integrands. Section 4.1 presents an approach based on direct control of the variation of the remainder (Theorem 4.6), while Section 4.2 explores a procedure based on a certain independence condition between the integrands and the 'future' of the remainder (Theorem 4.11). We suspect that both approaches should be more widely applicable, but here our focus is on showing that they work well for moving averages and correlated CTRWs. In turn, Section 4.3 provides two general results on the M1 convergence of stochastic integrals driven by such processes (Theorems 4.14 and 4.15).

We end the paper by applying the above analysis to address fundamental convergence questions for SDEs and stochastic delay differential equations (SDDEs). Section 5.1 completes the programme outlined in Section 1.1, showing that the solutions to the SDEs of interest, driven by uncorrelated CTRWs, indeed converge weakly in J1 to the corresponding SDEs driven by time-changed Lévy processes (Theorem 5.1). Finally, Section 5.2 shows how the independence framework of Section 4.2 can be used to obtain M1 convergence for certain SDDEs driven by strictly M1 convergent moving averages (Theorem 5.2).

The proofs of most results are postponed to Section 6. There, we give the proofs in the same order as they appear in the main body of the paper.

2. Stochastic integral convergence on Skorokhod space

This section recalls key elements of the general framework for stochastic integral convergence in the M1 and J1 topologies from [62]. Starting with the notation, we shall write $\mathbf{D}_{\mathbb{R}^d}[0,\infty)$ for the Skorokhod space consisting of all càdlàg paths $x:[0,\infty)\to\mathbb{R}^d$, for a given dimension $d\geq 1$. Moreover, we shall use d_{J1} and d_{M1} to refer to a fixed choice of metrics that induce, respectively, the J1 and M1 topologies on this space. For details on these topologies, we refer to [62, Appendix A]. The first key ingredient is a uniform regularity condition on the integrators.

Definition 2.1 (Good decompositions, [62, Def. 3.3])**.** Let $(X^n)_{n\geq 1}$ be a sequence of semimartingales on probability spaces $(\Omega^n,\mathcal{F}^n,\mathbb{F}^n,\mathbb{P}^n)$. The sequence is said to have *good decompositions (GD)* if, for the given filtrations $\mathbb{F}^n$, there exist decompositions

$$X^n = M^n + A^n, \quad M^n \text{ local martingales}, \quad A^n \text{ finite variation processes},$$

such that, for every $t>0$, we have

$$\lim_{R\to\infty}\limsup_{n\to\infty}\mathbb{P}^n\big(\mathrm{TV}_{[0,t]}(A^n)>R\big)=0 \qquad\text{and}\qquad \limsup_{n\to\infty}\mathbb{E}^n\big[\,|\Delta M^n_{t\wedge\tau^n_c}|\,\big]<\infty, \tag{GD}$$

for all $c>0$, where $\tau^n_c:=\inf\{s>0:|M^n|^*_s\geq c\}$. Here $\mathrm{TV}_{[0,t]}(A^n)$ denotes the total variation of A^n on $[0,t]$ and $\Delta M^n_t := M^n_t - M^n_{t-}$ denotes the jump of M^n at time t.

We note that (GD) plays a role analogous to that of the P-UT and UCV conditions used in [27, 28, 37] (for details on how these compare with (GD), see [62]). To address the interplay between integrands and integrators, we define a function $\hat{w}_\delta^T : \mathbf{D}_{\mathbb{R}^d}[0,\infty) \times \mathbf{D}_{\mathbb{R}^d}[0,\infty) \to \mathbb{R}_+$ of the largest consecutive increment within a δ period of time on $[0,T]$, namely

$$\hat{w}_\delta^T(x,y) := \sup\Big\{|x^{(i)}(s) - x^{(i)}(t)| \wedge |y^{(i)}(t) - y^{(i)}(u)| : s < t < u \leq (s+\delta),\, 1 \leq i \leq d\Big\},$$

where the supremum is restricted to $0 \leq s, u \leq T$. Here we have used the usual notation $a \wedge b := \min\{a,b\}$, and we have denoted by $x^{(i)}$ the i-th coordinate of x. In addition to the pivotal (GD) property, the second essential ingredient is the following condition.

Definition 2.2 (Asymptotically vanishing consecutive increments, [62, Def. 3.2])**.** Let $(X^n)_{n\geq 1}$ and $(H^n)_{n\geq 1}$ be d-dimensional càdlàg processes on given probability spaces $(\Omega^n, \mathbb{F}^n, \mathbb{P}^n)$. The sequence $(H^n, X^n)_{n\geq 1}$ is said to satisfy the *asymptotically vanishing consecutive increments* condition if, for every $\gamma > 0$ and $T > 0$, it holds that

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \mathbb{P}^n\big(\, \hat{w}_\delta^T(H^n,\, X^n) \;>\; \gamma\,\big) \;=\; 0. \tag{AVCI}$$

Whilst the above formulation is taken from [62, Sect. 3.1], we stress that the idea of enforcing (AVCI) comes from [27], as discussed in more detail in [62]. Both for intuition and applications, it is useful to keep in mind the following simple sufficient criteria.

Proposition 2.3 ([62, Prop. 3.8])**.** *In the setting of Theorem 2.4, the condition* (AVCI) *is satisfied if one of the following two criteria holds:*

1. *The pairs (H^n, X^n) converge together to (H, X) weakly in the J1 topology, meaning that the joint weak convergence $(H^n, X^n) \Rightarrow (H, X)$ holds on $(\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty),\, \mathrm{d}_{\mathrm{J1}})$.*
2. *The limiting processes H and X almost surely have no common discontinuities, that is,*

$$\mathrm{Disc}(H) \;\cap\; \mathrm{Disc}(X) \;=\; \emptyset \quad a.s.$$

We shall also point out that an alternative criterion to (AVCI) is developed in [62, Thm. 4.8] which we will use for the formulation of Theorem 4.15. However, for the majority of this paper, we shall rely on (AVCI). Equipped with (GD) and (AVCI), we have the following general result on the weak continuity properties of stochastic integrals on Skorokhod space.

Theorem 2.4 (Weak continuity of stochastic integrals [62, Thm. 3.6])**.** *For given filtered probability spaces $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, \mathbb{P}^n)$, consider a sequence of semimartingales $(X^n)_{n\geq 1}$ with good decompositions* (GD)*. Let $(H^n)_{n\geq 1}$ be any given sequence of adapted càdlàg processes for the same filtered probability spaces such that (i) there is joint weak convergence*

$$(H^n, X^n) \;\Rightarrow\; (H, X) \quad \text{on} \quad (\mathbf{D}_{\mathbb{R}^d}[0,\infty)\,,\, \tilde{\rho}\,) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty)\,,\, \rho)$$

with $\rho, \tilde{\rho} \in \{\mathrm{d}_{\mathrm{M1}}, \mathrm{d}_{\mathrm{J1}}\}$, for some càdlàg limits H and X, and (ii) the pairs (H^n, X^n) satisfy (AVCI)*. Then, X is a semimartingale in the filtration generated by the pair (H, X) and*

$$\left(X^n,\, \int_0^\bullet H^n_{s-}\, \mathrm{d}X^n_s \right) \;\Rightarrow\; \left(X,\, \int_0^\bullet H_{s-}\, \mathrm{d}X_s \right) \quad \text{on} \quad (\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty),\, \rho). \tag{2.1}$$

In (2.1), the integrals are understood to be defined component-wise. This also allows one to handle matrix valued processes and 'dot product' integrals, as per [62, Rem. 3.11]. As per [62, Rem. 3.9], one can also consider the so-called *weak* M1 topology in Theorem 2.4.

3. On moving averages, CTRWs, and their good decompositions

In this section, we first cover the classical results on functional CLTs for moving averages and CTRWs. We then examine their regularity as integrators, in the sense of good decompositions or a lack thereof, and, finally, derive our first results on weak convergence of stochastic integrals.

3.1. Basic definitions and stable scaling limits

We begin by recalling that a *moving average* (suitably re-scaled) is a continuous-time stochastic process of the form

$$X_t^n := \frac{1}{n^{\frac{1}{\alpha}}} \sum_{k=1}^{\lfloor nt \rfloor} \zeta_k, \qquad t \geq 0, \qquad n \geq 1, \tag{3.1}$$

where the innovations are given by

$$\zeta_i := \sum_{j=0}^{\infty} c_j \theta_{i-j}, \qquad i \geq 1, \tag{3.2}$$

for an i.i.d. sequence $\{\theta_k : -\infty < k < \infty\}$ of $\mathbb{R}^d$-valued random variables (on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$). Here, the θ_k are asummed to be in the *normal domain of attraction* of a non-degenerate α-stable random variable $\tilde{\theta}$ with $0 < \alpha < 2$, i.e.

$$n^{-\frac{1}{\alpha}} (\theta_1 + \ldots + \theta_n) \Rightarrow \tilde{\theta} \tag{3.3}$$

in $\mathbb{R}$. Recall that $\tilde{\theta}$ is said to have a strictly α-stable law on $\mathbb{R}$ if there exist i.i.d. copies $\tilde{\theta}_1, \tilde{\theta}_2, ...$ of $\tilde{\theta}$ such that $\tilde{\theta}_1 + \ldots + \tilde{\theta}_n \sim n^{\frac{1}{\alpha}} \tilde{\theta}$ for all $n \geq 1$. For $\alpha = 2$, (3.3) holds instead with $\tilde{\theta}$ being a non-degenerate Gaussian random variable. Throughout, we assume $c_j \geq 0$ for all j, and we require that $\sum_{j=0}^{\infty} c_j^{\rho} < \infty$, for some $0 < \rho < \alpha$. The latter ensures that the series (3.2) converges in $L^{\rho}(\Omega, \mathcal{F}, \mathbb{P})$—and, in fact, almost surely (see [3, 33]).

For a *zero-order moving average* (i.e., $c_0 > 0$ and $c_j = 0$ for all $j \geq 1$), it is a classical result of Skorokhod [59] that $X^n \Rightarrow Z$ in $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d_{J1}})$, where Z is a Brownian motion if $\alpha = 2$ or an α-stable Lévy process (with $Z_1 \sim \tilde{\theta}$) if $0 < \alpha < 2$. Avram & Taqqu [3] studied functional convergence in the general case, showing that $X^n \Rightarrow (\sum_{j=0}^{\infty} c_j) Z$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d_{M1}})$, if $\mathbb{E}[\theta_1] = 0$ when $1 < \alpha < 2$ or if the law of θ_1 is symmetric when $\alpha = 1$. We shall assume throughout that, if $1 < \alpha \leq 2$, then either $c_j = 0$ for all but finitely many $j \geq 0$ or the sequence $(c_j)_{j \geq 0}$ is monotone and $\sum_{j=0}^{\infty} c_j^{\rho} < \infty$ for some $\rho < 1$. Under these assumptions, the aforementioned restrictions in the result of Avram & Taqqu can be omitted (see e.g. [64, Thm. 4.7.1]). Finally, [3] also showed that if not just $c_0 > 0$ but also $c_j > 0$ for at least one $j \geq 1$, then the convergence *cannot* be strengthened to hold in the J1 topology.

Continuous-time random walks (CTRWs) are generalisations of moving averages (3.1), allowing for random waiting times J_i (with infinite mean) in between jumps. More precisely, let $J_1, J_2, ...$ be i.i.d. random variables in the normal domain of attraction of a β-stable random variable with $\beta \in (0,1)$, defined on the same probability space as the θ_k above. A CTRW (suitable re-scaled), then takes the form

$$X_t^n := \frac{1}{n^{\frac{\beta}{\alpha}}} \sum_{k=1}^{N(nt)} \zeta_k, \qquad N(nt) := \max\{m \geq 0 : L(m) \leq nt\}, \tag{3.4}$$

with $L_m := L(m) := J_1 + \ldots + J_m$ and $L(0) \equiv 0$. Note that $N(nt)$ gives the number of jumps up until time nt. In the literature two basic types are distinguished: a CTRW is said to be *uncorrelated* if $c_j = 0$ for all $j \geq 1$ and *correlated* otherwise. Further, we will also use the term *finitely correlated* whenever there exists $\mathcal{J} \geq 1$ such that $c_j = 0$ for all $j > \mathcal{J}$.

If the sequences $(J_i)_{i \geq 1}$ and $(\zeta_k)_{k \geq 1}$ are independent, the CTRW (3.4) is typically called *uncoupled*. For this (uncoupled) setting, Becker-Kern, Meerschaert & Scheffler [6] as well as Meerschaert, Nane & Xiao [43] extended the results of [3] for moving averages, showing that CTRWs exhibit a similar scaling-limit behaviour. That is, for $0 < \alpha \leq 2$, we have

$$n^{-\beta} N(n\bullet) \Rightarrow D^{-1} \tag{3.5}$$

on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{J1}})$ as well as

$$X^n \;\Rightarrow\; \Big(\sum_{j=0}^{\infty} c_j\Big) Z_{D^{-1}} \tag{3.6}$$

on the M1 Skorokhod space $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$, where Z is either a Brownian motion ($\alpha = 2$) or an α-stable Lévy process ($0 < \alpha < 2$), and the process

$$D_t^{-1} \;:=\; \inf\big\{\, s \geq 0 \,:\, D(s) \,>\, t \,\big\}$$

is the generalised inverse of a β-stable subordinator D with $\beta \in (0,1)$. Since D is strictly increasing, D^{-1} is a continuous process. Moreover, based on the arguments of [3] for moving averages, it is also shown in [43] how the convergence *cannot* be strengthened to J1 if there is $j \neq i$ such that $c_i, c_j > 0$. If $c_0 > 0$ is the only non-zero constant, then the M1 convergence (3.6) was first shown in [6] and it was later improved to hold in the J1 topology by [23].

One can also consider *coupled* CTRWs. Let (3.4) be uncorrelated with $c_0 = 1$, then the CTRWs are said to be *coupled* if the pairs of associated innovations and jump times $(\zeta_k, J_k)_{k\geq 1}$ constitute an i.i.d. sequence, while however, we allow for dependence of the components ζ_k and J_k of each pair. In this (coupled) framework, the results of [23, 32] yield

$$X^n \;\Rightarrow\; \big((Z^-)_{D^{-1}}\big)^+, \tag{3.7}$$

on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{J1}})$ for Z and D^{-1} as above and where we have used the notation $x^-(t) := x(t-)$ and $x^+(t) := x(t+)$. Should the CTRWs be uncoupled, it follows that Z and D arise as independent Lévy processes. Consequently, they almost surely have no common discontinuities, and the limit then simplifies to $Z_{D^{-1}}$ in agreement with (3.6) for uncorrelated CTRWs (for further details on this, see [23, Lem. 3.9]).

Remark 3.1 (Domain of attraction)**.** The above assumptions—on the jumps and the waiting times being in the normal domain of attraction of the respective stable laws—simply serve to ease notation. The convergence results discussed above also hold for random variables belonging only to the *strict domain of attraction.* Moreover, in the case of zero-order moving averages, Skorokhod [59] proved a similar convergence result for increments that are but in the domain of attraction of an α-stable law, where the limit then is an α-stable Lévy process or a Brownian motion with drift. Our subsequent results can all be extended to these more general settings in complete analogy with the procedure described next (in Section 3.2).

Remark 3.2 (Omitting the distant absolute past leaves results unchanged)**.** For moving averages and CTRWs, the dependence structure can be altered in such a way that there is but a dependence on finitely many innovations from the absolute past without affecting the convergence results. More precisely, if $\mathcal{J} \geq 1$ and we redefine

$$\zeta_i \;:=\; \sum_{j=0}^{i+\mathcal{J}} c_j \theta_{i-j},$$

then, if all other assumptions above remain in place, we still have $(\sum_{j=0}^{\infty} c_j)^{-1} X^n \Rightarrow Z_{D^{-1}}$ for X^n either a moving average or a correlated CTRW as in (3.1) or (3.4). This can be shown in full analogy to the proofs in [3, 43] with the remainder becoming asymptotically negligible.

Throughout the paper, if not stated otherwise, we will work on a given family of filtered probability spaces, denoted by $(\Omega, \mathcal{F}^n, \mathbb{F}^n, \mathbb{P})$, for $n \geq 1$, with filtrations $\mathbb{F}^n$ defined by

$$\mathcal{F}_t^n \;:=\; \sigma\Big(\sigma(\zeta^n_{N(ns)},\, N(ns) \,:\, 0 \leq s \leq t) \,\cup\, \mathcal{G}_t^n\Big), \qquad t \geq 0,\, n \geq 1, \tag{3.8}$$

where the $\mathcal{G}_s^n$ can be any system of measurable sets satisfying that, for X^n as in (3.4) with $c_j = 0$ for all $j \geq 1$, the independent increment property $X_t^n - X_s^n \perp\!\!\!\perp \mathcal{G}_s^n$ holds for all $0 \leq s < t$. As we will see below, the latter is essential in order to have good decompositions (GD) for uncorrelated CTRWs (see in particular Proposition 3.4). We note that each quantity $\zeta^n_{N(ns)}$ in (3.8) is, naturally, nothing else than $\sum_{k=1}^{\infty} \zeta_k^n \mathbf{1}_{\{N(ns)=k\}}$.

3.2. Good decompositions for uncorrelated CTRWs

In earlier works on stochastic integral convergence for uncorrelated CTRWs, only the uncoupled case is considered and the particular approaches necessitated quite restrictive assumptions, see [9, 52, 58] and the overview in [63, Ch. 6]. These works are all based on the general J1 theory developed in [28, 37] and hence rely on verifying the P-UT condition of [28] or the UCV condition of [37]. Both conditions imply (GD) in the present setting (see [62, Prop. B.3]).

In [58], a sufficient condition for P-UT and UCV is shown to hold for zero-order moving averages with $\alpha \in (1, 2]$ and innovations whose laws are symmetric. Through [44] and the time-change approach of [36], this is then enough to establish convergence for uncorrelated uncoupled CTRW integrators, as only a fixed continuous deterministic integrand is considered. Note that [58, Sect. 4.3] verifies the UCV condition directly for zero-order moving averages with $\alpha \in (0, 2]$ and without the symmetry assumption, but there is an error in the proof. In [9], a sufficient condition for P-UT and UCV is verified for uncorrelated, uncoupled CTRWs with $\alpha \in (1, 2]$, assuming centered innovations. Similarly, the analysis of zero-order moving averages in [52] relies on $\alpha \in (1, 2]$ and the innovations being centered.

None of the above arguments generalise to handle coupled CTRWs or the interesting critical case $\alpha = 1$ even for zero-order moving averages. It is worth briefly recounting the key step in the verification of the UCV condition in [52], as this is the closest to the approach we implement here. The setting is $\alpha > 1$ with deterministic waiting times (i.e., essentially $\beta = 1$) and the innovations θ_k being centered. Leading up to [52, Prop. 3], a decomposition similar to (3.9) is introduced. However, the proof of [52, Prop. 3] then exploits the existence of first moments of the θ_k to pass over to (in this case equivalently) proving $\sup_{n\geq 1} n\mathbb{E}[\zeta_1^n \mathbf{1}_{\{|\zeta_1^n|>a\}}] < \infty$. In general, if the innovations are not centered or if e.g. $\alpha \leq 1$, such a procedure fails to apply even for zero-order moving averages. Instead, exploiting the tail regularity of the θ_k and the convergence to a suitable time-changed Lévy process, we show that the decompositions (3.9) are good by virtue of Propositions 3.4 and 3.5 below. The proofs of these two propositions are provided in Appendix A.

Theorem 3.3 (Good decompositions in the uncorrelated case). *Let $(X^n)_{n\geq 1}$ be a sequence of uncorrelated coupled CTRWs or zero-order moving averages, as given by* (3.4) *or* (3.1) *with $c_0 = 1$ and $c_j = 0$ for all $j \geq 1$. Then, $(X^n)_{n\geq 1}$ has good decompositions* (GD).

We stress that the notion of good decompositions is dependent on the underlying filtration and here we consider the filtration (3.8). In order prove Theorem 3.3 for general coupled but uncorrelated CTRWs, we fix $a \geq 1$ and consider the decompositions

$$X_t^n = \sum_{k=1}^{N(nt)} \zeta_k^n = M_t^n + \sum_{k=1}^{N(nt)} \zeta_k^n \mathbf{1}_{\{|\zeta_k^n|>a\}} + N(nt)\,\mathbb{E}[\zeta_1^n \mathbf{1}_{\{|\zeta_1^n|\leq a\}}] \tag{3.9}$$

where, for simplicity, we have introduced $\zeta_k^n := n^{-\frac{\beta}{\alpha}}\zeta_k = n^{-\frac{\beta}{\alpha}}\theta_k$ so that $X_t^n = \sum_{k=1}^{N(nt)} \zeta_k^n$, and where we have defined

$$M_t^n := \sum_{k=1}^{N(nt)} \zeta_k^n \mathbf{1}_{\{|\zeta_k^n|\leq a\}} - N(nt)\mathbb{E}[\zeta_1^n \mathbf{1}_{\{|\zeta_1^n|\leq a\}}].$$

Proposition 3.4. *For each $n \geq 1$, the process M^n defined above is a martingale with respect to the filtration $(\mathcal{F}_t^n)_{t\geq 0}$ given by $\mathcal{F}_t^n := \sigma(\zeta^n_{N(ns)}, N(ns) : 0 \leq s \leq t)$, where $\zeta^n_{N(ns)}$ is understood as $\sum_{k=1}^{\infty} \zeta_k^n \mathbf{1}_{\{N(ns)=k\}}$. Furthermore, we have $|\Delta M^n| \leq 2a$, for all $n \geq 1$.*

Given that not only the number and size of large jumps of the uncorrelated CTRW are tight on compact time intervals but also the sequence $n^\beta N(n\bullet)$ (as a result of their tightness on the Skorokhod space, namely (3.5) and (3.6)), whether or not the X^n admit good decompositions ultimately depends on whether or not the supremum $\sup_{n\geq 1} n^\beta \mathbb{E}[\zeta_1^n \mathbf{1}_{\{|\zeta_1^n|\leq a\}}]$ is finite.

Define the truncation function $h(x) = x\,\mathbf{1}_{\{|x|\leq a\}} + \operatorname{sgn}(x)\,a\,\mathbf{1}_{\{|x|>a\}}$. Then, the condition $\sup_{n\geq 1} n^\beta \mathbb{E}[\zeta_1^n \mathbf{1}_{\{|\zeta_1^n|\leq a\}}] < \infty$ is equivalent to $\sup_{n\geq 1} n^\beta \mathbb{E}[h(\zeta_1^n)] < \infty$, since $n^\beta \mathbb{P}(|\zeta_1^n| > a) \to a^{-\alpha}$ as $n \to \infty$ (noting that $\mathbb{P}(|\theta_1| > x) = \mathcal{O}(x^{-\alpha})$ is known to be satisfied for random variables in the normal domain of attraction of an α-stable law, see e.g. [16],[64, Thm. 4.5.2]). It is a stronger version of the concept of the so called *infinitesimality* property described in more general settings by, e.g., [25, Def. VII.2.33]. We note that similar results to the ones below also hold under weaker assumptions in this more general framework, in particular without the assumption of identical distributions.

Proposition 3.5. *In the above setting, $\lim_{n\to\infty} n^\beta \mathbb{E}[h(\zeta_1^n)] = b$, for some $b \in \mathbb{R}$.*

By the above observations, this directly implies Theorem 3.3. Having established (GD) for zero-order moving averages and uncorrelated CTRWs, we can now state the following result on weak convergence of stochastic integrals driven by these very processes.

Theorem 3.6 (Weak integral convergence in the uncorrelated case)**.** *Let X^n be zero-order moving averages or uncorrelated CTRWs* (3.9) *with $0 < \alpha \leq 2$, defined on filtered probability spaces* (3.8)*. Further, let H^n be càdlàg adapted processes for the same filtered probability spaces. If $(H^n, X^n) \Rightarrow (H, X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{J1}})$ with X given by Z, $Z_{D^{-1}}$ or $((Z^-)_{D^{-1}})^+$, and if the (H^n, X^n) satisfy* (AVCI)*, then X is a semimartingale in the natural filtration generated by (H, X) and it holds that*

$$\left(X^n, \int_0^\bullet H^n_{s-}\,\mathrm{d}X^n_s\right) \implies \left(X, \int_0^\bullet H_{s-}\,\mathrm{d}X_s\right) \qquad \text{on} \qquad (\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty), \mathrm{d}_{\mathrm{J1}}). \tag{3.10}$$

Remark 3.7. We note that (AVCI) in Theorem 3.6 can be replaced by the conditions (a) & (b) set out in [62, Thm. 4.7]. In particular, if we are faced with uncorrelated and uncoupled CTRWs X^n, then these condition are indeed satisfied provided the integrands H^n, at each time t, depend only on the trajectory of X^n up to this time (where part (a) can be shown on behalf of [62, Lem. 4.11]). In specific cases, one might also be able to readily verify (AVCI) directly. For example, if we consider $H_t^n = \sum_{i=1}^\infty g_{n,i}(t_i^n, X^n_{t_i^n})\,\mathbf{1}_{(t_i^n, t_{i+1}^n]}(t)$, with equicontinuous $g_{n,i}$, we can make use of the alternative J1 tightness criteria given in [8, Thm. 12.4] and the J1 tightness of the integrators in order to show (AVCI).

3.3. Correlated CTRWs and moving averages do not blend in

As we discuss below, one cannot expect moving averages and correlated CTRWs to have good decompositions in general. However, there is no such problem when $0 < \alpha < 1$. The following result confirms that, in this case, these processes do indeed possess good decompositions (GD).

Proposition 3.8 (Good decompositions for $0 < \alpha < 1$)**.** *If $0 < \alpha < 1$, moving averages as in* (3.1) *and correlated CTRWs as in* (3.4) *are processes of tight total variation on compact time intervals and therefore admit good decompositions.*

Due to Proposition 3.8, we can extend Theorem 3.6 to cover integrators which are uncoupled correlated CTRWs or moving averages with $0 < \alpha < 1$. The proof of Proposition 3.8 can be found in Section 6.1 and the next result then follows from Theorem 2.4.

Theorem 3.9 (M1 integral convergence for $0 < \alpha < 1$)**.** *Theorem 3.6 also holds if the X^n are moving averages or uncoupled correlated CTRWs with $0 < \alpha < 1$, where either $X = (\sum_{j=0}^{\infty} c_j)Z$ or $X = (\sum_{j=0}^{\infty} c_j)Z_{D^{-1}}$, provided we replace replace* d_{J1} *with* d_{M1}.

In the case $1 \le \alpha \le 2$, the situation is very different from that of uncorrelated CTRWs or zero-order moving averages: even a single-delay correlation structure of the X^n generally does *not* translate into admitting good (GD) and the counterexample in [62, Prop. 4.5] demonstrates that this can result in a severe failure of the desired convergence (3.10), even for integrands H^n which converge to the zero process almost surely in the uniform norm.

In these circumstances, to preserve integral convergence we can seek to compensate the lack of (GD) by imposing additional restrictions on the interplay of integrands and integrators. Here we pursue such an extended framework for weak integral convergence on Skorokhod space, covering integrators that are—in a certain sense—close to admitting (GD).

4. A generalised framework for good decompositions

Considering the correlated CTRW X^n defined in (3.4) with $1 \le \alpha \le 2$, on a probability space with filtration (3.8), we can decompose $\psi^{-1}X^n = U^n + V^n$ with

$$U_t^n \;:=\; U_t^{n,1} \,+\, U_t^{n,2} \;:=\; \frac{1}{n^{\frac{\beta}{\alpha}}} \sum_{k=1}^{N(nt)} \theta_k \;+\; \frac{\psi^{-1}}{n^{\frac{\beta}{\alpha}}} \sum_{k=0}^{\infty} \Bigg(\sum_{j=k+1}^{N(nt)+k} c_j \Bigg) \theta_{-k}, \tag{4.1}$$

$$V_t^n \;:=\; -\frac{\psi^{-1}}{n^{\frac{\beta}{\alpha}}} \sum_{k=1}^{N(nt)} \Bigg(\sum_{j=k}^{\infty} c_j \Bigg) \theta_{N(nt)-k+1}, \quad \psi := \sum_{j=0}^{\infty} c_j. \tag{4.2}$$

We recall that the first summand $U^{n,1}$ of U^n is nothing else than an uncorrelated uncoupled CTRW and so, by Theorem 3.3, possesses good decompositions. If, by suitable assumptions, we ensure that the second summand of U^n is of tight total variation on compact time intervals, then it will come as no surprise that weak integral convergence can be achieved by simply controlling the interplay of the integrands H^n and the processes V^n. We are going to show that the second summand of $U^{n,2}$ of U^n is of tight total variation on compact time intervals, if we impose a mild technical condition on the tail summability of the c_j, more precisely

$$\begin{cases} \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} c_j \,<\, \infty, & \text{if } 1 < \alpha \,\le\, 2; \\ \sum_{i=1}^{\infty} \Big(\sum_{j=i}^{\infty} c_j \Big)^{\rho} \;<\; \infty \;\text{ for some } 0 < \rho < 1, & \text{if } \alpha = 1. \end{cases} \tag{TC}$$

Remark 4.1. By Fubini's theorem, $\sum_{k=1}^{\infty} k\, c_k < \infty$ is a simple sufficient criterion for (TC) in the case $1 < \alpha \le 2$, and so is $\sum_{k=1}^{\infty} k\, c_k^{\rho} < \infty$ for some $0 < \rho < 1$ in the case $\alpha = 1$.

Lemma 4.2. *Under* (TC)*, the processes U^n in* (4.1) *have good decompositions* (GD).

Proof. As outlined earlier, it suffices to show that the second summand of U^n is of tight total variation on compact time intervals. If $1 < \alpha \le 2$, choose $\rho = 1$ and, in the case $\alpha = 1$, let $0 < \rho < 1$ such that (TC) is satisfied. A simple application of Markov's inequality, monotone convergence and the identical distribution of the θ_k yield

$$\mathbb{P}\Big(\mathrm{TV}_{[0,t]}(U^{n,1}) \,>\, R \Big) \le \mathbb{P}\Bigg(\frac{\psi^{-1}}{n^{\frac{\beta}{\alpha}}} \sum_{k=0}^{\infty} \Big(\sum_{\ell=1}^{N(nt)} c_{k+\ell} \Big) |\theta_{-k}| \,>\, R \Bigg) \le \frac{\tilde{c}\, \psi^{-\rho}}{R^{\rho}\, n^{\frac{\rho\beta}{\alpha}}} \mathbb{E}[|\theta_0|^{\rho}] \to 0$$

as $R \to \infty$ or $n \to \infty$, where $\tilde{c} := \sum_{i=1}^{\infty} (\sum_{j=i}^{\infty} c_j)^{\rho}$, since the ρ-th moment of θ_0 exists as its law is in the domain of attraction of an α-stable distribution with $\rho < \alpha$. ☐

4.1. Direct control of variation

The most direct approach to controlling integrals against the V^n amounts to having the H^n 'tame' the total variation of the V^n. If the H^n are pure jump processes, this becomes particularly simple, leading to Proposition 4.4 below.

Definition 4.3 (Random countable partition)**.** We will call $\pi := \{s_k : k = 0, 1, 2, ...\} \cup \{T\}$ a *(countable) partition of* $[0, T]$ if $\bigcup_{k=0}^{\infty} [s_k, s_{k+1}) = [0, T)$ and $[s_k, s_{k+1}) \cap [s_\ell, s_{\ell+1}) = \emptyset$ for all $k \neq \ell$. In addition, if the s_k are random variables, the partition is said to be *random.* Further, we denote the mesh size of such partition by $|\pi| := \sup_{k=0,1,...} |s_{k+1} - s_k|$.

Towards the next proposition, let $(H^n)_{n\geq 1}$ be a sequence of adapted pure jump càdlàg integrands such that, for every $T > 0$, the set $\mathrm{Disc}_{[0,T]}(H^n) = \{s_k : k = 0, 1, 2, ...\} \cup \{T\}$ is a countable partition of $[0, T]$. Further assume $X^n = U^n + V^n$, V^n are semimartingales adapted to the same filtration as the H^n, where the U^n have (GD), $X^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0, \infty), \rho)$ with $\rho \in \{\mathrm{d_{J1}}, \mathrm{d_{M1}}\}$ and $U^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0, \infty), \mathrm{d_{M1}})$ for some X, and

$$\mathbb{P}\Big(\sum_{k=0}^{\infty} |V^n_{s_{k+1}} - V^n_{s_k}| \; > \; \lambda\Big) \xrightarrow[n\to\infty]{} 0, \quad \text{for each } T, \lambda > 0. \tag{4.3}$$

Proposition 4.4 (Pure jump integrands and control of variation)**.** *Let H^n, X^n be as stated above. If $(H^n, X^n) \Rightarrow (H, X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0, \infty), \mathrm{d_{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0, \infty), \mathrm{d_{M1}})$ and the (H^n, X^n) satisfy* (AVCI)*, then it holds*

$$\Big(X^n, \int_0^\bullet H^n_{s-} \, \mathrm{d}X^n_s\Big) \quad \Longrightarrow \quad \Big(X, \int_0^\bullet H_{s-} \, \mathrm{d}X_s\Big) \qquad \text{on} \qquad (\mathbf{D}_{\mathbb{R}^{2d}}[0, \infty), \mathrm{d_{M1}}).$$

The approach of controlling directly the activity of the V^n is not restricted to pure jump integrands. Indeed, we can ask for the integrands to be Lipschitz continuous in such a way that they exhibit enough inertia to not be able to react in a critical way to changes of the V^n and therefore not 'pick up' too much of the latter's variation through integration.

Definition 4.5 (GD modulo controllable activity)**.** Let $(X^n)_{n\geq 1}$ be a sequence of d-dimensional semimartingales on filtered probability spaces $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, \mathbb{P}^n)$ and let X be defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We say the sequence $(X^n)_{n\geq 1}$ has *good decompositions modulo a weakly asymptotically negligible process of n^γ-controllable activity on an $n^{-\lambda}$-fine partition*—abbreviated as $\mathrm{GD \, mod \, CA}(\gamma, \lambda)$—with $\gamma, \lambda > 0$, if there exist processes $(U^n)_{n\geq 1}$, $(V^n)_{n\geq 1}$ on the same filtered probability spaces as the X^n such that

(i) the U^n are semimartingales having (GD) and $U^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0, \infty), \mathrm{d_{M1}})$;

(ii) the V^n are adapted, càdlàg pure jump processes of finite variation such that, for every $T > 0$ there exist random partitions $\pi^n := \pi^n(\omega)$ of $[0, T]$ with $|\pi^n| \leq n^{-\lambda}$ as well as $\mathrm{Disc}_{[0,T]}(V^n) \subseteq \pi^n$ almost surely, and it holds

$$n^{-\gamma} \sum_{s \, \in \, \pi^n} |V^n_s| \xrightarrow[n\to\infty]{\mathbb{P}^n} 0$$

(iii) $X^n = U^n + V^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0, \infty), \mathrm{d_{M1}})$.

Theorem 4.6 (Lipschitz integrands and GD mod CA)**.** *Let $(X^n)_{n\geq 1}$ be a sequence of d-dimensional semimartingales on filtered probability spaces $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, \mathbb{P}^n)$ which are $\mathrm{GD \, mod \, CA}(\gamma, \tilde{\gamma})$ as in Definition 4.5. Further let $(H^n)_{n\geq 1}$ be a sequence of adapted càdlàg processes on the same filtered probability spaces and suppose that, for every $n \geq 1$ and $T > 0$, $H^n|_{[0,T]}$ is almost surely Lipschitz continuous with Lipschitz constant $C_T \, n^{\tilde{\gamma}-\gamma}$, i.e.*

$$|H^n_t - H^n_s| \; \leq \; C_T \, n^{\tilde{\gamma}-\gamma} \, |t - s|$$

for all $0 \le s \le t \le T$*, where* $C_T > 0$ *only depends on* T*. Suppose further that both* (H^n, U^n) *and* (H^n, X^n) *satisfy* (AVCI)*. If* $(H^n, X^n) \Rightarrow (H, X)$ *on* $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty), \rho)$*, where* $\rho \in \{\mathrm{d}_{\mathrm{J1}}, \mathrm{d}_{\mathrm{M1}}\}$*, then* X *is a semimartingale in the natural filtration generated by* (H, X) *and*

$$\Big(X^n, \int_0^\bullet H^n_{s-}\, \mathrm{d}X^n_s\Big) \quad \Longrightarrow \quad \Big(X, \int_0^\bullet H_{s-}\, \mathrm{d}X_s\Big) \qquad \text{on} \qquad (\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty), \rho).$$

Remark 4.7. By definition of the consecutive increment function $\hat{w}$ in (AVCI), we note that, since $X^n = U^n + V^n$, showing (H^n, X^n) and (H^n, U^n) satisfy (AVCI) is equivalent to establishing (AVCI) for (H^n, X^n) and (H^n, V^n) or (H^n, U^n) and (H^n, V^n).

As one would hope, moving averages and correlated CTRWs constitute a natural class of processes which enjoy the GD mod CA property.

Proposition 4.8 (CTRWs are GD mod CA)**.** *Moving averages* (3.1) *and correlated CTRWs* (3.4) *with* $1 \le \alpha \le 2$ *and* (TC)*, are* GD mod CA(γ, β) *for all* $\gamma > (\beta - \beta/\alpha)$.

4.2. Control through independence

Instead of restricting the class of admissible integrands to such processes which act as a direct control to the activity of the V^n, as pursued in Section 4.1, there is another more probabilistic approach that suggests itself: if we impose that the integrands must not anticipate the 'future' behaviour of the integrator remainders V^n (i.e., adequate independence) and the (conditional) expectation of the latter is suitably centered around zero, then this should offer enough control for a weak continuity result of stochastic integrals. In the sequel, we will provide a precise framework for the implementation of this idea. Throughout, we write $|X|^*_t := \sup_{0 \le s \le t} |X_s|$ for the running supremum of X over the time interval $[0, t]$.

Definition 4.9 (GD modulo processes controllable by independence)**.** Let $(X^n)_{n \ge 1}$ be a sequence of d-dim semimartingales on filtered probability spaces $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, \mathbb{P}^n)$ and let X be a process on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We say that the sequence $(X^n)_{n \ge 1}$ has *good decompositions modulo a weakly asymptotically negligible processes controllable through suitable independence of the integrands*—abbreviated as GD mod CI—if there exist processes $(U^n)_{n \ge 1}$, $(\tilde{U}^n)_{n \ge 1}$, $(V^{n,i})_{n \ge 1}$, $i \ge 1$, which are defined on the same filtered probability spaces as the X^n as well as $f : \mathbb{N} \to (0, \infty)$ and $\lambda, \mu > 1$ such that

- (i) the U^n, $\tilde{U}^n$ are semimartingales having (GD) and $U^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$;
- (ii) the $V^{n,i}$ are pure jump semimartingales, with finitely many jumps on compact time intervals, and let $\sigma^{n,i}_1 \le \sigma^{n,i}_2 \le \ldots$ be stopping times such that $\mathrm{Disc}(V^{n,i}) \subseteq \{\sigma^{n,i}_k : k \ge 1\}$ and denote $\Lambda^{n,i}(t) := \sup\{k \ge 1 : t \le \sigma^{n,i}_k\}$. Further, it holds:
  - (ii.i) for every $T > 0$ and $i \ge 1$, the sequence of random variables $\sup_{i \ge 1} |\Lambda^{n,i}(\bullet)/f(n)|^*_T$ is tight in $\mathbb{R}$;
  - (ii.ii) for every $n, k \ge 1$, $\sum_{i=1}^\infty |V^{n,i}_{\sigma^{n,i}_k}|$ is integrable and we have

$$\mathbb{E}^n\left[V^{n,i}_{\sigma^{n,i}_k} \mid \mathcal{V}_{n,k-1}, \mathcal{V}_{n,k-2}, \ldots, \mathcal{V}_{n,1}\right] \;=\; 0;$$

  for all $i \ge 1$, where $\mathcal{V}_{n,k} := \sigma\Big(V^{n,j}_{\sigma^{n,j}_k} : j \ge 1\Big)$;

(ii.iii) for every $n,k\geq 1$ we have

$$\limsup_{n\geq 1}\sum_{k=0}^{K\,f(n)}\mathbb{E}^n\Big[\,(\tilde{V}_k^{n,>})^{\,\lambda}\Big]\;<\;\infty\qquad\text{and}$$
$$\limsup_{n\geq 1}\sum_{k=0}^{K\,f(n)}\mathbb{E}^n\Big[\,(\tilde{V}_k^{n,\leq})^{\,\mu}\Big]\;<\;\infty,\quad\text{for each } K>0,$$

where $\tilde{V}_k^{n,>}:=\sum_{i=1}^\infty|V^{n,i}_{\sigma_k^{n,i}}|\,\mathbf{1}_{\{|V^{n,i}_{\sigma_k^{n,i}}|>1\}}$ and $\tilde{V}_k^{n,\leq}:=\sum_{i=1}^\infty|V^{n,i}_{\sigma_k^{n,i}}|\,\mathbf{1}_{\{|V^{n,i}_{\sigma_k^{n,i}}|\leq 1\}}$;

(iii) $X^n=U^n+\tilde{U}^n+\sum_{i=1}^\infty V^{n,i}\Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty),\mathrm{d}_{\mathrm{M1}})$.

Moving averages or correlated uncoupled CTRWs with centered innovations and $1<\alpha\leq 2$ provide a natural vast group of processes which are GD mod CI under (TC).

Proposition 4.10 (CTRWs with $1<\alpha\leq 2$ are GD mod CI)**.** *Let X^n be moving averages as in* (3.1) *or a correlated CTRW as in* (3.4) *with $1<\alpha\leq 2$. If $\mathbb{E}[\theta_0]=0$ and* (TC) *hold, then the sequence X^n is* GD mod CI *with*

- *if X^n is a CTRW: $\sigma_k^{n,i}=\sigma_k^n=\sum_{\ell=1}^k J_\ell/n$ and*

$$V_t^{n,i}\;=\;-\,\frac{(\sum_{j=0}^\infty c_j)^{-1}}{n^{\frac{\beta}{\alpha}}}\Big(\sum_{\ell=i}^\infty c_\ell\Big)\,\theta_{N(nt)-i+1}\,\mathbf{1}_{\{N(nt)\geq i\}},\quad\textit{for all } i,k\geq 1;$$

- *if X^n is a moving average: $\sigma_k^{n,i}=\sigma_k^n=k/n$ and*

$$V_t^{n,i}\;=\;-\,\frac{(\sum_{j=0}^\infty c_j)^{-1}}{n^{\frac{\beta}{\alpha}}}\Big(\sum_{\ell=i}^\infty c_\ell\Big)\,\theta_{\lfloor nt\rfloor-i+1}\,\mathbf{1}_{\{\lfloor nt\rfloor\geq i\}},\quad\textit{for all } i,k\geq 1.$$

Theorem 4.11 (Independent integrands and GD mod CI)**.** *Let $(X^n)_{n\geq 1}$ be a sequence of d-dimensional semimartingales on filtered probability spaces $(\Omega^n,\mathcal{F}^n,\mathbb{F}^n,\mathbb{P}^n)$ which are* GD mod CI *and let $(H^n)_{n\geq 1}$ be a sequence of adapted càdlàg processes on the same filtered probability spaces. Suppose that for every $n,k,i\geq 1$ it holds*

$$\sigma\left(H^n_{t\wedge\sigma_\ell^{n,j}}\;:\;t\geq 0,\,j\geq 1,\,\ell\leq k\right)\;\perp\!\!\!\perp\;V^{n,i}_{\sigma_{k-1}^{n,i}}\,,\;V^{n,i}_{\sigma_k^{n,i}}\tag{4.4}$$

and that the pairs (H^n,X^n) satisfy (AVCI) *or the criteria set out in [62, Thm. 4.7].*

If $(H^n,X^n)\Rightarrow(H,X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty),\mathrm{d}_{\mathrm{M1}})\times(\mathbf{D}_{\mathbb{R}^d}[0,\infty),\rho)$, where $\rho\in\{\mathrm{d}_{\mathrm{J1}},\mathrm{d}_{\mathrm{M1}}\}$, then X is a semimartingale in the natural filtration generated by (H,X) and it holds

$$\Big(X^n,\int_0^\bullet H^n_{s-}\;\mathrm{d}X^n_s\Big)\quad\Longrightarrow\quad\Big(X,\int_0^\bullet H_{s-}\;\mathrm{d}X_s\Big)\qquad\textit{on}\qquad(\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty),\rho).$$

Remark 4.12. A close inspection of the proof of Theorem 4.11 reveals that if instead of (ii.ii) in Definition 4.9, we assume just

$$\mathbb{E}^n\Big[V^{n,i,\leq}_{\sigma_k^{n,i}}\mid\mathcal{V}_{n,k-1},\mathcal{V}_{n,k-2},...,\mathcal{V}_{n,1}\Big]\;=\;0\tag{4.5}$$

for every $n,k,i\geq 1$, then, in particular, we gain some degree of freedom towards (ii.iii), where it suffices to replace the first bound by the direct control

$$\lim_{M\to\infty}\;\limsup_{n\to\infty}\;\mathbb{P}^n\bigg(\sum_{k=1}^{Kf(n)}\sum_{i=1}^\infty|V^{n,i,>}_{\sigma_k^{n,i}}|\;\geq\;M\bigg)\;=\;0\tag{4.6}$$

for all $K > 0$. Under this alternative set of assumptions, Theorem 4.11 is clearly still valid. Indeed, in this case, (4.6) fully controls the second probability term in (6.21) and we proceed in full analogy to the proof of Theorem 4.11 for the remaining part.

Corollary 4.13 (CTRWs with $\alpha = 1$ are GD mod CI)**.** *Let X^n be a moving averages as in* (3.1) *or a correlated CTRW as described* (3.4) *with $\alpha = 1$. If the law of θ_0 is symmetric around zero and $\sum_{i=1}^{\infty}\sum_{k=i}^{\infty} c_k < \infty$, then the sequence X^n is* GD mod CI.

4.3. Final results for moving averages and correlated CTRWs

From Theorem 4.6 and Propositions 4.8, we immediately obtain the following result on weak integral convergence of Lipschitz integrands with respect to moving averages and CTRWs with $1 \leq \alpha \leq 2$.

Theorem 4.14 (Weak integral convergence for CTRWs I)**.** *Let X^n be a moving average as in* (3.1) *or an uncoupled correlated CTRWs as in* (3.4) *with $1 \leq \alpha \leq 2$ and* (TC)*. Furthermore, let H^n be processes adapted to the filtration* (3.8) *and let $\gamma \in (0, \beta/\alpha)$ such that for every $T > 0$ there exists $C_T > 0$ with*

$$|H^n_s - H^n_t| \;\leq\; C_T\, n^{\gamma}\, |t-s| \quad \textit{for all} \quad 0 \leq s,t \leq T.$$

If $(H^n, X^n) \Rightarrow (H, X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ with $X = (\sum_{j=0}^{\infty} c_j)Z$ or $X = (\sum_{j=0}^{\infty} c_j)Z(D^{-1}(\bullet))$ and the (H^n, X^n) and (H^n, U^n) satisfy (AVCI)*, where U^n denotes the corresponding moving average or (correlated) CTRW, then X is a semimartingale in the natural filtration generated by (H, X) and it holds*

$$\left(X^n, \int_0^{\bullet} H^n_{s-}\;\mathrm{d}X^n_s\right) \;\Longrightarrow\; \left(X, \int_0^{\bullet} H_{s-}\;\mathrm{d}X_s\right) \qquad \textit{on} \qquad (\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty), \mathrm{d}_{\mathrm{M1}}).$$

Just as we have used Theorem 4.6 and Propositions 4.8 to deduce Theorem 4.14, we can similarly obtain a tailored result for moving averages and CTRWs with $1 \leq \alpha \leq 2$ on behalf of Theorem 4.11 and Proposition 4.10 (respectively Corollary 4.13). Towards a simplification of notation for the next theorem, we will comprehend L_k as the quantity defined before (3.4) if we consider CTRWs and $L_k = k$ if we consider moving averages.

Theorem 4.15 (Weak integral convergence for CTRWs II)**.** *Let X^n be a moving averages as defined in* (3.1) *or an uncoupled correlated CTRW as in* (3.4) *with $1 \leq \alpha \leq 2$, which are adapted to filtrations $\mathbb{F}^n$ and* (TC) *holds. Moreover, assume that we have $\mathbb{E}[\theta_0] = 0$ if $1 < \alpha \leq 2$ and that the law of θ_0 is symmetric if $\alpha = 1$. Furthermore, let H^n be càdlàg processes adapted to the filtration* (3.8) *such that for all $n, k \geq 1$,*

$$H^n_{\bullet \wedge \frac{L_k}{n}} \;\perp\!\!\!\perp\; \sigma\left(\theta_{k-\ell}\,,\, J_{k+i} \;:\; i \geq 1,\, \ell = 0,1,2,...,k \wedge (\mathcal{J}+1)\right) \tag{4.7}$$

where $\mathcal{J} = \sup\{j \geq 1 : c_i = 0 \;\text{ for }\; i > j\}$. If $(H^n, X^n) \Rightarrow (H, X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ with $X = (\sum_{j=0}^{\infty} c_j)Z$ or $X = (\sum_{j=0}^{\infty} c_j)Z(D^{-1}(\bullet))$, then X is a semimartingale in the natural filtration generated by (H, X) and it holds

$$\left(X^n, \int_0^{\bullet} H^n_{s-}\;\mathrm{d}X^n_s\right) \;\Longrightarrow\; \left(X, \int_0^{\bullet} H_{s-}\;\mathrm{d}X_s\right) \qquad \textit{on} \qquad (\mathbf{D}_{\mathbb{R}^{2d}}[0,\infty), \mathrm{d}_{\mathrm{M1}}).$$

Proof. To prove the theorem, according to Theorem 4.11 and Proposition 4.10, we only need to verify (4.4) as well as the alternative conditions to (AVCI) given in [62, Thm. 4.7(a)&(b)]. On behalf of (4.7), it is straightforward to show (4.4) under the choice of $V^{n,i}$ and $\sigma^{n,i}_k$ given in (6.16). With regards to the alternative condition to (AVCI) given in [62, Thm. 4.7], choosing $\sigma^n_k = L_k/n$ for CTRWs and $\sigma^n_k = k/n$ for moving averages, (b) follows immediately from the the

M1 tightness criterion involving the modulus of continuity w'' in [62, Def. A.7] by noting that X^n restarted from σ_k^n is in itself a moving average or uncoupled correlated CTRW, which then is tight on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ according to Section 3.1. Turning to condition (a), it is direct to use [62, Lem. 4.11] together with (4.7) and the fact that X^n is a moving average or an uncoupled CTRW, i.e. the waiting times are independent of the innovations. □

Remark 4.16 (Weaker assumptions on the integrands)**.** It must be pointed out that the convergence $(H^n, X^n) \Rightarrow (H, X)$ on $(\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbf{D}_{\mathbb{R}^d}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ can be significantly weakened and the results in Theorems 3.6, 4.14 and 4.15 (as well as the more general results in Proposition 4.4, Theorem 4.6 and Theorem 4.11) still hold true. We refer to [62, Prop. 3.22] for more details.

Remark 4.17 (Sum of admissible integrands)**.** A close inspection of the proofs of Theorem 4.14 and Theorem 4.15 as well as the proofs of [62, Prop. 5.3 & Thm. 3.6] reveals that if we are given integrands $H^n = H^{1,n} + H^{2,n}$, where $H^{1,n}$ satisfies the assumption of Theorem 4.14 and $H^{2,n}$ meets the assumptions of Theorem 4.15 (where we can replace the convergence condition with the relaxed assumptions described in Remark 4.16), then we obtain the desired integral convergence result from Theorem 4.14/ 4.15 for H^n.

5. Applications to SDE and SDDE models of anomalous diffusion

In Section 1.1, we discussed the connection between fractional Fokker–Planck–Kolmogorov equations and SDEs driven by time-changed Lévy processes. As a concrete example, consider a spherically symmetric Lévy process Z (with characteristic exponent $\Psi(\xi) = -|\xi|^2$) time-changed by the inverse of a β-stable subordinator D. Under suitable assumptions, an SDE of the form

$$\mathrm{d}X_t = \sigma(X_{t-})\,\mathrm{d}Z_{D_t^{-1}}, \quad X_0 = x,$$

will then have transition densities $p(t,x)$ governed by

$$D_t^{\beta} p(t,x) = -\kappa(\alpha)(-\Delta)^{\frac{\alpha}{2}}\big(\sigma(x)^{\alpha} p(t,x)\big), \quad p(0,x) = \delta_x,$$

where $-(-\Delta)^{\frac{\alpha}{2}}$ is a fractional Laplacian and $\kappa(\alpha)$ is the appropriate diffusivity constant. In order to connect the continuum formulations with CTRW driven models, a natural approach is to exploit the results on stochastic integral convergence established earlier in the paper, similarly to the classical results of [60] and [37, Sect. 5]. Beyond establishing a rigorous theoretical link to CTRW formulations, this also provides a natural numerical scheme for the simulation of SDEs driven by time-changed Lévy processes and, consequently, the associated fractional Fokker–Planck–Kolmogorov equations through a Monte Carlo procedure.

While Section 5.1 implements the above, Section 5.2 proceeds to consider SDDEs. Specifically, our aim is to illustrate that, for such equations, it is possible to have functional limit theorems on Skorokhod space even if the driving CTRWs are strictly M1 convergent. This is achieved by exploiting the framework of Section 4.2. To keep the analysis manageable, we focus on moving averages rather than more general correlated CTRWs. In turn, the limiting equations are driven by a standalone Lévy process rather than a time-changed one.

5.1. Functional limit theorems for SDEs driven by CTRWs

A series of papers [29, 31, 39] have investigated weak and strong approximation schemes for SDEs of the general form

$$\begin{cases} \mathrm{d}X_t = \mu(D_t^{-1}, X_t)_-\,\mathrm{d}D_t^{-1} + \sigma(D_t^{-1}, X_t)_-\,\mathrm{d}B_{D_t^{-1}} \\ X_0 = x, \end{cases} \tag{5.1}$$

where B is a Brownian motion B and D is a subordinator. Motivated by stable limit theorems, we focus on stable subordinators for concreteness, but the main thing is simply that the subordinator is strictly increasing. A recent work [30] extends the results of [31] to allow for a traditional drift term, but only of the specific form $b(D_t^{-1})\,\mathrm{d}t$. In general, $\mathrm{d}t$ terms are problematic because they break the useful duality with non-time-changed SDEs explored in [36].

As discussed in [63, Ch. 6.6], the inverse subordinator and time-changed driver in (5.1) are not Markovian and do not enjoy independent or stationary increments, so one cannot rely directly on the usual tools for classical SDEs such as the Euler method. On the other hand, CTRW approximations present themselves as a natural alternative (see e.g. [44, Ch. 5] for the simulation of CTRWs and their use in simulating time-changed Lévy processes).

Section 4 of the aforementioned work [36] studies theoretical properties, such as existence and uniqueness, for the more general class of SDEs

$$\begin{cases} \mathrm{d}X_t = b(t, D_t^{-1}, X_t)_-\,\mathrm{d}t \;+\; \mu(t, D_t^{-1}, X_t)_-\,\mathrm{d}D_t^{-1} \;+\; \sigma(t, D_t^{-1}, X_t)_-\,\mathrm{d}Z_{D_t^{-1}} \\ X_0 = x, \end{cases} \tag{S}$$

for suitable Z and D. As we have done throughout, here we take Z to be an α-stable Lévy process and D^{-1} to be the inverse of a β-stable subordinator, so that our analysis aligns with the stable limit theory recalled in Section 3.1. We are then interested in connecting (S) with the approximating SDEs

$$\begin{cases} \mathrm{d}X_t^n = b(t, D_t^n, X_t^n)_-\,\mathrm{d}t \;+\; \mu(t, D_t^n, X_t^n)_-\,\mathrm{d}D_t^n \;+\; \sigma(t, D_t^n, X_t^n)_-\,\mathrm{d}Z_t^n \\ X_0 = x, \end{cases} \tag{S_n}$$

where $D^n := n^{-\beta}N(n\bullet)$ and where the Z^n denote uncorrelated uncoupled CTRWs defined as in (3.4) with $c_j = 0$ for all $j \geq 1$. Of course, one could also consider weakly convergent initial conditions independent of the other stochastic inputs.

In terms of structural assumptions, we take the functions $b, \mu, \sigma : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$ to be continuous as well as satisfying, for all $T, R > 0$, a strictly sublinear growth condition

$$\sup_{0\leq t\leq T} \sup_{|\tilde{y}|\leq R} \left(|b(t,\tilde{y},y)| \vee |\mu(t,\tilde{y},y)| \vee |\sigma(t,\tilde{y},y)|\right) \;\leq\; K\,|y|^p + C, \tag{5.2}$$

where the constants $K, C > 0$ and the exponent $p \in (0,1)$ may depend on T, R. Here we have used the usual notation $a \vee b := \max\{a, b\}$. Imposing strict sublinearity for the growth condition in the space variable serves as a tool to prove suitable tightness of the solutions X^n on the Skorokhod space. In typical SDE configurations, one would often allow for linear growth by resorting to Gronwall type arguments. For example, one could aim to rely on the general stochastic version of Gronwall's lemma in [18, Thm. 1.2], but the lack in predictability and integrability of the integrators for classical upper bounds does not permit a direct application.

With the above assumptions, we obtain the following functional central limit theorem for SDEs driven by CTRWs. It gives existence of a solution to the SDE (S) and shows that the limits points of (S_n) are supported on the solution set. The proof is given in Section 6.4.

Theorem 5.1 (Convergence of SDEs driven by CTRWs)**.** *Any subsequence of the solutions $(X^n)_{n\geq 1}$ to* (S_n) *has a further subsequence converging weakly on* $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{J1}})$ *to a solution X of* (S)*. If there is uniqueness in law for* (S)*, then $(X^n)_{n\geq 1}$ itself converges weakly to this unique limit.*

5.2. Stochastic delay differential equations driven by moving averages

For a given delay parameter $r > 0$, consider the stochastic delay differential equation

$$\begin{cases} \mathrm{d}X_t = b(t, X_{t-r})_-\,\mathrm{d}t + \sigma(t, X_{t-r})_-\,\mathrm{d}Z(t) \\ X_s = \eta_s, \quad s \in [-r, 0] \end{cases} \tag{$\check{\mathrm{S}}$}$$

whose driver, Z, is a stable Lévy process. Such a model can capture phenomena driven by anomalous diffusion in the space variable, where the rates of change for the drift and diffusion depend on the state of the system some time into the past.

When Z is a Brownian motion, discrete approximation schemes for delayed systems such as (Š) have been examined in [65]. We also note that [24] studies convergence of stochastic integrals for SDDEs, again in the Brownian setting, but there the focus is on characterising the small delay limit. Here we are interested instead in the stability of (Š) with respect to the driver Z when this arises as the weak scaling limit of moving averages in the M1 topology. To this end, we shall rely on Theorem 4.15. One could also consider the corresponding systems driven by time-changed Lévy processes, and hence study the weak convergence in terms of general CTRWs. However, in order to keep the treatment succinct, we prefer to illustrate the procedure with the simpler moving averages as drivers.

Consider the SDDE from (Š) with initial condition $\eta \in \mathbf{D}_{\mathbb{R}}[-r,0]$ and continuous $b,\sigma : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ satisfying for any $T>0$ the boundedness condition

$$\sup_{0\le t\le T}\ \sup_{x\in\mathbb{R}} |b(t,x)| \vee |\sigma(t,x)| \ <\ \infty. \tag{5.3}$$

Let $(X^n)_{n\ge 1}$ be a sequence of solutions to the SDDEs

$$\begin{cases} \mathrm{d}X^n_t &= \ b(t, X^n_{t-r})_-\ \mathrm{d}t \ +\ c^{-1}\,\sigma(t, X^n_{t-r})_-\ \mathrm{d}Z^n_t \\ X^n_s &= \ \eta(s), \quad s\in[-r,0] \end{cases} \tag{Š$_n$}$$

where $c := \sum_{k=0}^{\mathcal{J}} c_k$ and the Z^n are moving averages as defined in (3.1) with $c_j = 0$ for all $j > \mathcal{J}$ and some $\mathcal{J}\ge 1$ as well as $\mathbb{E}[\theta_0] = 0$ if $1<\alpha\le 2$ and the law of θ_0 being symmetric around zero if $\alpha = 1$. Then, we know from Section 3.1, that $c^{-1}Z^n \Rightarrow Z$ on the M1 Skorokhod space. Notice that the Z^n are pure jump processes with finitely many jumps on compact time intervals and therefore solutions to (Š$_n$) do not only exist but are also explicit. We are interested in the weak convergence of solutions of (Š$_n$) to a solution of (Š) on on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$.

To prove this result, we must first show some form of relative compactness for the sequence $(X^n)_{n\ge 1}$ of solutions to (Š$_n$). Unfortunately, the development of general convenient criteria established on the M1 space fell short of the one on its J1 counterpart (e.g. Aldous' criteria [8, Thm. 16.10] or Rebolledo's criteria [56]). Still, a useful condition for relative compactness on the M1 space has been given by Avram & Taqqu [2, Thm. 1]. This condition allows it to reduce the classical criteria involving the modulus of continuity w'' (see e.g. [62, Def. A.7]) to a version based on intervals with respect to fixed times if one can ensure a certain uniform bound which is essentially scaled by the length of the intervals, and we will be able to use this in order to show the M1 relative compactness of the X^n.

Once this has been shown, it is straightforward to deduce that any subsequence has a further subsequence which converges to some limit (a priori dependent on the specific subsequence) and such that the integrands on the right-hand side of (Š$_n$) satisfy the weaker integrand conditions mentioned in Remark 4.16. Clearly, (AVCI) is trivially satisfied for the drift integral in (Š$_n$) due to Proposition 2.3, and thus these converge suitably according to [62, Prop. 3.22]. On the other hand, if the integrators and integrands of the diffusion integral meet assumption (4.7) of Theorem 4.15, then Remark 4.16 yields additionally the desired convergence of these integrals. Since the limit of the drift integrals is continuous, we ultimately deduce the desired convergence of both sides of (Š$_n$) on the M1 space. The proof is given in Section 6.4.

Theorem 5.2 (Convergence of SDDEs driven by moving averages)**.** *For any subsequence of the solutions X^n of (Š$_n$), there is a further subsequence converging weakly on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ to a solution X of (Š). If there is uniqueness in law for (Š), then $X^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$, where X is the unique solution.*

One can easily extend the previous result to SDDEs of the type

$$\begin{cases} \mathrm{d}X_t &= b(t, X_{t-r})_- \ \mathrm{d}t \ + \ [\sigma(t, X_{t-r})_- \ + \ \tilde{\sigma}(t, X_{[t-r,t]})_-] \ \mathrm{d}Z_t \\ X_u &= \eta(u), \quad u \in [-r, 0] \end{cases} \tag{$\tilde{\mathrm{S}}$}$$

where Z, b, σ and η are defined as before in ($\check{\mathrm{S}}$) and $X_{[t-r,t]}$ denotes the path segment $(X_s)_{t-r \le s \le t}$. Here $\tilde{\sigma}$ is of the form

$$\tilde{\sigma}(t, \ X_{[t-r,t]}) \ = \ \int_{t-r}^{t} \Phi(t, s, X_s) \ \mathrm{d}s$$

where $\Phi : [0, \infty) \times [-r, \infty) \times \mathbb{R} \to \mathbb{R}$ is such that for every $T > 0$

$$\sup_{t \in [0,T]} \ \sup_{s \in [t-r,t]} \ \sup_{x \in \mathbb{R}} |\Phi(t, s, x)| \ < \ \infty.$$

We let Φ be continuous in the third component (i.e., $x \mapsto \Phi(t, s, x)$ is continuous for fixed t, s) and Lipschitz on compacts in the first component: for all $T > 0$ there is $L_T > 0$ such that

$$|\Phi(t, s, x) - \Phi(\tilde{t}, s, x)| \ \le \ L_T \ |t - \tilde{t}|$$

for any $0 \le t \le \tilde{t} \le T$, $x \in \mathbb{R}$ and $\tilde{t} - r \le s \le t$. Obviously, the map $x \mapsto \tilde{\sigma}(\bullet, x_{[\bullet-r,\bullet]})$ is continuous from $(\mathbf{D}_{\mathbb{R}}[0, \infty), \mathrm{d}_{\mathrm{M1}})$ into $(\mathbf{C}_{\mathbb{R}}[0, \infty), |\cdot|_{\infty}^{*})$. Indeed, relative compactness follows from the Arzelà-Ascoli theorem, and pointwise convergence can easily be shown by dominated convergence, the continuity of Φ in the third component and the fact that $x_n \to x$ in $(\mathbf{D}_{\mathbb{R}}[0, \infty), \mathrm{d}_{\mathrm{M1}})$ implies in particular $x_n(s) \to x(s)$ for all but countably many $s \in [0, \infty)$. An example of such a $\tilde{\sigma}$ could be the convolution with a Lipschitz kernel $\rho : [-r, 0] \to \mathbb{R}$, that is $\Phi(t, s, X_s) := \rho(t - s) \, f(X_s)$ for $t - r \le s \le t$, where f is bounded and continuous.

Corollary 5.3 (Another class of SDDEs)**.** *Denote by* ($\tilde{\mathrm{S}}^{\mathrm{n}}$) *the approximating SDDEs for* ($\tilde{\mathrm{S}}$) *in analogy to how* ($\check{\mathrm{S}}_n$) *and* ($\check{\mathrm{S}}$) *relate. For any subsequence of the solutions X^n of* ($\tilde{\mathrm{S}}^{\mathrm{n}}$)*, there is a further subsequence which converges weakly on* $(\mathbf{D}_{\mathbb{R}}[0, \infty), \, \mathrm{d}_{\mathrm{M1}})$ *to a solution X of* ($\tilde{\mathrm{S}}$)*. If* ($\tilde{\mathrm{S}}$) *is unique in law, then $X^n \Rightarrow X$ on* $(\mathbf{D}_{\mathbb{R}}[0, \infty), \, \mathrm{d}_{\mathrm{M1}})$*, where X is the unique solution.*

6. Proofs of results from Sections 3, 4, and 5

6.1. Proofs pertaining to Section 3.3

Lemma 6.1. *Zero-order moving averages and uncoupled uncorrelated CTRWs (i.e.,* (3.1) *or* (3.4) *with $c_j = 0$ for all $j \ge 1$) are of tight total variation on compacts.*

Proof of Lemma 6.1. We are only going to prove the lemma for CTRWs. The proof for moving averages can be conducted in a very similar way. Let X^n be defined as in (3.4) with $c_0 = 1$ and $c_j = 0$ for all $j \ge 1$. As a pure jump process, the total variation of X^n is

$$\mathrm{TV}_{[0,t]}(X^n) \ = \ \sum_{k=1}^{N(nt)} |\zeta_k / n^{\frac{\beta}{\alpha}}| \, .$$

Fix $t > 0$ and let $\varepsilon > 0$. Given the tightness from (3.5), there exists $K_{t,\varepsilon} > 0$ such that

$$\mathbb{P}\Big(\sup_{s \in [0,t]} \ n^{-\beta} N(ns) \ > \ K_{t,\varepsilon}\Big) \ = \ \mathbb{P}\Big(N(nt) \ > \ n^{\beta} K_{t,\varepsilon}\Big) \ \le \ \frac{\varepsilon}{3}.$$

Hence,

$$\mathbb{P}\Big(\mathrm{TV}_{[0,t]}(X^n) \ > \ C_{t,\varepsilon} \Big) \ \le \ \frac{\varepsilon}{3} \ + \ \mathbb{P}\Big(n^{-\frac{\beta}{\alpha}} \sum_{k=1}^{\lfloor n^\beta K_{t,\varepsilon} \rfloor} |\zeta_k| \ > \ C_{t,\varepsilon} \Big)$$

$$\le \ \frac{\varepsilon}{3} \ + \ \mathbb{P}\Big(\sum_{k=1}^{\lfloor n^\beta K_{t,\varepsilon} \rfloor} |\zeta_k| \ > \ C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \ , \ \max_{k=1,\ldots,\lfloor n^\beta K_{t,\varepsilon} \rfloor} |\zeta_k| \le C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \Big) \tag{6.1}$$

$$+ \ \mathbb{P}\Big(\max_{k=1,\ldots,\lfloor n^\beta K_{t,\varepsilon} \rfloor} |\zeta_k| \ > \ C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \Big) \tag{6.2}$$

and we are going to determine a suitable $C_{t,\varepsilon} > 0$ such that (6.1) and (6.2) are each bounded by $\varepsilon/3$, then implying that the sequence X^n has tight total variation on $[0, t]$. Firstly,

$$(6.2) \ \le \ K_{t,\varepsilon}\, n^\beta \ \mathbb{P}\left(|\zeta_1| \ > \ C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \right) \ = \ \frac{K_{t,\varepsilon}}{C_{t,\varepsilon}^\alpha} \left[C_{t,\varepsilon}^\alpha\, n^\beta \ \mathbb{P}\left(|\zeta_1| \ > \ C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \right) \right].$$

Due to the alternative characterisation of the domain of attraction of an α-stable law (see [7, (1.1)] or [16, p. 312/313]), the tail probabilities of ζ_1 behave with regularity $\mathbb{P}(|\zeta_1| > x) \sim x^{-\alpha}$. Hence, we can choose $n_0 \ge 1$ such that for all $n \ge n_0$ it holds

$$C_{t,\varepsilon}^\alpha\, n^\beta \ \mathbb{P}\left(|\zeta_1| \ > \ C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}} \right) \ \le \ 2$$

implying that for all $n \ge n_0$,

$$(6.2) \ \le \ \frac{2K_{t,\varepsilon}}{C_{t,\varepsilon}^\alpha}\,. \tag{6.3}$$

Concerning (6.1), we will make use of the Fuk–Nagaev inequality (see [7, Theorem 5.1(i)] and take into account the remark in [7, p.12] that the inequality only requires the bound $\mathbb{P}(|\zeta_1| > x) \le cx^{-\alpha}$) with $X_k = \zeta_k$ and $x = y = C_{t,\varepsilon} n^{\frac{\beta}{\alpha}}$ to obtain

$$(6.1) \ \le \ e\lambda^{-1} \lfloor n^\beta K_{t,\varepsilon} \rfloor \, (C_{t,\varepsilon}\, n^{\frac{\beta}{\alpha}})^{-\alpha} \ \le \ e\lambda^{-1}\, \frac{K_{t,\varepsilon}}{C_{t,\varepsilon}^\alpha} \tag{6.4}$$

for some constant $\lambda > 0$ independent of all other quantities. Therefore, choosing $n \ge n_0$ and $C_{t,\varepsilon} > 0$ large enough, we are able to bound (6.1) and (6.2) by $\varepsilon/3$ on behalf of (6.3) and (6.4). Since all X^n are of local finite variation, the family $\{\mathrm{TV}_{[0,t]}(X^n) \ : \ n < n_0\}$ is obviously tight and we deduce the existence of $C_{t,\varepsilon} > 0$ such that $\mathbb{P}(\mathrm{TV}_{[0,t]}(X^n) \ > \ C_{t,\varepsilon}) \le \varepsilon$. □

Proof of Proposition 3.8. Consider the decompositions (4.1) and (4.2). Note that

$$\mathrm{TV}_{[0,t]}(U^n) \ \le \ \mathrm{TV}_{[0,t]}(U^{n,1}) \ + \ n^{-\beta/\alpha} \sum_{k=0}^{\infty} \Big(\sum_{\ell=1}^{N(nt)} c_{k+\ell} \Big) \, |\theta_{-k}|$$

and $\mathrm{TV}_{[0,t]}(V^n) \le \mathrm{TV}_{[0,t]}(U^{n,1})$, where $U_t^{n,1} = n^{-\beta/\alpha} \sum_{k=1}^{N(nt)} \theta_k$ denotes the uncorrelated CTRW. Since $\mathrm{TV}_{[0,t]}(U^{n,1})$ is tight by Lemma 6.1, it only remains to show that $n^{-\beta/\alpha} \sum_{k=0}^{\infty} (\sum_{\ell=1}^{N(nt)} c_{k+\ell}) |\theta_{-k}|$ is tight on compact time intervals: the case of only finitely many $c_j \neq 0$ is straightforward. Assume now that there are infinitely many $c_j \neq 0$ and for the sake of simplicity that the sequence $(c_j)_{j \ge 1}$ is non-increasing. Let $t \ge 0$, $\varepsilon > 0$ as well as $0 < \rho < \alpha < 1$ such that $\sum_{j=0}^{\infty} c_j^\rho < \infty$, and note that then $\theta_0 \in L^\rho(\Omega, \mathcal{F}, \mathbb{P})$. Therefore, since $\alpha < 1$ and hence $\beta - \beta/\alpha < 0$, we have

$$\begin{aligned}
&\mathbb{P}\Big(n^{-\beta/\alpha} \sum_{k=0}^{\infty} \Big(\sum_{\ell=1}^{N(nt)} c_{k+\ell}\Big)\, |\theta_{-k}| \;>\; R\Big) \\
&\leq\; \frac{\varepsilon}{2} \;+\; \mathbb{P}\Big(n^{-\beta/\alpha} \sum_{k=0}^{\infty} \Big(\sum_{\ell=1}^{K_\varepsilon n^\beta} c_{k+\ell}\Big)\, |\theta_{-k}| \;>\; R\Big) \\
&\leq\; \frac{\varepsilon}{2} \;+\; \mathbb{P}\Big(\sum_{k=0}^{\infty} c_k\, |\theta_{-k}| \;>\; n^{\beta/\alpha-\beta} \frac{R}{K_\varepsilon}\Big) \\
&\leq\; \frac{\varepsilon}{2} \;+\; \Big(\frac{K_\varepsilon}{R}\Big)^\rho n^{(\beta-\beta/\alpha)\rho}\, \mathbb{E}\Big[\Big(\sum_{k=0}^{\infty} c_k\, |\theta_{-k}|\Big)^\rho\Big] \\
&\leq\; \frac{\varepsilon}{2} \;+\; \Big(\frac{K_\varepsilon}{R}\Big)^\rho n^{(\beta-\beta/\alpha)\rho}\, \mathbb{E}[|\theta_0|^\rho] \sum_{k=0}^{\infty} c_k^\rho \;\xrightarrow[n\to\infty]{}\; \frac{\varepsilon}{2}
\end{aligned}$$

by monotone convergence and the identical distribution of the θ_k, where $K_\varepsilon > 0$ is such that $\sup_{n\geq 1} \mathbb{P}(n^{-\beta} N(nt) > K_\varepsilon) \leq \varepsilon/2$. Thus, we have obtained the desired tightness. $\square$

6.2. Proofs pertaining to Section 4.1

Proof of Theorem 4.6. We proceed by mimicing the proof of [62, Thm. 3.6] and refer to the latter for definitions of the quantities used subsequently. Note that it suffices to show that X is a semimartingale with respect to the natural filtration generated by the pair (H, X) and we obtain a convenient bound for the quantity (T4) defined in the proof of the proof of [62, Thm. 3.6]. Denote U^n and V^n the decomposition processes of X^n according to Definition 4.5. Given that $(H^n, X^n) \Rightarrow_{f.d.d.} (H, X)$ on a co-countable set of times due to the convergence on the Skorokhod space, also $(H^n, U^n) \Rightarrow_{f.d.d.} (H, X)$, and $(|U^n|^*_T)_{n\geq 1}$ is tight as a result of the tightness of the U^n on the Skorokhod space (Definition 4.5(i)), hence [62, Prop. 3.5] yields that X is a semimartingale with respect to the natural filtration generated by (H, X) since the U^n have (GD). Now, in order to derive the desired bound on (T4) from the proof of [62, Thm. 3.6], note that it is enough to show

$$\lim_{\varepsilon\to 0}\; \lim_{m\to\infty}\; \limsup_{n\to\infty}\; \mathbb{E}^n\Big[\,\Big|\int_0^\bullet (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-})\; \mathrm{d}X^n_s\,\Big|^*_T \;\wedge\; 1\Big] \;=\; 0$$

for each $T > 0$. Since $X^n = U^n + V^n$ and the U^n have (GD), it only remains to bound

$$\lim_{\varepsilon\to 0}\; \lim_{m\to\infty}\; \limsup_{n\to\infty}\; \mathbb{E}^n\Big[\,\Big|\int_0^\bullet (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-})\; \mathrm{d}V^n_s\,\Big|^*_T \;\wedge\; 1\Big] \;=\; 0 \tag{6.5}$$

as for $\mathbb{E}^n[\,|\int_0^\bullet (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-})\, \mathrm{d}U^n_s\,|^*_T \wedge 1]$ we can proceed as in the proof of [62, Thm. 3.6] (and just replace X^n by U^n in this proof). Fix $T > 0$ and let $\pi^n = \{s^n_k : k = 0, 1, ...\}$ be the respective partitions of $[0, T]$ from Definition 4.5(ii). For simplicity of notation, we will denote $s_k := s^n_k$. First note that for every $n \in \mathbb{N}$ and any càdlàg process G of finite variation, integration by parts for Lebesgue-Stieltjes integration yields

$$\int_0^t G_{s-}\; \mathrm{d}V^n_s \;=\; V^n_t G_t \;-\; V^n_0 G_0 \;-\; \sum_{k=0}^{\infty} V^n_{s_{2k}}\; [G_{s_{2k+1}\wedge t} - G_{s_{2k}\wedge t}]$$

for all $t \geq 0$. Now, choosing $G := H^n - H^{n\,|\,m,\varepsilon}$, where we recall that by definition $|H^n - H^{n\,|\,m,\varepsilon}|^*_T \leq \varepsilon$, we obtain

$$
\begin{aligned}
A_n \;&:=\; \Big| \int_0^\bullet (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-}) \,\mathrm{d}V^n_s \Big|^*_T \\
&\le\; 2\varepsilon\, |V^n|^*_T \;-\; \sum_{k=0}^{\infty} |V^n_{s_k}| \; |(H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}) \;-\; (H^n_{s_{k+1}} - H^{n\,|\,m,\varepsilon}_{s_{k+1}})|.
\end{aligned}
$$

Since $H^{n\,|\,m,\varepsilon}$ by definition is constant between the partition points of $\rho^{m,\varepsilon}(H^n)$, we have

$$
\begin{aligned}
A_n \;&\le\; 2\varepsilon\, |V^n|^*_T \;+\; \sum_{\substack{k\ge 0 \\ \nexists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |(H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}) \;-\; (H^n_{s_{k+1}} - H^{n\,|\,m,\varepsilon}_{s_{k+1}})| \\
&\qquad +\; \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |(H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}) \;-\; (H^n_{s_{k+1}} - H^{n\,|\,m,\varepsilon}_{s_{k+1}})| \\
&\le\; 2\varepsilon\, |V^n|^*_T \;+\; \sum_{\substack{k\ge 0 \\ \nexists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |H^n_{s_k} \;-\; H^n_{s_{k+1}}| \\
&\qquad +\; \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |(H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}) \;-\; (H^n_{s_{k+1}} - H^{n\,|\,m,\varepsilon}_{s_{k+1}})| \\
&\le\; 2\varepsilon\, |V^n|^*_T \;+\; 2C_T\, n^{-\gamma} \sum_{t\,\in\,\zeta^n} |V^n_t| \;+\; \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}| \qquad (6.6)
\end{aligned}
$$

where we have made use of the Lipschitz continuity of the H^n and $s_{k+1} - s_k \le |\zeta^n| \le n^{-\tilde{\gamma}}$. Since $V^n = X^n - U^n$, $(|V^n|^*_T)_{n\ge 1}$ is tight in $\mathbb{R}$ as a result of $(|U^n|^*_T)_{n\ge 1}, (|X^n|^*_T)_{n\ge 1}$ being tight in $\mathbb{R}$ (due to the tightness of U^n, X^n on the Skorokhod space), and by Definition 4.5(ii), for the first two terms of (6.6) it holds

$$
\lim_{\varepsilon\to 0}\; \limsup_{n\to\infty}\; \mathbb{P}^n\Big(\varepsilon\, |V^n|^*_T \;+\; 2C_T\, n^{-\gamma} \sum_{t\,\in\,\pi^n} |V^n_t| \;>\; \gamma\Big) \;=\; 0 \qquad (6.7)
$$

for all $\gamma > 0$. Considering the last term of (6.6), we continue bounding it by

$$
\begin{aligned}
&\sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_{k+1}}| \\
&\le \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\setminus\rho^m\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}| \;+\; \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^m\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}| \; |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}| \\
&\le \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\setminus\rho^m\,:\, s_k < p \le s_{k+1}}} \big(|V^n_{s_k} - V^n_{\min\,\{t\,\in\,\rho^m\,:\, t\,\ge\, s_k\}}| + |V^n_{\min\,\{t\,\in\,\rho^m\,:\, t\,\ge\, s_k\}}|\big) \; |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}| \\
&\qquad +\; \varepsilon \sum_{\substack{k\ge 0 \\ \exists\, p\,\in\,\rho^m\,:\, s_k < p \le s_{k+1}}} |V^n_{s_k}|\,. \qquad (6.8)
\end{aligned}
$$

Since $V^n = X^n - U^n$ and both X^n, U^n converge weakly in M1 to X (and therefore their finite-dimensional distributions converge to those of X along a co-countable subset), there exists a dense subset $D \subseteq [0,T]$ such that the finite-dimensional distributions of V^n converge to 0 along D. Let π^J, $J \ge 2$ be such that $\pi^J = \{r_{J,i} : 0 = r_{J,1} < r_{J,2} < \ldots < r_{J,J} = T,\, r_{J,i} \in D\}$ such that $|\pi^J| := \max_{2\le i\le J} |r_{J,i+1} - r_{J,i}| \to 0$ as $J \to \infty$. Recall that the number of partition points of the deterministic partition ρ^m is bounded by some constant k_m only depending on m and T. Then, we continue by estimating

$$\begin{aligned}
\sum_{\substack{k\ge 0\\ \exists\, p\,\in\,\rho^m\,:\,s_k<p\le s_{k+1}}} |V^n_{s_k}| \;\le\; & \sum_{\substack{k\ge 0\\ \exists\, p\,\in\,\rho^m\,:\,s_k<p\le s_{k+1}}} \Big(\big(|V^n_{s_k} - V^n_{\max\{t\in\pi^J\,:\,t\le s_k\}}| \,\vee\, |V^n_{s_k} - V^n_{\min\{t\in\pi^J\,:\,t\ge s_k\}}|\big) \\
& \qquad\qquad + \big(|V^n_{\max\{t\in\pi^J\,:\,t\le s_k\}}| \,\vee\, |V^n_{\min\{t\in\pi^J\,:\,t\ge s_k\}}|\big)\Big) \\
\le\; & k_m \Big(w''(V^n, |\pi^J|) \;+\; 3 \max_{t\,\in\,\pi^J} |V^n_t| \Big) \\
\le\; & k_m \Big(w''(U^n, |\pi^J|) \;+\; w''(X^n, |\pi^J|) \;+\; 3 \max_{t\,\in\,\pi^J} |V^n_t| \Big)
\end{aligned} \tag{6.9}$$

for all $J \in \mathbb{N}$. In particular, we have used for the second inequality that, by the triangle inequality, $|V_{r_2}-V_{r_1}|\vee|V_{r_2}-V_{r_3}| \le |V_{r_3}-V_{r_2}|+w''(V^n|_{[0,T]},\theta)$ for $0 \le r_1 < r_2 < r_3 \le T$, $|r_3-r_1| \le \theta$, where w'' denotes the modulus of continuity (as e.g. defined in [62, Def. A7]). Moreover, for the third inequality, we have employed that $V^n = X^n - U^n$ and the property $w''(X^n - U^n, \theta) \le w''(X^n, \theta) + w''(U^n, \theta)$. Further, the number of partition points of $\rho^{m,\varepsilon}(H^n) \setminus \rho^m$ is bounded by $N^T_\varepsilon(H^n)$, where N^T_ε is the maximal number of ε-increments of H^n on $[0, T]$ as defined in [62, Eq. (3.9)]. On behalf of the bound (6.9), (6.8) can be further estimated by

$$\begin{aligned}
& \sum_{\substack{k\ge 0\\ \exists\, p\,\in\,\rho^{m,\varepsilon}(H^n)\,:\,s_k<p\le s_{k+1}}} |V^n_{s_k}| \;\, |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}| \\
\le\;& N^T_\varepsilon(H^n) \sup_{\substack{0\,\le\, s\,\le\, u\,\le\, r\,\le\, T\\ |s-r|\,\le\, 2|\rho^m|}} |H^n_s - H^n_u| \;\, |V^n_u - V^n_r| \;+\; 2\varepsilon\, N^T_\varepsilon(H^n) \sum_{p=1}^{k_m} |V^n_{t^m_p}| \\
& \quad + \;\varepsilon\, k_m \max_{t\,\in\,\pi^J} |V^n_t| \;+\; \varepsilon\, k_m \Big(w''(X^n|_{[0,T]},\, |\pi^J|) \;+\; w''(U^n|_{[0,T]},\, |\pi^J|) \Big) \\
\le\;& 2N^T_\varepsilon(H^n)\, (|H^n|^*_T \,+\, |V^n|^*_T)\; \hat{w}^T_{2|\rho^m|}(H^n,\, V^n) \\
& \qquad\qquad + \;\varepsilon\, N^T_\varepsilon(H^n) \sum_{p=1}^{k_m} |V^n_{t^m_p}| \;+\; \varepsilon\, k_m \max_{t\,\in\,\pi^J} |V^n_t| \\
& \qquad\qquad + \;\varepsilon\, k_m \Big(w''(X^n|_{[0,T]},\, |\pi^J|) \;+\; w''(U^n|_{[0,T]},\, |\pi^J|) \Big)
\end{aligned} \tag{6.10}$$

where the t^m_p denote the partition points of the partition ρ^m (which without loss of generality can be assumed to be in the dense set D). Furthermore, $\hat{w}$ is the consecutive increment function from Definition 2.2. The last inequality is based on the fact that $|xy| \le (|x| \vee |y|)(|x| \wedge |y|)$. Combining (6.6) with (6.10), for all $J \ge 1$ we obtain

$$\begin{aligned}
\Big| \int_0^\bullet (H^n_{r-} - H^{n\,|\,m,\varepsilon}_{r-})\; \mathrm{d}V^n_r \Big|^*_T \;\le\; & 2\varepsilon\, |V^n|^*_T \;+\; 2C\, n^{-\gamma} \sum_{t\,\in\,\zeta^n} |V^n_t| \\
& + \;2N^T_\varepsilon(H^n)\, (|H^n|^*_T \,+\, |V^n|^*_T)\; \hat{w}^T_{|\rho^m|}(H^n,\, V^n) \\
& + \;2\varepsilon\, N^T_\varepsilon(H^n) \sum_{p=1}^{k_m} |V^n_{t^m_p}| \;+\; \varepsilon\, k_m \max_{t\,\in\,\pi^J} |V^n_t| \\
& + \;\varepsilon\, k_m \Big(w''(X^n|_{[0,T]},\, |\pi^J|) \;+\; w''(U^n|_{[0,T]},\, |\pi^J|) \Big).
\end{aligned} \tag{6.11}$$

We know that $|H^n|^*_T$ is tight since the H^n are tight in M1. The same holds for $|V^n|^*_T$ as $V^n = X^n - U^n$ and both X^n and U^n are tight in M1. In addition, for fixed $\varepsilon > 0$, it is known that $N_\varepsilon^T(H^n)$ is tight (see e.g. [62] Cor. A.9). In addition, we have assumed (AVCI) for the sequence (H^n, V^n) which is nothing else than

$$\lim_{m\to\infty} \limsup_{n\to\infty} \ \mathbb{P}^n\left(\hat{w}^T_{|\rho^m|}(H^n,\, V^n) \ > \ \gamma\right) \ = \ 0 \tag{6.12}$$

for all $\gamma > 0$ and $T \geq 1$. Furthermore, for fixed $m \geq 1$, $J \geq 1$, due to the convergence of the finite-dimensional distributions of V^n to 0 along D, it holds that

$$\sum_{p=1}^{k_m} |V^n_{t^m_p}| \ \xrightarrow[n\to\infty]{\mathbb{P}^n} \ 0 \qquad \text{and} \qquad \max_{t\,\in\,\pi^J} |V^n_t| \ \xrightarrow[n\to\infty]{\mathbb{P}^n} \ 0. \tag{6.13}$$

Finally it is well-known that the tightness of the X^n and U^n on the M1 Skorokhod space implies in particular

$$\lim_{J\to\infty} \limsup_{n\to\infty} \ \mathbb{P}^n\left(w''(X^n|_{[0,T]},\, |\pi^J|) + w''(U^n|_{[0,T]},\, |\pi^J|) \ > \ \gamma\right) \ = \ 0. \tag{6.14}$$

With regards to (6.11), we now employ (6.7), (6.12), (6.13) and (6.14) to deduce

$$\lim_{\varepsilon\to 0} \ \lim_{m\to\infty} \ \limsup_{n\to\infty} \ \mathbb{P}^n\left(\left|\int_0^\bullet (H^n_{r-} - H^{n\,|\,m,\varepsilon}_{r-}) \ \mathrm{d}V^n\right|^*_T \ > \ \gamma\right) \ = \ 0$$

for all $\gamma > 0$, which immediately yields (6.5). □

Proof of Proposition 4.4. Proceeding analogously to the proof of Theorem 4.6 above, it suffices to note that for every $\varepsilon > 0$, it almost surely holds

$$\left|\int_0^\bullet (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-}) \ \mathrm{d}V^n_s\right|^*_T \ \leq \ \sum_{k=0}^\infty |H^n_{s_k} - H^{n\,|\,m,\varepsilon}_{s_k}|\,|V^n_{s_{k+1}} - V^n_{s_k}| \ \leq \ \varepsilon \sum_{k=0}^\infty |V^n_{s_{k+1}} - V^n_{s_k}|$$

since, given that the H^n are pure jump, so are the $H^n - H^{n\,|\,m,\varepsilon}$ with $\mathrm{Disc}_{[0,T]}(H^n - H^{n\,|\,m,\varepsilon}) \subseteq \mathrm{Disc}_{[0,T]}(H^n)$. Thus,

$$\mathbb{P}^n\left(\left|\int (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-}) \ \mathrm{d}V^n\right|^*_T \ \geq \ \lambda\right) \ \leq \ \mathbb{P}^n\left(\sum_{k=0}^\infty |V^n_{s_{k+1}} - V^n_{s_k}| \ \geq \ \frac{\lambda}{\varepsilon}\right) \ \xrightarrow[n\to\infty]{} \ 0$$

by (4.3) for each $\lambda, \varepsilon > 0$. □

Proof of Proposition 4.8. By Lemma 4.2 and the convergence results for uncorrelated or correlated uncoupled CTRWs in (3.6), it remains to show (ii) of Definition 4.5. To this end, fix $T > 0$, $\gamma > \beta - \beta/\alpha$ and let $\rho = 1$ if $1 < \alpha \leq 2$ and $0 < \rho < 1$ such that (TC) and $\rho(\gamma + \beta/\alpha) > \beta$ hold if $\alpha = 1$. First note that for every random variable $s : \Omega \to [0, T]$, by monotone convergence and identical distribution of the θ_k, we obtain

$$\mathbb{E}\left[\,|V^n_s|^\rho\,\right] \ \leq \ \frac{\tilde{c}\,\psi^{-\rho}}{n^{\frac{\rho\beta}{\alpha}}}\,\mathbb{E}[\,|\theta_0|^\rho\,] \tag{6.15}$$

where $\tilde{c} := \sum_{i=1}^\infty (\sum_{j=i}^\infty c_j)^\rho < \infty$. Recall in particular that the ρ-th moment of θ_0 exists as its law is in the domain of attraction of an α-stable random variable where $1 \leq \alpha \leq 2$. For every $n \geq 1$ define (random) partitions ζ^n by

$$\pi^n \ := \ \underbrace{\left\{\frac{k}{n^\beta}T \ : \ k = 0, 1, ..., n^\beta\right\}}_{=:\pi^{n,1}} \ \cup \ \underbrace{\left\{\, 0 < s \leq T \ : \ \Delta N(ns) \neq 0 \,\right\}}_{=:\pi^{n,2}}$$

and the second set $\pi^{n,2}$ contains precisely the jumping times of V^n on $[0,T]$. Moreover, the mesh size of the partitions satisfy $|\pi^n| \leq n^{-\beta}$ and each of the elements of π^n is independent of the Z_k. In addition, since $t \mapsto N(nt)$ is a counting process,we have

$$\operatorname{card}(\pi^{n,2}) \;=\; \operatorname{card}\Big(\big\{\, 0 < s \leq T \;:\; \Delta N(ns) \neq 0 \,\big\}\Big) \;=\; N(nT).$$

Now, let $\varepsilon > 0$. Then, by tightness of $n^{-\beta}N(n\bullet)$, there exists $K_\varepsilon > 0$ such that $\mathbb{P}(\, N(nT) > K_\varepsilon n^\beta) \leq \varepsilon$. Hence, we have

$$\begin{aligned}
&\mathbb{P}\Big(n^{-\gamma} \sum_{s \,\in\, \pi^n} |V_s^n| \;>\; \lambda \Big) \;\leq\; \mathbb{P}\Big(\sum_{s \,\in\, \pi^{n,1}} |V_s^n| \;>\; \frac{\lambda n^\gamma}{2} \Big) \\
&\quad + \;\mathbb{P}\Big(\sum_{s \,\in\, \pi^{n,2}} |V_s^n| \;>\; \frac{\lambda n^\gamma}{2}\;,\; N(nT) \;\leq\; K_\varepsilon n^\beta \Big) + \;\mathbb{P}\,\Big(N(nT) \;>\; K_\varepsilon n^\beta \Big) \\
&\quad \leq\; \mathbb{P}\left(\sum_{k=0}^{n^\beta} |V^n_{\frac{k}{n^\beta}T}| \;>\; \frac{\lambda n^\gamma}{2} \right) \;+\; \mathbb{P}\left(\sum_{k=1}^{K_\varepsilon n^\beta} |V^n_{L_k \wedge T}| \;>\; \frac{\lambda n^\gamma}{2} \right) \;+\; \varepsilon
\end{aligned}$$

where $L_k := \sum_{i=1}^k J_i$ is defined before (3.4). Then, by Markov's inequality and (6.15)—since the times $kn^{-\beta}T$ and $L_k \wedge T$, $k = 0, 1, 2, \ldots$, are independent of the θ_ℓ—this implies

$$\begin{aligned}
\mathbb{P}\Big(n^{-\gamma} \sum_{s \,\in\, \zeta^n} |V_s^n| \;>\; \lambda \Big) \;&\leq\; \Big(\frac{2}{\lambda n^\gamma}\Big)^\rho \Big(\sum_{k=1}^{n^\beta} \mathbb{E}[\,|V^n_{\frac{k}{n^\beta}T}|^\rho\,] \;+\; \sum_{k=1}^{K_\varepsilon n^\beta} \mathbb{E}[\,|V^n_{L_k \wedge T}|^\rho\,] \Big) \;+\; \varepsilon \\
&\leq\; \frac{2^\rho \tilde{c}\, \psi^{-\rho}(K_\varepsilon + 1) n^\beta}{\lambda^\rho n^{\rho(\gamma + \frac{\beta}{\alpha})}}\; \mathbb{E}[\,|\theta_0|^\rho\,] \;+\; \varepsilon
\end{aligned}$$

for all $n \geq 1$. Since $\gamma > \beta - \beta/\alpha$ by assumption, and ρ is such that $\rho(\gamma + \beta/\alpha) > \beta$, we get

$$\mathbb{P}\Big(n^{-\gamma} \sum_{s \,\in\, \zeta^n} |V_s^n| \;>\; \lambda \Big) \;\leq\; \frac{2^\rho \tilde{c}\, \psi^{-\rho}(K_\varepsilon + 1) n^\beta}{\lambda^\rho n^{\rho(\gamma + \frac{\beta}{\alpha})}}\; \mathbb{E}[\,|\theta_0|^\rho\,] \;+\; \varepsilon \quad \xrightarrow[n\to\infty]{} \quad \varepsilon$$

and, as $\varepsilon > 0$ was arbitrary, this yields (ii) of Definition 4.5 for $\mathrm{GD} \bmod \mathrm{CA}(\gamma, \beta)$. $\square$

6.3. Proofs pertaining to Section 4.2

Proof of Propositon 4.10. We will only prove the claim for correlated uncoupled CTRWs, the proof for moving averages can be conducted in full analogy. Set $\tilde{U}^n \equiv 0$ and note that $\psi^{-1}X^n = U^n + \tilde{U}^n + \sum_{i=1}^\infty V^{n,i}$, where U^n is defined as in (4.1). Due to Lemma 4.2 and the convergence results for uncorrelated and correlated uncoupled CTRWs in (3.6), only (ii) of Definition 4.9 is still to be shown.

With stopping times $\sigma_k^{n,i} = \sigma_k^n = \sum_{\ell=1}^k J_\ell / n$, $k \geq 1$, we obtain by definition of $N(nt)$,

$$V^{n,i}_{\sigma_k^{n,i}} \;=\; -\,\frac{1}{\psi n^{\frac{\beta}{\alpha}}} \Big(\sum_{\ell=i}^\infty c_\ell \Big)\, \theta_{k-i+1} \tag{6.16}$$

if $k \geq i$ and it is equal to zero otherwise. Hence (ii.ii) of Definition 4.9 follows from the independence of the θ_k and the fact that they are centered as well as the integrability of the θ_k. Since $\Lambda^{n,i}(t) = N(nt)$, by definition of the $\sigma_k^{n,i}$ and given that $|N(n\bullet)/n^\beta|_T^*$ is tight in $\mathbb{R}$ for every $T > 0$, we obtain (ii.i) of Definition 4.9 with $f(n) = n^\beta$. Thus, it only remains to verify (ii.iii). To this end, let $k \geq 1$ and $\gamma > 0$ such that $\alpha - \gamma > 1$. Denote $\tilde{c}_i := \psi^{-1} \sum_{\ell=i}^\infty c_\ell \leq 1$, $\tilde{c} := \sum_{i=1}^\infty \tilde{c}_i$.

Then, using Jensen's inequality, monotone convergence and the identical distribution of the θ_k, we obtain

$$\mathbb{E}\left[\,(\tilde{V}_k^{n,>})^{\,\alpha-\gamma}\,\right] \;\leq\; \tilde{c}^{\alpha-\gamma}\, n^{-\frac{\beta(\alpha-\gamma)}{\alpha}}\, \mathbb{E}\left[\,|\theta_0\,|^{\,\alpha-\gamma}\; \mathbf{1}_{\{|\theta_0|\,>\,n^{\frac{\beta}{\alpha}}\}}\,\right] \tag{6.17}$$

where we recall that $\tilde{V}_k^{n,>} = \sum_{i=1}^{\infty} |V^{n,i}_{\sigma_k^{n,i}}|\, \mathbf{1}_{\{|V^{n,i}_{\sigma_k^{n,i}}|>1\}}$. Further, we continue by

$$\begin{aligned}
\mathbb{E}\Big[\,|\theta_0\,|^{\,\alpha-\gamma}\; \mathbf{1}_{\{|\theta_0|\,>\,n^{\frac{\beta}{\alpha}}\}}\Big] \;&=\; \int_{(n^{\frac{\beta}{\alpha}},\,\infty)} x^{\alpha-\gamma}\; \mathrm{d}\Phi(x) \;+\; \int_{(-\infty,\,-n^{\frac{\beta}{\alpha}})} (-x)^{\alpha-\gamma}\; \mathrm{d}\Phi(x)\\
&=\; -\int_{(n^{\frac{\beta}{\alpha}},\,\infty)} x^{\alpha-\gamma}\; \mathrm{d}(1-\Phi)(x) \;-\; \int_{(-\infty,\,-n^{\frac{\beta}{\alpha}})} (-x)^{\alpha-\gamma}\; \mathrm{d}(1-\Phi)(x)
\end{aligned} \tag{6.18}$$

where we have denoted the cdf of θ_0 by Φ. Integration by parts now yields

$$\begin{aligned}
\int_{n^{\frac{\beta}{\alpha}}}^{\infty} x^{\alpha-\gamma}\; \mathrm{d}(1-\Phi)(x) &= \Big[x^{\alpha-\gamma}(1-\Phi)(x)\Big]_{n^{\frac{\beta}{\alpha}}}^{\infty} \;-\; (\alpha-\gamma)\int_{n^{\frac{\beta}{\alpha}}}^{\infty} x^{\alpha-\gamma-1}\,(1-\Phi)(x)\,\mathrm{d}x\\
&\geq\; -n^{\frac{\beta(\alpha-\gamma)}{\alpha}}\; n^{-\beta}(1+\mathcal{O}(1)) \;-\; (\alpha-\gamma)\int_{n^{\frac{\beta}{\alpha}}}^{\infty} x^{\alpha-\gamma-1}\,x^{-\alpha}(1+\mathcal{O}(1))\,\mathrm{d}x\\
&\geq\; -\frac{3}{2}n^{-\frac{\beta\gamma}{\alpha}} \;+\; \frac{3(\alpha-\gamma)}{2\gamma}\,n^{-\frac{\beta\gamma}{\alpha}} \;=\; -\frac{3\alpha}{2\gamma}\,n^{-\frac{\beta\gamma}{\alpha}}
\end{aligned}$$

for n large enough, where we have used that $(1-\Phi)(x) \leq \mathbb{P}(|\theta_0| > x) \sim x^{-\alpha}$ and $\mathcal{O}(1)$ is the Landau notation for an asymptotically vanishing function in n or x respectively. Analogously, we achieve the same lower bound for the second integral of (6.18). Hence, this implies

$$\mathbb{E}\left[\,|\theta_0\,|^{\,\alpha-\gamma}\; \mathbf{1}_{\{|\theta_0|\,>\,n^{\frac{\beta}{\alpha}}\}}\,\right] \;\leq\; \frac{3\alpha}{\gamma}\,n^{-\frac{\beta\gamma}{\alpha}}$$

for n large enough, and combining this with (6.17), we deduce

$$\sum_{k=0}^{K\,n^{\beta}} \mathbb{E}\left[\,(\tilde{V}_k^{n,>})^{\,\alpha-\gamma}\,\right] \;\leq\; \frac{3\alpha}{\gamma}\,(Kn^{\beta}+1)\;n^{-\beta}\,\tilde{c}^{\alpha-\gamma} \;\leq\; \frac{6\alpha K\tilde{c}^{\alpha-\gamma}}{\gamma} \;<\; \infty$$

for n large enough, which yields the first bound in (ii.iii) in Definition 4.9 with $\lambda = \alpha-\gamma$. For the second quantity in (ii.iii), we obtain a similar bound with $\mu = \alpha+\gamma$, by proceeding similarly to the first uniform bound, using that $\mathbb{E}[|\theta_0|^{\alpha+\gamma-1}] < \infty$ and $\tilde{c}_i \leq 1$. □

Proof of Theorem 4.11. As in the proof of Theorem 4.6, it suffices to show that

$$\lim_{\varepsilon\downarrow 0}\;\limsup_{m\to\infty}\;\limsup_{n\to\infty}\;\mathbb{P}^n\bigg(\,\Big|\int_0^{\bullet} (H^n_{s-} - H^{n\,|\,m,\varepsilon}_{s-})\;\mathrm{d}V^n_s\,\Big|^*_T \;>\; \eta\bigg) \;=\; 0 \tag{6.19}$$

for every $T,\eta > 0$ with $V^n = \sum_{i\geq 1} V^{n,i}$. For this we exploit the main ideas of the proof of [3, Lem. 2(b)] and adapt it to our setting. Fix $T > 0$, $\eta > 0$ and denote $\tilde{H}^{n\,|\,m,\varepsilon} := H^n - H^{n\,|\,m,\varepsilon}$. For all $\varepsilon > 0$ and $n, m \geq 1$ it holds that

$$\begin{aligned}
\int_0^t \tilde{H}^{n\,|\,m,\varepsilon}_{s-}\;\mathrm{d}V^n_s \;&=\; \sum_{i=1}^{\infty}\sum_{k=1}^{\Lambda^{n,i}(t)} \tilde{H}^{n\,|\,m,\varepsilon}_{\sigma_k^{n,i}-}\;\Delta V^{n,i}_{\sigma_k^{n,i}}\\
&=\; \sum_{i=1}^{\infty}\sum_{k=1}^{\Lambda^{n,i}(t)} \tilde{H}^{n\,|\,m,\varepsilon}_{\sigma_k^{n,i}-}\;V^{n,i}_{\sigma_k^{n,i}} \;-\; \sum_{i=1}^{\infty}\sum_{k=1}^{\Lambda^{n,i}(t)} \tilde{H}^{n\,|\,m,\varepsilon}_{\sigma_k^{n,i}-}\;V^{n,i}_{\sigma_{k-1}^{n,i}}\,.
\end{aligned}$$

Let $\delta > 0$. By (ii.i) of Definition 4.9 there exists $K_\delta > 0$ such that

$$\mathbb{P}^n\bigg(\bigg|\int_0^\bullet \tilde{H}^{n\,|\,m,\varepsilon}_{s-}\,\mathrm{d}V^n_s\bigg|^*_T > \eta\bigg) \;\le\; \delta \;+\; \mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta}\bigg|\sum_{k=1}^{j}\sum_{i=1}^{\infty}\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\,V^{n,i}_{\sigma^{n,i}_k}\bigg| \;>\; \frac{\eta}{2}\bigg)$$
$$+\;\mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta}\bigg|\sum_{k=1}^{j}\sum_{i=1}^{\infty}\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\,V^{n,i}_{\sigma^{n,i}_{k-1}}\bigg| \;>\; \frac{\eta}{2}\bigg). \quad (6.20)$$

We begin by bounding the first probability summand in (6.20) and the bound for the second probability term can then be obtained in full analogy. To this end, we define

$$Z^{n,\le}_k \;:=\; \sum_{i=1}^{\infty}\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\Big(V^{n,i,\le}_{\sigma^{n,i}_k} \;-\; \mathbb{E}^n\Big[V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,k-1},\,\dots,\,\mathcal{V}_{n,1}\Big]\Big)$$
$$Z^{n,>}_k \;:=\; \sum_{i=1}^{\infty}\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\Big(V^{n,i,>}_{\sigma^{n,i}_k} \;+\; \mathbb{E}^n\Big[V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,k-1},\,\dots,\,\mathcal{V}_{n,1}\Big]\Big)$$

where we have denoted $V^{n,i,\le}_{\sigma^{n,i}_k} := V^{n,i}_{\sigma^{n,i}_k}\,\mathbf{1}_{\{|V^{n,i}_{\sigma^{n,i}_k}|\le 1\}}$ as well as $V^{n,i,>}_{\sigma^{n,i}_k} := V^{n,i}_{\sigma^{n,i}_k}\,\mathbf{1}_{\{|V^{n,i}_{\sigma^{n,i}_k}|>1\}}$, and $\mathcal{V}_{n,\ell}$ is defined as in (ii.ii) of Definition 4.9. Let us remark that $Z^{n,\le}_k + Z^{n,>}_k = \sum_{i\ge 1}\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\,V^{n,i}_{\sigma^{n,i}_k}$. The discrete-time processes $j \mapsto \sum_{k=1}^{j} Z^{n,\le}_k$ and $j \mapsto \sum_{k=1}^{j} Z^{n,>}_k$ are both martingales with respect to the filtration $\{\sigma(\mathcal{V}_{n,\ell},\mathcal{H}_{n,\ell} : \ell \le j)\}_{j\ge 1}$, where $\mathcal{H}_{n,\ell} := \sigma(\,H^n_{t\wedge\sigma^{n,i}_\ell} : t \ge 0, i \ge 1)$. To see this, it suffices to note that for $k > j$ we have

$$\mathbb{E}^n\big[Z^{n,\le}_k\,|\,\mathcal{V}_{n,\ell},\mathcal{H}_{n,\ell}\,:\,\ell \le j\big] = \mathbb{E}^n\Big[\mathbb{E}^n\big[\,Z^{n,\le}_k\,|\,\mathcal{V}_{n,\ell},\mathcal{H}_{n,\ell},\mathcal{H}_{n,k} : \ell \le j\big]\,|\,\mathcal{V}_{n,\ell},\mathcal{H}_{n,\ell} : \ell \le j\,\Big]$$

and, by dominated convergence, since the $\tilde{V}^{n,\le}_k$ —as defined in (ii.iii) of Definition 4.9—are integrable due to (ii.iii), it is enough to show

$$\mathbb{E}^n\bigg[\,\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\Big(V^{n,i,\le}_{\sigma^{n,i}_k} \;-\; \mathbb{E}^n\Big[V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,k-1},\,\dots,\,\mathcal{V}_{n,1}\Big]\Big)\;|\;\mathcal{V}_{n,\ell}\,,\,\mathcal{H}_{n,\ell},\,\mathcal{H}_{n,k}\,:\,\ell \le j\,\bigg] \;=\; 0$$

for all $i \ge 1$. Based on the measurability of $\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}$ with respect to $\mathcal{H}_{n,k}$ as well as assumption (4.4), that is the independence of $V^{n,i}_{\sigma^{n,i}_k}$ and $\mathcal{H}_{n,\ell}$ for all $\ell \le k$, we obtain

$$\begin{aligned}
&\mathbb{E}^n\bigg[\,\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\Big(V^{n,i,\le}_{\sigma^{n,i}_k} \;-\; \mathbb{E}^n\Big[V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,k-1},\,\dots,\,\mathcal{V}_{n,1}\Big]\Big)\;|\;\mathcal{V}_{n,\ell}\,,\,\mathcal{H}_{n,\ell},\,\mathcal{H}_{n,k}\,:\,\ell \le j\,\bigg]\\
&\quad=\;\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\bigg(\mathbb{E}^n\bigg[\,V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,\ell}\,,\,\mathcal{H}_{n,\ell},\,\mathcal{H}_{n,k}\,:\,\ell \le j\bigg] \;-\; \mathbb{E}^n\bigg[\,V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,\ell}\,:\,\ell \le j\bigg]\bigg)\\
&\quad=\;\tilde{H}^{n\,|\,m,\varepsilon}_{\sigma^{n,i}_k-}\bigg(\mathbb{E}^n\bigg[\,V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,\ell}\,:\,\ell \le j\bigg] \;-\; \mathbb{E}^n\bigg[\,V^{n,i,\le}_{\sigma^{n,i}_k}\,|\,\mathcal{V}_{n,\ell}\,:\,\ell \le j\bigg]\bigg) = 0.
\end{aligned}$$

We proceed similarly for the process $j \mapsto \sum_{k=1}^{j} Z^{n,>}_k$ by using $V^{n,i,>}_{\sigma^{n,i}_k} = V^{n,i}_{\sigma^{n,i}_k} - V^{n,i,\le}_{\sigma^{n,i}_k}$ and (ii.ii) from Definition 4.9. Since both $j \mapsto \sum_{k=1}^{j} Z^{n,\le}_k$ and $j \mapsto \sum_{k=1}^{j} Z^{n,>}_k$ are discrete-time martingales, we can apply Doob's maximal inequality to obtain

$$\begin{aligned}
&\mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta} \Big| \sum_{k=1}^{j} \sum_{i=1}^{\infty} \tilde{H}^{n\,|\,m,\varepsilon}_{\sigma_k^{n,i}-} \, V^{n,i}_{\sigma_k^{n,i}} \Big| \; > \; \frac{\eta}{2} \bigg) \\
&\leq \; \mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta} \Big| \sum_{k=1}^{j} Z_k^{n,\leq} \Big| \; > \; \frac{\eta}{4} \bigg) \; + \; \mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta} \Big| \sum_{k=1}^{j} Z_k^{n,>} \Big| \; > \; \frac{\eta}{4} \bigg) \qquad (6.21) \\
&\leq \; \Big(\frac{4}{\eta}\Big)^{\mu} \, \mathbb{E}^n\bigg[\Big| \sum_{k=1}^{f(n)K_\delta} Z_k^{n,\leq} \Big|^{\mu} \bigg] \; + \; \Big(\frac{4}{\eta}\Big)^{\lambda} \, \mathbb{E}^n\bigg[\Big| \sum_{k=1}^{f(n)K_\delta} Z_k^{n,>} \Big|^{\lambda} \bigg]
\end{aligned}$$

and by Bahr-Esseen's martingale-differences inequality (cf. [5, Theorem 2]), this can be further estimated by

$$\leq \; \big(\frac{2^{\frac{1}{\mu}}4}{\eta}\big)^{\mu} \sum_{k=1}^{f(n)K_\delta} \mathbb{E}^n\Big[\Big| Z_k^{n,\leq} \Big|^{\mu} \Big] \; + \; \big(\frac{2^{\frac{1}{\lambda}}4}{\eta}\big)^{\lambda} \sum_{k=1}^{f(n)K_\delta} \mathbb{E}^n\Big[\Big| Z_k^{n,>} \Big|^{\lambda} \Big]. \qquad (6.22)$$

Finally, with $\tilde{V}_k^{n,\leq}$ defined as in (ii.iii) Definition 4.9, by a simple applications of Jensen's inequality we deduce

$$\mathbb{E}^n\Big[\Big| Z_k^{n,\leq} \Big|^{\mu} \Big] \; \leq \; (2\varepsilon)^{\mu} \, \mathbb{E}^n[\, (\tilde{V}_k^{n,\leq})^{\mu} \,] \quad \text{for all } \; n, k \geq 1,$$

since $|\tilde{H}^{n\,|\,m,\varepsilon}| \; \leq \; \varepsilon$ by its definition. Analogously we obtain the estimate $\mathbb{E}^n[|Z_k^{n,>}|^{\lambda}] \; \leq \; 2^{\lambda}\varepsilon^{\lambda}\, \mathbb{E}^n[(\tilde{V}_k^{n,>})^{\lambda}]$. Hence, based on (6.22), this yields

$$\begin{aligned}
&\limsup_{n\to\infty} \; \mathbb{P}^n\bigg(\sup_{j=1,\dots,f(n)K_\delta} \Big| \sum_{k=1}^{j} \sum_{i=1}^{\infty} \tilde{H}^{n\,|\,m,\varepsilon}_{\sigma_k^{n,i}-} \, V^{n,i}_{\sigma_k^{n,i}} \Big| \; > \; \frac{\eta}{2} \bigg) \\
&\leq \; \max\,\{\varepsilon^{\lambda}, \varepsilon^{\mu}\}\; C_{\lambda,\mu,\eta} \; \limsup_{n\to\infty} \sum_{k=1}^{f(n)K_\delta} \Big(\mathbb{E}^n[\,(\tilde{V}_k^{n,\leq})^{\mu}\,] \, + \, \mathbb{E}^n[\,(\tilde{V}_k^{n,>})^{\lambda}\,] \Big) \qquad (6.23)
\end{aligned}$$

where $C_{\lambda,\mu,\eta} := 2(8/\eta)^{\lambda} + 2(8/\eta)^{\mu}$. Then, the limit superior part of (6.23) is finite by (ii.iii) of Definition 4.9 and (6.23) tends to 0 as $\varepsilon \to 0$. Analogously to the above procedure, one can achieve the same bound for the second probability term of (6.20) and hence deduce (6.19). □

Proof of Corollary 4.13. We only need to verify the conditions (4.5) and (4.6) from Remark 4.12, as the remaining properties of Definition 4.9 follow in the exact same way as outlined in the proof of Proposition 4.10 (we recall, with $f(n) = n^{\beta}$). Due to the symmetry of the law of the θ_k, (4.5) is obtained immediately with the $V^{n,i}$ and $\sigma_k^{n,i}$ as in (6.16). Moreover, since $t \mapsto \sum_{k=1}^{\lfloor n^{\beta}t\rfloor} n^{-\beta/\alpha}\theta_k$ is a subsequence of the zero-order moving averages (3.1) converging weakly in the J1 Skorokhod space, to every $\delta > 0$ there exists $\Gamma_\delta > 0$ such that

$$\sup_{n\geq 1} \; \mathbb{P}\bigg(\sum_{k=1}^{Kn^{\beta}} \mathbf{1}_{\{|\theta_k| > n^{\beta/\alpha}\}} \, > \, \Gamma_\delta \, , \; n^{-\beta/\alpha}|\theta_k| \, > \, \Gamma_\delta \bigg) \; \leq \; \delta$$

as the maximal number of large oscillations as well as the absolute size of jumps is tight see e.g. [62, Thm. A.8 & Cor. A.9] However, this yields

$$\begin{aligned}
\mathbb{P}\bigg(\sum_{k=1}^{Kn^{\beta}} \sum_{i=1}^{\infty} |V^{n,i,>}_{\sigma_k^{n,i}}| \; \geq \; M \bigg) \; &\leq \; \mathbb{P}\bigg(\tilde{c} \sum_{k=1}^{Kn^{\beta}} n^{-\beta/\alpha}|\theta_k|\, \mathbf{1}_{\{|\theta_k| > n^{\beta/\alpha}\}} \; \geq \; M \bigg) \\
&\leq \; \delta \; + \; \mathbb{P}\big(\, \tilde{c}\,\Gamma_\delta^2 \; \geq \; M \, \big) \; = \; \delta
\end{aligned}$$

for M large enough, where $\tilde{c} = \psi^{-1}\sum_{i=1}^{\infty}\sum_{\ell=i}^{\infty} c_i$. A δ was arbitrary, we deduce (4.6). □

6.4. Proofs pertaining to Sections 5.1 and 5.2

Lemma 6.2. *The solutions X^n to* (S_n) *are stochastically bounded uniformly in n, that is* $\lim_{\eta\to\infty}\sup_{n\geq 1}\mathbb{P}(|X^n|^*_T > \eta) = 0$ *for all $T \geq 0$.*

Proof. To alleviate notation, we will only provide a proof for the case $C \equiv 0$. On behalf of the decomposition (3.9), we can write $Z^n = M^n + A^n$, where the M^n are local martingales with $|\Delta M^n| \leq 1$ and the A^n are of tight total variation. Thus, with $\tilde{\sigma} := |b| \vee |\mu| \vee |\sigma|$ and $\tilde{A}^n := \mathrm{Id} + D^n + A^n$, and the elementary inequality $ac \leq (a^2+c^2)/2$, we obtain

$$(X^n_t)^2 \;\leq\; \frac{3}{2}\Big[\Big(\int_0^t \sigma(s,D^n_s,X^n_s)_-\,\mathrm{d}M^n\Big)^2 \;+\; \int_0^t \tilde{\sigma}(s,D^n_s,X^n_s)^2_-\,\mathrm{d}\,\mathrm{TV}_{[0,s]}(\tilde{A}^n)\Big], \tag{6.24}$$

for $t \geq 0$. Let $T,\varepsilon,\eta > 0$ and define stopping times $\tau_n := \inf\{t>0 : |X^n_t| > \eta\} \wedge T$ and $\rho_n := \inf\{t > 0 : |D^n_t| > R\} \wedge T$ with $R > 0$ such that $\sup_{n\geq 1}\mathbb{P}(\rho_n \leq T) \leq \varepsilon$ (this is indeed possible since, according to (3.5), the D^n converge in the Skorokhod space and are therefore stochastically bounded). Then, we may continue by

$$\begin{aligned}\mathbb{P}(|X^n|^*_T > \eta) \;\leq\; \mathbb{P}(|X^n_{\tau_n\wedge\rho_n}|^2 > \eta^2) + \varepsilon \;\leq\; &\mathbb{P}\Big(\Big(\int_0^{\tau_n\wedge\rho_n}\sigma(s,D^n_s,X^n_s)_-\,\mathrm{d}M^n\Big)^2 > \frac{\eta^2}{3}\Big)\\ &+ \mathbb{P}\Big(K^2\int_0^{\tau_n\wedge\rho_n}|X^n_{s-}|^{2p}\,\mathrm{d}\,\mathrm{TV}_{[0,s]}(\tilde{A}^n) > \frac{\eta^2}{3}\Big) + \varepsilon\\ \leq\; &\mathbb{P}\Big(\Big(\int_0^{\tau_n\wedge\rho_n}\sigma(s,D^n_s,X^n_s)_-\,\mathrm{d}M^n\Big)^2 > \frac{\eta^2}{3}\Big)\\ &+ \mathbb{P}\Big(\mathrm{TV}_{[0,T]}(\tilde{A}^n) > \frac{\eta^{2-2p}}{3K^2}\Big) + \varepsilon\end{aligned} \tag{6.25}$$

using (6.24), the strict sublinear growth bound (5.2) as well as that $|X^n_s| \leq \eta$ for all $s < \tau_n$. An application of Lenglart's inequality [25, Lem. I.3.30b] to the first term of (6.25) (with the quadratic variation as L-domination process; applicable since the integral with respect to M^n is a local martingale) yields

$$\mathbb{P}(|X^n|^*_T > \eta) \leq \frac{3\gamma+3}{\eta^2} + \mathbb{P}\Big([M^n]_T > \frac{\gamma}{\eta^{2p}K^2}\Big) + \mathbb{P}\Big(\mathrm{TV}_{[0,T]}(\tilde{A}^n) > \frac{\eta^{2-2p}}{3K^2}\Big) + \varepsilon \tag{6.26}$$

for all $\gamma > 0$, where we made use of the strict sublinear growth bound (5.2). Clearly, the $\tilde{A}^n$ are of tight total variation on $[0,T]$ and the tightness of the $[M^n]_T$ follows from another application of Lenglart's inequality [62, Lem. 5.4], the good decompositions of the Z^n and their tightness in the Skorokhod space. Choosing $\gamma \in (2p,2)$ yields $\gamma/\eta^2 \to 0$ and $\gamma/\eta^{2p} \to \infty$ as $\eta \to \infty$, so we deduce the claim from (6.26) and the fact that $\varepsilon > 0$ was arbitrary. □

Proof of Theorem 5.1. As for Lemma 6.2, purely for simplicity of notation, we assume $C \equiv 0$ in (5.2). To avoid repetition, we note that we can proceed similarly to the proof of Theorem 5.2 below, and so it suffices to jump directly to the verification of the following properties:

1. the X^n are tight on the space $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{J1}})$;
2. as well as there being convergence

$$\int_0^\bullet \mu(s,D^n_s,X^n_s)_-\,\mathrm{d}D^n_s \Rightarrow \int_0^\bullet \mu(s,D^{-1}_s,Y_s)_-\,\mathrm{d}D^{-1}_s \quad \text{and}$$
$$\int_0^\bullet \sigma(s,D^n_s,X^n_s)_-\,\mathrm{d}Z^n_s \Rightarrow \int_0^\bullet \sigma(s,D^{-1}_s,Y_s)_-\,\mathrm{d}Z_{D^{-1}_s}$$

whenever $(X^n,D^n) \Rightarrow (Y,D^{-1})$ and $(X^n,D^n,Z^n) \Rightarrow (Y,Z)$ for a càdlàg process Y.

For the second convergence of the second claim, as in the proof of Theorem 5.2, it is straightforward to show that the $(\sigma(\bullet, D^n_\bullet, X^n_\bullet), Z^n)$ satisfy the assumptions of [62, Prop. 3.22]. The good decompositions of the integrators follow from Theorem 3.3. According to the brief remark after [62, Thm. 4.8], we can replace (AVCI) by the conditions (a)&(b) set out in [62, Thm. 4.8]. Taking $\sigma^n_k := L_k/n$ the jump times of Z^n, condition (b) follows as detailed in the remark after [62, Ex. 4.11]. In terms of (a), we note that $H^n := \sigma(\bullet, D^n, X^n_\bullet)$ is adapted to the filtration generated by $\{\theta_{N(ns)}, N(ns) : 0 \le s \le t\}$, $t \ge 0$. Thus, we deduce

$$(H^n_{t_1}, ..., H^n_{t_\ell})^{-1}(A_1 \times ... \times A_\ell) \cap \{q_2 \,<\, \sigma^n_{k+1}\} \cap \{\sigma^n_k \,\le\, q_1\} \,\in\, \sigma\big(\theta_1, ..., \theta_k, J_1, ..., J_{k+1}\big)$$

for all $A_1, ..., A_\ell \in \mathscr{B}(\mathbb{R})$, $n, k \ge 1$, $q_1, q_2 \in \mathbb{Q}$ and $t_1, ..., t_\ell \le q_2$, while $Z^n_{\sigma^n_{k+1}+\bullet} - Z^n_{\sigma^n_k}$ is $\sigma(\theta_{k+i}, J_{k+1+i} : i \ge 1)$–measurable. Applying [62, Lem. 4.12] yields condition (a) and hence the second convergence. The convergence of the first integrals follows directly from [62, Prop. 3.22] and the fact that (AVCI) is satisfied by Proposition 2.3 due to continuity of D^{-1}.

In order to show 1, we employ Aldous' J1 tightness criterion given in [8, Thm. 16.10], requiring us to verify that for every $T > 0$ and $\gamma > 0$ it holds

$$\lim_{R\to\infty} \sup\nolimits_{n\ge1} \mathbb{P}(|X^n|^*_T \,>\, R) = 0 \quad \text{and} \quad \lim_{\delta\downarrow0} \limsup_{n\to\infty} \sup\nolimits_\tau \mathbb{P}(|X^n_{\tau+\delta} - X^n_\tau| > \eta) = 0 \qquad (6.27)$$

where the inner supremum on the right-hand side runs over all $\mathbb{F}^n$–stopping times τ which are bounded by T. The first part of (6.27) follows immediately from Lemma 6.2. Towards the second part of (6.27), let $\varepsilon > 0$ and fix $\eta, T > 0$. Choose $R > 0$ large enough such that $\sup_{n\ge1} \mathbb{P}(|X^n|^*_{T+1} \vee |D^n|^*_{T+1} > R) \le \varepsilon/4$ and define stopping times $\rho_n := \inf\{t > 0 : |X^n_t| \vee |D^n_t| > R\}$. Then, for any $\mathbb{F}^n$-stopping time τ bounded by T and $\delta \le 1$, we have

$$\mathbb{P}(|X^n_{\tau+\delta} - X^n_\tau| \,>\, \eta) \;\le\; \frac{\varepsilon}{4} \,+\, \mathbb{P}(|X^n_{(\tau+\delta)\wedge\rho_n} - X^n_{\tau\wedge\rho_n}| \,>\, \eta). \qquad (6.28)$$

By (S_n), the definition of ρ_n and the strict sublinear growth condition (5.2), the second term on the right side of (6.28) can be further estimated by

$$\begin{aligned}\mathbb{P}(|X^n_{(\tau+\delta)\wedge\rho_n} - X^n_{\tau\wedge\rho_n}| \,>\, \eta) \;\le\; \mathbf{1}_{\delta > \frac{\eta}{3KR^p}} \,+ \mathbb{P}\Big(D^n_{(\tau+\delta)\wedge\rho_n} - D^n_{\tau\wedge\rho_n} \,>\, \frac{\eta}{3KR^p}\Big) \\ +\, \mathbb{P}\Big(\Big|\int_{\tau\wedge\rho_n}^{(\tau+\delta)\wedge\rho_n} \sigma(s, D^n_s, X^n_s)_- \,\mathrm{d}Z^n_s\Big| \,>\, \frac{\eta}{3}\Big) \qquad (6.29)\end{aligned}$$

where we recall that $D^n = n^{-\beta}N(n\bullet)$. While the first term on the right side of (6.29) disappears uniformly in n when $\delta \to 0$, this is also true for the second term. Indeed, recall that since $D^n \Rightarrow D^{-1}$ on the J1 space and D^{-1} is continuous, it holds in particular that

$$\lim_{\delta\downarrow0} \limsup_{n\to\infty} \mathbb{P}\Big(\sup_{\substack{0\le s\le t\le T+1\\ t-s\le\delta}} (D^n_t - D^n_s) \,>\, \lambda\Big) \;=\; 0 \qquad (6.30)$$

for all $\lambda > 0$. Thus, it suffices to investigate the convergence of the third term on the right-hand side of (6.29). Choose $a > 0$ such that $(3aKR^p/\eta)^2 \le \varepsilon/4$, set $\tilde{C}_a := \sup_{n\ge1} n^\beta \mathbb{E}[\zeta^n_1 \mathbf{1}_{\{|\zeta^n_1|>a\}}] < \infty$ (cf. Proposition 3.5), and let $Z^n = M^{n,a} + A^{n,a}$ be good decompositions (3.9) so that $|\Delta M^{n,a}| \le a$. Making use of the concrete form of the good decompositions as well as again the definition of ρ_n and the strict sublinear growth condition (5.2), we obtain

$$\begin{aligned}\mathbb{P}\Big(\Big|\int_{\tau\wedge\rho_n}^{(\tau+\delta)\wedge\rho_n} \sigma(s, D^n_s, X^n_s)_- \,\mathrm{d}Z^n_s\Big| \,>\, \frac{\eta}{3}\Big) \;\le\; \mathbb{P}\Big(\Big|\int_{\tau\wedge\rho_n}^{(\tau+\delta)\wedge\rho_n} \sigma(s, D^n_s, X^n_s)_- \,\mathrm{d}M^{n,a}_s\Big|^2 \,>\, \frac{\eta^2}{9}\Big) \\ +\, \mathbb{P}\Big(|\Delta Z^n|^*_{T+\delta} \sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{|\zeta^n_k|\,>\,a\}} \,>\, \frac{\eta}{3KR^p}\Big) \,+\, \mathbb{P}\Big(D^n_{\tau+\delta} - D^n_\tau \,>\, \frac{\eta}{3KR^p\tilde{C}_\alpha}\Big). \quad (6.31)\end{aligned}$$

For the first probability term of (6.31), similarly to the proof of Lemma 6.2 we apply Lenglart's inequality [25, Lem. I.3.30b] in order to obtain

$$\mathbb{P}\Big(\Big|\int_{\tau\wedge\rho_n}^{(\tau+\delta)\wedge\rho_n} \sigma(s, D_s^n, X_s^n)_- \,\mathrm{d}M_s^{n,a}\Big|^2 \;>\; \frac{\eta^2}{9}\Big) \;\le\; \frac{9(\gamma + (aKR^p)^2)}{\eta^2} + \mathbb{P}\Big([M^{n,a}]_{\tau+\delta} - [M^{n,a}]_\tau \;>\; \frac{\gamma}{K^2R^{2p}}\Big)$$

for all $\gamma > 0$. Choose $\gamma > 0$ such that $9\gamma/\eta^2 \le \varepsilon/4$. Hence, we can further bound (6.31) by

$$\begin{aligned}
&\mathbb{P}\Big(\Big|\int_{\tau\wedge\rho_n}^{(\tau+\delta)\wedge\rho_n} \sigma(s, D_s^n, X_s^n)_- \,\mathrm{d}Z_s^n\Big| \;>\; \frac{\eta}{3}\Big) \\
&\le \frac{\varepsilon}{2} + \mathbb{P}\Big([M^{n,a}]_{\tau+\delta} - [M^{n,a}]_\tau \;>\; \frac{\gamma}{K^2R^{2p}}\Big) + \mathbb{P}\Big(|Z^n|^*_{T+\delta} \sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{|\zeta_k^n|>a\}} \;>\; \frac{\eta}{3KR^p}\Big) \\
&\quad + \mathbb{P}\Big(D^n_{\tau+\delta} - D^n_\tau \;>\; \frac{\eta}{3KR^p\tilde{C}_\alpha}\Big) \\
&\le \frac{\varepsilon}{2} + \mathbb{P}\Big(\sum_{k=N(n\tau)}^{N(n(\tau+\delta))} (\zeta_k^n)^2\, \mathbf{1}_{\{|\zeta_k^n|\le a\}} \;>\; \frac{\gamma}{3(aKR^p)^2}\Big) + \mathbb{P}\Big(n^{-\beta} D^n_{T+\delta} \;>\; \frac{\gamma^2}{12a\tilde{C}_aK^2R^{2p}}\Big) \\
&\quad + \mathbb{P}\Big(n^{-\beta} D^n_{T+\delta} \;>\; \frac{\gamma}{6\tilde{C}_a^2K^2R^{2p}}\Big) + \mathbb{P}\Big(|Z^n|^*_{T+\delta} \sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{|\zeta_k^n|>a\}} \;>\; \frac{\eta}{3KR^p}\Big) \\
&\quad\quad + \mathbb{P}\Big(D^n_{\tau+\delta} - D^n_\tau \;>\; \frac{\eta}{3KR^p\tilde{C}_\alpha}\Big),
\end{aligned} \tag{6.32}$$

where we have made use of the specific form of $M^{n,a}$ given after (3.9). Since the Z^n are J1 tight, we can choose $C_\varepsilon > 0$ such that $\mathbb{P}(|Z^n|^*_{T+\delta} > C_\varepsilon) \le \varepsilon/4$. We will examine each term of (6.32) individually. Clearly, since $D^n_{T+\delta}$ are tight, we directly deduce

$$\limsup_{n\to\infty} \Big[\mathbb{P}\Big(n^{-\beta}\, D^n_{T+\delta} \;>\; \frac{\gamma^2}{12a\tilde{C}_aK^2R^{2p}}\Big) + \mathbb{P}\Big(n^{-\beta} D^n_{T+\delta} \;>\; \frac{\gamma}{6\tilde{C}_a^2K^2R^{2p}}\Big)\Big] \;=\; 0.$$

As $\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \mathbb{P}(D^n_{\tau+\delta} - D^n_\tau \;>\; \eta/3KR^p\tilde{C}_a) = 0$ due to (6.30), it only remains to consider the second and fifth terms of (6.32). For the latter, by independence of $N(n\bullet)$ and the ζ_k^n (recall the CTRW is uncoupled) as well as the identical distribution of the ζ_k^n,

$$\begin{aligned}
\mathbb{P}\Big(|Z^n|^*_{T+\delta} \sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{|\zeta_k^n|>a\}} \;>\; \frac{\eta}{3KR^p}\Big) \;&\le\; \frac{\varepsilon}{4} + \mathbb{P}\Big(\sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{|\zeta_k^n|>a\}} \;>\; \frac{\eta}{3KR^pC_\varepsilon}\Big) \\
&\le\; \frac{\varepsilon}{4} + \mathbb{P}\Big(D^n_{\tau+\delta} - D^n_\tau \;>\; \lambda\Big) + \mathbb{P}\Big(\sum_{k=0}^{\lfloor n^\beta\lambda\rfloor} \mathbf{1}_{\{|\zeta_k^n|>a\}} \;>\; \frac{\eta}{3KR^pC_\varepsilon}\Big) \\
&\le\; \frac{\varepsilon}{4} + \mathbb{P}\Big(D^n_{\tau+\delta} - D^n_\tau \;>\; \lambda\Big) + \mathbb{P}\Big(\tilde{N}_a^\lambda(\tilde{Z}^{n^\beta}) \;>\; \frac{\eta}{3KR^pC_\varepsilon}\Big)
\end{aligned}$$

for any $\lambda < 1$, where $\tilde{Z}^n := \sum_{k=0}^{\lfloor n\bullet\rfloor} \zeta_k^n$ is tight in J1 as a zero-order moving average and $\tilde{N}_a^T(\tilde{Z}^n)$ the maximal number of a-increments of $\tilde{Z}^n$ on $[0,T]$ (see e.g. [62, (3.9)]). According to the classical tightness criterion based on the J1 modulus of continuity (see e.g. [62, Def. A.7]), we deduce that $\lim_{\lambda\downarrow 0}\limsup_{n\to\infty} \mathbb{P}(\tilde{N}_a^\lambda(\tilde{Z}^{n^\beta}) > \eta/(3KR^pC_\varepsilon)) = 0$. Further, again due to (6.30), $\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \mathbb{P}(D^n_{\tau+\delta} - D^n_\tau \;>\; \lambda) = 0$, and thus

$$\lim_{\delta\downarrow 0}\, \limsup_{n\to\infty}\, \mathbb{P}\Big(|Z^n|^*_{T+\delta} \sum_{k=N(n\tau)}^{N(n(\tau+\delta))} \mathbf{1}_{\{\zeta_k^n>a\}} \;>\; \frac{\eta}{3KR^p}\Big) \;=\; \frac{\varepsilon}{4}.$$

Finally, for the second term of (6.32) we can proceed similarly to how we did for the previous (fifth) term, noting that $[\tilde{Z}^n] = \sum_{k=0}^{\lfloor n\bullet\rfloor}(\zeta_k^n)^2$ is tight on the J1 space as a consequence of the short comment after [62, Cor. 3.13]. $\square$

Proof of Theorem 5.2. Suppose we have weak relative compactness of the pairs $((X^n, c^{-1}Z^n))_{n\geq 1}$ on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})^2$, as we will establish later in the proof. Then we know that, to every subsequence $((X^{n_k}, c^{-1}Z^{n_k}))_{k\geq 1}$, there is a further subsubsequence $((X^{n_{k_\ell}}, c^{-1}Z^{n_{k_\ell}}))_{\ell\geq 1}$ which converges weakly in $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})^2$ to some a càdlàg limit (Y, Z). Of course, Y may depend on the choice of the subsubsequence, while the scaling limit Z is given in Section 3.1. To simplify the notation, we denote the subsequence by $(X^n, c^{-1}Z^n)$. Due to the continuity of b and σ, we can then deduce that

1. for all $\ell \geq 1$ and $t_1, ..., t_\ell \in \Lambda$, where Λ is a co-countable subset of $[0,\infty)$, it holds
$$\begin{aligned}(c^{-1}Z^n,\, &\sigma(t_1, X^n_{t_1-r}),\, \sigma(t_2, X^n_{t_2-r}),\, ...,\, \sigma(t_\ell, X^n_{t_\ell-r}))\\ &\Rightarrow\ (Z,\, \sigma(t_1, Y_{t_1-r}),\, \sigma(t_2, Y_{t_2-r}),\, ...,\, \sigma(t_\ell, Y_{t_\ell-r}))\end{aligned}$$
on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbb{R}^\ell, |\cdot|)$;
2. for any $T > 0$, the sequence $(|\sigma(\bullet, X^n_{\bullet-r})|^*_T)_{n\geq 1}$ is tight;
3. for any $\delta > 0$, the number of δ-increments of the $(\sigma(\bullet, X^n_{\bullet-r}))_{n\geq 1}$ over any interval $[0,T]$ is tight (see [62, (3.9)] for a precise definition);

and analogously for the sequence $b(\bullet, X^n_{\bullet-r})$. Now, for $n \geq 1$ such that $(\mathcal{J}+2)/n < r$, we have that, for each $k \geq 1$ the random variable $X^n_{(k/n)-r}$ is $\sigma(\theta_1, ..., \theta_{k-\mathcal{J}-2}, J_1, ..., J_{k-\mathcal{J}-2})$-measurable and therefore $\sigma(\bullet, X^n_{\bullet-r})$ satisfies (4.7). Then, according to [62, Prop. 3.22] and Remark 4.16, we obtain that

$$\begin{aligned}\int_0^\bullet b(s, X^n_{s-r})_-\, \mathrm{d}s \ &\Rightarrow\ \int_0^\bullet b(s, Y_{s-r})_-\, \mathrm{d}s \quad \text{and}\\ c^{-1}\int_0^\bullet \sigma(s, X^n_{s-r})_-\, \mathrm{d}Z^n_s \ &\Rightarrow\ \int_0^\bullet \sigma(s, Y_{s-r})_-\, \mathrm{d}Z_s,\end{aligned}$$

as $n \to \infty$ on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$. Since $t \mapsto \int_0^t b(s, Y_{s-r})_-\, \mathrm{d}s$ is continuous, we obtain that

$$\begin{aligned}X^n \ = \ X^n_0 + \int_0^\bullet b(s, X^n_{s-r})_-\, \mathrm{d}s \ + \ c^{-1}\int_0^\bullet \sigma(s, X^n_{s-r})_-\, \mathrm{d}Z^n_s\\ \Rightarrow \quad Y_0 + \int_0^\bullet b(s, Y_{s-r})_-\, \mathrm{d}s \ + \ \int_0^\bullet \sigma(s, Y_{s-r})_-\, \mathrm{d}Z(s)\end{aligned}$$

as $n \to \infty$ in $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$. Uniqueness of weak limits implies that Y satisfies

$$Y_t \ = \ Y_0 \ + \ \int_0^t b(s, Y_{s-r})_-\, \mathrm{d}s \ + \ \int_0^t \sigma(s, Y_{s-r})_-\, \mathrm{d}Z(s),$$

so it realises a solution to (Š). If there is uniqueness in law for (Š), then Y is the unique solution. Since any subsequence has a further subsequence converging to this same limit, we then obtain the weak convergence $X^n \Rightarrow X$ on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})$.

At this point, it only remains to show that $((X^n, c^{-1}Z^n))_{n\geq 1}$ is relatively compact in the weak topology on $(\mathbf{D}_{\mathbb{R}}[0,\infty), \mathrm{d}_{\mathrm{M1}})^2$, which—due to Prokhorov's Theorem—equates to proving the tightness of these pairs on the given space. However, as the tightness of the $c^{-1}Z^n$ has already been established as a consequence of (3.6), it is enough to treat the tightness of the X^n

separately (see e.g. the discussion after [62, Rem. 3.10]). Clearly, for this it suffices to show the M1 tightness of the processes

$$\Xi^n_\bullet \;:=\; c^{-1}\int_0^\bullet \sigma(s,X^n_{s-r})_-\;\mathrm{d}Z^n_s \;=\; c^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\sigma\Big(\frac{j}{n},\,X^n_{\frac{j}{n}-r}\Big)_-\;\Delta Z^n_{\frac{L_j}{n}}$$
$$=\;\frac{1}{cn^{\frac{\beta}{\alpha}}}\sum_{j=1}^{\lfloor nt\rfloor}\left(\sigma\Big(\frac{j}{n},\,X^n_{\frac{j}{n}-r}\Big)_-\;\sum_{i=0}^{\mathcal{J}}c_i\,\theta_{j-i}\right)$$

as the the tightness of the drift term follows directly on behalf of (5.3). We will now show that for fixed $T>0$ and for any $\eta>0$ it holds

$$\lim_{\lambda\searrow 0}\;\limsup_{n\to\infty}\;\mathbb{P}\left(w''(\Xi^n,\lambda)\;>\;\eta\right)\;=\;0$$

where w'' is the M1 modulus of continuity defined as in [62, Def. A.7]. However, according to [2, Corollary 1], for this it is enough to prove that for any fixed $0\le t_1<t<t_2$, $\eta>0$ and n large enough it holds

$$\mathbb{P}\left(\nu(\Xi^n,t_1,t,t_2)\;>\;\eta\right)\;\le\;C(\gamma,\beta)\,\eta^{-\gamma}\,(t_2-t_1)^{1+\beta} \tag{6.33}$$

for some constants $\gamma>0$, $\beta\ge 0$ and $C(\gamma,\beta)>0$, and $\nu(\Xi^n,t_1,t,t_2):=\big\|\Xi^n_t-[\Xi^n_{t_1},\Xi^n_{t_2}]\big\|:=\inf\{|\Xi^n_t-\alpha\Xi^n_{t_1}-(1-\alpha)\Xi^n_{t_2}|:\alpha\in[0,1]\}$ as well as

$$\lim_{\lambda\searrow 0}\;\limsup_{n\to\infty}\;\mathbb{P}\bigg(\sup_{0\le s,t\le\lambda}\;|\,\Xi^n_t\;-\;\Xi^n_s|\;>\;\eta\bigg)\;=\;0. \tag{6.34}$$

Similar to the proof of Proposition 4.10 and Theorem 4.11, we can now establish that for any $m,k\ge 0$ and large enough $n\ge 1$ there is a bound of the type

$$\mathbb{P}\bigg(\sup_{1\le i\le k}\;\Big|\Xi^n_{\frac{m+i}{n}}\;-\;\Xi^n_{\frac{m}{n}}\Big|\;>\;\eta\bigg)\;\le\;C\,\eta^{\gamma}\,\frac{k}{n} \tag{6.35}$$

with $C,\gamma>0$ constants which do not depend on n or k. From this, we can easily deduce (6.34). In addition, it will also be the key ingredient for establishing (6.33). To this end, fix $0\le t_1<t<t_2\le T$ as well as $\eta>0$ and $n\ge 1$ large enough. Without loss of generality assume that $t_1\in[\ell/n,(\ell+1)/n)$, $t\in[r/n,(r+1)/n)$ and $t_2\in[p/n,(p+1)/n)$ for some $\ell,r,p\ge 1$ with $\ell<r<p$. If Ξ^n_t lies in the interval with endpoints $\Xi^n_{t_1}$ and $\Xi^n_{t_2}$, then we have $\|\Xi^n_t-[\Xi^n_{t_1},\,\Xi^n_{t_2}]\|=0$. Therefore,

$$\begin{aligned}
\mathbb{P}\left(\nu(\Xi^n,t_1,t,t_2)\;>\;\eta\right)\;&=\;\mathbb{P}\left(\Xi^n_{t_1}-\Xi^n_t>\eta\,,\,\Xi^n_{t_2}\ge\Xi^n_{t_1}\right)\;+\;\mathbb{P}\left(\Xi^n_t-\Xi^n_{t_1}>\eta\,,\,\Xi^n_{t_2}\le\Xi^n_{t_1}\right)\\
&\quad+\;\mathbb{P}\left(\Xi^n_{t_2}-\Xi^n_t>\eta\,,\,\Xi^n_{t_2}<\Xi^n_{t_1}\right)\;+\;\mathbb{P}\left(\Xi^n_t-\Xi^n_{t_2}>\eta\,,\,\Xi^n_{t_2}>\Xi^n_{t_1}\right)\\
&\le\;2\mathbb{P}\left(|\Xi^n_{t_1}-\Xi^n_t|>\eta\right)\;+\;2\mathbb{P}\left(|\Xi^n_{t_2}-\Xi^n_t|>\eta\right)\\
&\le\;2\mathbb{P}\bigg(\Big|\Xi^n_{\frac{r}{n}}\;-\;\Xi^n_{\frac{\ell}{n}}\Big|>\eta\bigg)\;+\;2\mathbb{P}\bigg(\Big|\Xi^n_{\frac{p}{n}}\;-\;\Xi^n_{\frac{\ell}{n}}\Big|>\eta\bigg)\\
&\le\;2\mathbb{P}\bigg(\sup_{1\le i\le r-\ell}\;\Big|\Xi^n_{\frac{\ell+i}{n}}\;-\;\Xi^n_{\frac{\ell}{n}}\Big|>\eta\bigg)\;+\;2\mathbb{P}\bigg(\sup_{1\le i\le p-r}\;\Big|\Xi^n_{\frac{r+i}{n}}\;-\;\Xi^n_{\frac{r}{n}}\Big|>\eta\bigg)\\
&\le\;2C\,\eta^{\gamma}\,\frac{(r-\ell)+(p-r)}{n}\;\;\le\;\;4C\,\eta^{\gamma}\frac{p-(\ell+1)}{n}\;\;\le\;\;4C\,\eta^{\gamma}(t_2-t_1)
\end{aligned}$$

where we have used (6.35). Finally, the tightness of the running supremum of Ξ^n on compacts follows immediately from (6.35), so we obtain the tightness of the sequence Ξ^n. □

We shall point out that the weak relative compactness of the X^n was only needed to deduce the conditions 1-3 in the above proof. Therefore, for different sequences of drivers or different classes of differential equations it may indeed be more tractable to directly attempt to show these conditions. In particular, it is worth noting that, if these conditions are satisfied, then the limiting integrand (with respect to convergence of finite-dimensional distributions) admits a càdlàg modification and therefore, without loss of generality, we can assume it to be càdlàg. We comment on this in the following remark.

Remark 6.3 (Càdlàg limits)**.** Let $(X^n)_{n\geq 1}$ be a sequence of stochastic processes taking values in $\mathbf{D}_{\mathbb{R}^d}[0,T]$ and let $\Lambda \subset [0,T]$ (which includes T) be a dense subset of $[0,T]$. Suppose that all finite-dimensional distributions along Λ of the X^n converge to those of some limit Y. Then, in all generality, Y obviously does not possess a càdlàg modification. However, if we assume that for every $\delta > 0$ the maximal number of δ-increments $(N^T_\delta(X^n))_{n\geq 1}$ is tight (where N^T_δ is defined in [62, (3.9)]), this becomes true. To see this, it suffices to consider a deterministic counterexample: let

$$x_n := \mathbf{1}_{[1\,,\,1+\frac{1}{n})} + \sum_{k=1}^{n-1} (-1)^k \; \mathbf{1}_{[1+\frac{1}{k+1}\,,\,1+\frac{1}{k})} \quad \text{and} \quad x := \sum_{k=1}^{\infty} (-1)^k \; \mathbf{1}_{[1+\frac{1}{k+1}\,,\,1+\frac{1}{k})}$$

and note that x_n converges pointwise on $[0,2]\setminus\{1\}$ to x, yet x cannot be made càdlàg at 1 without loosing pointwise convergence on a dense subset of $[0,2]$. Clearly, in this example we see that it is the exploding number of large oscillations of the sequence $(x_n)_{n\geq 1}$ which causes this problem. If we assume that for every $\delta > 0$, the maximal number of δ-increments $(N^T_\delta(X^n))_{n\geq 1}$ is tight, then this is no longer an issue. Let $Q \subseteq \Lambda$ be countable and dense in $[0,T]$. Obviously, the inclusion

$$\left\{ \exists\, t \in [0,T) \text{ such that } \lim_{\substack{q \downarrow\downarrow t \\ q\in Q}} Y_q \text{ or } \lim_{\substack{q \uparrow\uparrow t \\ q\in Q}} Y_q \text{ do not exist} \right\} \subseteq \bigcup_{k=1}^{\infty} \left\{ \bar{N}^Q_{\frac{1}{k}}(Y) = \infty \right\} \tag{6.36}$$

holds true, where $\bar{N}^Q_\delta(Y)$ is the maximal number of δ-increments of Y on $Q \subseteq [0,T]$. If we can show that each of the sets on the right of (6.36) has probability zero, we can define

$$\tilde{Y}_t \;:=\; \begin{cases} Y_t, & t \in Q \\ \lim_{q\downarrow\downarrow t,\, q\in Q} Y_q, & t \in [0,T]\setminus Q. \end{cases} \tag{6.37}$$

Y has left and right limits and $\tilde{Y}$ is càdlàg (where the latter follows from the fact that Y is right-continuous in probability restricted to Q, which can be proven similarly to the subsequent part). In order to do so, fix $\delta > 0$ and let $\{q_m : m \geq 1\}$ be an enumeration of the set Q. Note that $\{\bar{N}^Q_\delta(Y) = \infty\}$ is a subset of

$$\bigcap_{K\geq 1} \; \bigcup_{m=K+1}^{\infty} \; \bigcup_{\substack{q_{i_1} \leq q_{i_2} \leq \ldots \leq q_{i_K} \\ \{i_1,\ldots,i_K\} \subseteq \{1,\ldots,m\}}} \left\{ |Y_{q_{i_{j+1}}} - Y_{q_{i_j}}| \;>\; \delta \;,\; \forall j = 1,\ldots,K-1 \right\}$$

where the last union on the right-hand side can be written as $\{(Y_{q_1},\ldots,Y_{q_m}) \in A_{m,K}\}$ for open subsets $A_{m,K} \subseteq \mathbb{R}^n$. Thus, by continuity and monotonicity of the probability measure, it suffices for our purposes to find, for every $\varepsilon > 0$, a $K_\varepsilon \geq 1$ such that

$$\mathbb{P}\left((Y_{q_1},\ldots,Y_{q_m}) \;\in\; A_{m,K_\varepsilon} \right) \;\leq\; \varepsilon, \quad \text{for all } \; m \geq K_\varepsilon.$$

This follows by Portmanteau's Theorem, due to the convergence of the finite-dimensional distributions and the tightness condition imposed on $(N^T_\delta(X^n))_{n\geq 1}$. Finally, we note that enlarging the set to Q by any countable subset of Λ does not change the process $t \mapsto \lim_{q \downarrow\downarrow t, q \in Q} Y_q$ (up to a set of probability zero). Therefore, once we establish that $\{t : \mathbb{P}(\tilde{Y}_t \neq Y_t) > 0\}$ is countable, (6.37) yields a càdlàg modification of Y. Indeed, this holds true as for each $m \geq 1$, the set $\{t : \mathbb{P}(|Y_t - \tilde{Y}_t| > 1/m) > 1/m\}$ must be finite, otherwise there existed a sequence t_k, $k \geq 1$ with this property and enlarging Q by this sequence would produce a contradiction to Y being almost surely right-continuous at all points of this enlarged Q.

Proof of Corollary 5.3. Define the processes

$$\Lambda_t^n \;:=\; c^{-1} \int_0^t \tilde{\sigma}(s, X^n_{[s-r,s]})_- \;\mathrm{d}Z_s^n,$$

where $t \mapsto \tilde{\sigma}(t, X_{[t-r,t]})$ is Lipschitz continuous on compacts (a.s.). Rewriting $c\Lambda_t^n$ as

$$\begin{aligned}\sum_{k=1}^{\lfloor nt \rfloor} \Gamma^n_{\frac{k}{n}-} \left(Z^n_{\frac{k}{n}} - Z^n_{\frac{k-1}{n}} \right) \;&=\; \sum_{k=1}^{\lfloor nt \rfloor} \left(\Gamma^n_{\frac{k}{n}} \, Z^n_{\frac{k}{n}} \;-\; \Gamma^n_{\frac{k-1}{n}} \, Z^n_{\frac{k-1}{n}} \right) \;+\; \sum_{k=1}^{\lfloor nt \rfloor} \left(\Gamma^n_{\frac{k-1}{n}} \;-\; \Gamma^n_{\frac{k}{n}} \right) Z^n_{\frac{k-1}{n}} \\ &=\; \Gamma^n_{\frac{\lfloor nt \rfloor}{n}} \, Z^n_{\frac{\lfloor nt \rfloor}{n}} \;+\; \sum_{k=1}^{\lfloor nt \rfloor} \left(\Gamma^n_{\frac{k-1}{n}} \;-\; \Gamma^n_{\frac{k}{n}} \right) Z^n_{\frac{k-1}{n}}\end{aligned}$$

with $\Gamma_t^n := \tilde{\sigma}(t, X^n_{[t-r,t]})$, it is straightforward to obtain the tightness of $(|\Lambda^n|_T^*)_{n\geq 1}$, $(|\Gamma^n|_T^*)_{n\geq 1}$ and $(N_\delta^T(\Lambda^n))_{n\geq 1}$, $(N_\delta^T(\Gamma^n))_{n\geq 1}$ for any $T, \delta > 0$ (where N_δ^T is again defined as in [62, (3.9)]) by using the tightness of the Z^n as well as the boundedness condition and the Lipschitz continuity of the Γ^n. Proceeding as in the second part of the proof of Theorem 5.2, we obtain the tightness of the Ξ^n on $(\mathbf{D}_\mathbb{R}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ and thus, in particular, the tightness of $(|\Xi^n|_T^*)_{n\geq 1}$ and $(N_\delta^T(\Xi^n))_{n\geq 1}$ for any $T, \delta > 0$. Hence, we deduce the tightness of $(|X^n|_T^*)_{n\geq 1}$ and, since $N_\delta^T(x+y) \leq N_{\delta/2}^T(x) + N_{\delta/2}^T(y)$, also the tightness of $(N_\delta^T(X^n))_{n\geq 1}$. On behalf of a diagonal sequence argument, we can identify càdlàg processes $Y, \tilde{Y}$ and a subsequence $(c^{-1}Z^n, X^n, \Gamma^n)_{n\geq 1}$ such that

$$(c^{-1}Z^n, X^n_{t_1}, ..., X^n_{t_\ell}, \Gamma^n_{t_1}, ..., \Gamma^n_{t_\ell}) \;\Rightarrow\; (Z, Y_{t_1}, ..., Y_{t_\ell}, \tilde{Y}_{t_1}, ..., \tilde{Y}_{t_\ell})$$

on $(\mathbf{D}_\mathbb{R}[0,\infty), \mathrm{d}_{\mathrm{M1}}) \times (\mathbb{R}^{2\ell}, |\cdot|)$ for all $\ell \geq 1$ and $t_1, ..., t_\ell$ in some countable dense subset of $[0,\infty)$, where the càdlàg property of Y, $\tilde{Y}$ is due to Remark 6.3. Since the Γ^n are Lipschitz continuous with a constant independent of n, they are equicontinuous and converge weakly in $(\mathbf{C}_\mathbb{R}[0,\infty), |\cdot|_\infty^*)$ on the basis of the Arzèla-Ascoli Theorem, which even implies that $\tilde{Y}$ can be chosen almost surely continuous. Thus, X^n and Γ^n satisfy the conditions of [62, Prop. 3.22] and so does $\sigma(\bullet, X^n_{\bullet - r})$. As shown in the proof of Theorem 5.2, $(\sigma(\bullet, X^n_{\bullet-r}), c^{-1}Z^n)$ meet the conditions of 4.16 with respect to Theorem 4.15 while $(\Gamma^n, c^{-1}Z^n)$ satisfy the assumptions of 4.16 with respect to Theorem 4.14 (where the required (AVCI) conditions are immediate from the continuity of $\tilde{Y}$ and Proposition 1). Thus, we deduce

$$\begin{aligned}X^n \;&=\; \int_0^\bullet b(s, X^n_{s-r}) \;\mathrm{d}s \;+\; c^{-1} \int_0^\bullet \sigma(s, X^n_{s-r}) + \Gamma^n_s \;\mathrm{d}Z^n_s \\ &\Rightarrow\; \int_0^\bullet b(s, Y_{s-r}) \;\mathrm{d}s \;+\; \int_0^\bullet \sigma(s, Y_{s-r}) + \tilde{Y}_s \;\mathrm{d}Z_s \qquad (6.38)\end{aligned}$$

and hence $X^n \Rightarrow Y$ on $(\mathbf{D}_\mathbb{R}[0,\infty), \mathrm{d}_{\mathrm{M1}})$. The map $x \mapsto \tilde{\sigma}(\bullet, x_{[\bullet - r, \bullet]})$ from $(\mathbf{D}_\mathbb{R}[0,\infty), \mathrm{d}_{\mathrm{M1}})$ into $(\mathbf{C}_\mathbb{R}[0,\infty), |\cdot|_\infty^*)$ is continuous and thus the continuous mapping theorem yields $\tilde{\sigma}(\bullet, X^n_{[\bullet-r,\bullet]}) \Rightarrow \tilde{\sigma}(\bullet, Y_{[\bullet-r,\bullet]})$ on $(\mathbf{C}_{\mathbb{R}^d}[0,\infty), |\cdot|_\infty^*)$. Since weak limits are unique, we obtain $\tilde{Y} = \sigma(\bullet, Y_{[\bullet-r,\bullet]})$. Therefore, (6.38) becomes

$$X^n \;\Rightarrow\; \int_0^\bullet b(s, Y_{s-r}) \;\mathrm{d}s \;+\; \int_0^\bullet \sigma(s, Y_{s-r}) + \sigma(s, Y_{[s-r,s]}) \;\mathrm{d}Z_s.$$

where the right-hand side is a solution of $(\tilde{\mathrm{S}})$. Again by uniqueness of weak limits, Y thus realises a solution of $(\tilde{\mathrm{S}})$. If $(\tilde{\mathrm{S}})$ is unique in law, then Y is the unique solution. $\square$

A. Appendix: Proofs of auxiliary results from Section 3.2

In this appendix, we give the proofs of Propositions 3.4 and 3.5 which went into the proof of Theorem 3.3 on the good decompositions (GD) of uncorrelated CTRWs.

Proof of Proposition 3.4. Let $n \geq 1$ and $0 \leq s \leq t$. To show integrability of X_t^n, we shall apply Wald's identity [13, Thm. 4.1.5] (where the required discrete filtration will be $\{\sigma((\zeta_k^n, J_k) : 1 \leq k \leq m)\}_{m\geq 1}$, $N(nt)$ the stopping time, and the summands are the $\zeta_k^n \mathbf{1}_{\{|\zeta_k^n|\leq a\}}$). First, we establish the integrability of $N(nt)$. Note that

$$0 \leq N(nt) \leq \tilde{N}_b(nt) := \max\Big\{m \geq 0 \;:\; \sum_{k=1}^m b\,\mathbf{1}_{\{J_k>b\}} \;\leq\; nt\Big\} \;=\; \sum_{m=1}^{\infty} \mathbf{1}_{\left\{\sum_{k=1}^m \mathbf{1}_{\{J_k>b\}} \leq \frac{nt}{b}\right\}}$$

for any $b \geq 0$. Choose $b > nt$ and note that then

$$\mathbb{E}[\tilde{N}_b(nt)] \leq \sum_{m=1}^{\infty} \mathbb{P}\Big(\sum_{k=1}^m \mathbf{1}_{\{J_k>b\}} \;<\; 1\Big) = \sum_{m=1}^{\infty} \mathbb{P}\Big(\bigcap_{k=1}^m \{J_k \leq b\}\Big) = \sum_{m=1}^{\infty} \mathbb{P}(J_1 \leq b)^m < \infty$$

since $\mathbb{P}(J_1 \leq b) < 1$ as otherwise $\mathbb{E}[J_1] \neq \infty$. Hence, $0 \leq \mathbb{E}[N(nt)] \leq \mathbb{E}[\tilde{N}_b(nt)] < \infty$. Therefore, we can apply Wald's theorem to obtain

$$\mathbb{E}[|M_t^n|] \leq \mathbb{E}\Bigg[\sum_{k=1}^{N(nt)} |\zeta_k^n|\,\mathbf{1}_{\{|\zeta_k^n|\leq a\}}\Bigg] + \mathbb{E}[N(nt)]\;\mathbb{E}[\zeta_1^n\,\mathbf{1}_{\{|\zeta_1^n|\leq a\}}] \leq 2a\mathbb{E}[N(nt)] < \infty$$

Clearly, M^n is adapted to the filtration $(\mathcal{F}_t^n)_{t\geq 0}$ and $\mathbb{E}[M_t^n - M_s^n \,|\, \mathcal{F}_s^n]$ equals

$$\begin{aligned}
&\mathbb{E}\Bigg[\sum_{k=N(ns)+1}^{N(nt)} \zeta_k^n\,\mathbf{1}_{\{|\zeta_k^n|\leq a\}} \;\Bigg|\; \mathcal{F}_s^n\Bigg] - \mathbb{E}[N(nt) - N(ns)\,|\,\mathcal{F}_s^n]\mathbb{E}[\zeta_1^n\,\mathbf{1}_{\{|\zeta_1^n|\leq a\}}] \\
&= \;\mathbb{E}\Bigg[\sum_{k=N(ns)+1}^{N(nt)} \zeta_k^n\,\mathbf{1}_{\{|\zeta_k^n|\leq a\}}\Bigg] \;-\; \mathbb{E}[N(nt) - N(ns)]\;\mathbb{E}[\zeta_1^n\,\mathbf{1}_{\{|\zeta_1^n|\leq a\}}],
\end{aligned}$$

as $\sum_{k=N(ns)+1}^{N(nt)} \zeta_k^n\,\mathbf{1}_{\{|\zeta_k^n|\leq a\}} = \sum_{k=1}^{N(nt)-N(ns)} \zeta_{N(ns)+k}^n\,\mathbf{1}_{\{|\zeta_{N(ns)+k}^n|\leq a\}}$ and $N(nt) - N(ns)$ are both independent of $\mathcal{F}_s^n$ by definition of the filtration and the pairwise independence of the pairs (ζ_k^n, J_k). Applying again Wald's identity yields the claim. The bound on the jumps of the M^n follows from $|\Delta N(nt)| \leq 1$ and the boundedness of the $\zeta_k^n\,\mathbf{1}_{\{|\zeta_k^n|\leq a\}}$. □

Towards the proof of Proposition 3.5, we first need some preliminary results.

Theorem A.1 (A partial version of [25, Thm. VII.2.9]). *Let $(\mu_n)_{n\geq 0}$ be infinitely divisible probability measures on $\mathbb{R}$ with characteristics (b_n, c_n, ν_n) – i.e. the characteristic function of μ_n is of the form $\varphi_{\mu_n}(u) = \exp(\psi_{b_n,c_n,\nu_n}(u))$ with*

$$\psi_{b_n,c_n,\nu_n}(u) \;=\; iub_n \;-\; \frac{1}{2}c_n u^2 \;+\; \int_{\mathbb{R}\setminus\{0\}} e^{iux} - 1 - iuh(x)\;\nu_n(\mathrm{d}x) \tag{A.1}$$

where $b_n \in \mathbb{R}$, $c_n \geq 0$ and ν_n a measure on $\mathbb{R}$ such that $\nu_n(\{0\}) = 0$ and $\int x^2 \wedge 1\,\nu_n(\mathrm{d}x) < \infty$. If $\mu_n \Rightarrow \mu_0$, then $b_n \to b_0$, $c_n \to c_0$ and $\langle g, \nu_n\rangle \to \langle g, \nu_0\rangle$ for all continuous, bounded functions g satisfying $f(x) = \mathcal{O}(x^2)$ for $x \to 0$.

Proof. The proof is conducted as in [25, Thm. VII.2.9], but we give it for completeness, noting that less work is required for our version. We definine a convenient transformation of ψ_{b_n,c_n,ν_n} in such a way that the summands in b_n and $h(x)$ vanish. More precisely, let

$$\tilde{\varphi}_n(u) \;:=\; \psi_{b_n,c_n,\nu_n}(u) \;-\; \frac{1}{2}\int_{-1}^{1} \psi_{b_n,c_n,\nu_n}(u+s)\,\mathrm{d}s$$

A short calculation with a simple application of Fubini's theorem gives us

$$\tilde{\varphi}_n(u) \;=\; \frac{1}{6}c_n \;+\; \int_{\mathbb{R}\setminus\{0\}} e^{iux}(1-x^{-1}\sin(x))\,\nu_n(\mathrm{d}x)$$

and therefore $\tilde{\varphi}_n$ is the characteristic function of the non-negative finite[1] measure $\tilde{\mu}_n(\mathrm{d}x) := (c_n/6)\,\delta_{\{0\}}(\mathrm{d}x) + (1-\sin(x)/x)\,\nu_n(\mathrm{d}x)$ with total mass $\tilde{\mu}_n(\mathbb{R}) = \tilde{\varphi}_n(0)$. Since $\mu_n \Rightarrow \mu_0$, it holds $\tilde{\varphi}_n(u) \to \tilde{\varphi}_0(u)$ for all $u \in \mathbb{R}$ due to its definition and the uniform convergence of the ψ_{b_n,c_n,ν_n} on compacts. This, in turn, implies $\tilde{\mu}_n \Rightarrow \tilde{\mu}_0$ as without loss of generality we may assume $\tilde{\mu}_n(\mathbb{R}) \neq 0$ and then the $\tilde{\varphi}_n(0)^{-1}\tilde{\mu}_n(\mathrm{d}x)$ are probability measures with characteristic function $\tilde{\varphi}_n(0)^{-1}\tilde{\varphi}_n$. Finally, for each continuous, bounded $g(x) = o(x^2)$ for $x \to 0$, we have that $x \mapsto g(x)/(1-x^{-1}\sin(x))\,\mathbf{1}_{\{x\neq 0\}} =: h(x)$ is continuous and bounded, and hence

$$\int g(x)\,\nu_n(\mathrm{d}x) \;=\; \int h(x)\,\tilde{\mu}_n(\mathrm{d}x) \;\xrightarrow[n\to\infty]{}\; \int h(x)\,\tilde{\mu}_0(\mathrm{d}x) \;=\; \int g(x)\,\nu_0(\mathrm{d}x)$$

which yields the last part of the claim. Then, since the real and imaginary part of $f(x) := e^{iux} - 1 - iuh(x)$ are both $o(x^2)$ as $x \to 0$, we get $\langle f,\nu_n\rangle \to \langle f,\nu_0\rangle$ and so, from (A.1),

$$iub_n \;-\; \frac{1}{2}c_n u^2 \;\xrightarrow[n\to\infty]{}\; iub_0 \;-\; \frac{1}{2}c_0 u^2$$

for all $u \in \mathbb{R}$, thus $b_n \to b_0$ and $c_n \to c_0$. $\square$

The next lemma formalises the fact that the random variables $Z^n := \zeta_1^n - \mathbb{E}[h(\zeta_1^n)]$ asymptotically possess very beneficial properties.

Lemma A.2. *Let $Z^n = \zeta_1^n - \mathbb{E}[h(\zeta_1^n)]$ and denote $\varphi_{Z^n}(u) = \mathbb{E}[e^{iuZ^n}]$ the characteristic function of Z^n. Then, it holds that*

(a) $\sup_{n\geq 1} n^\beta \mathbb{P}(|Z^n| > c) < \infty$ for all $c > 0$;

(b) $n^\beta \mathbb{E}[h(Z^n)] \to 0$ as $n \to \infty$.

(c) for all $\gamma > 0$, $\sup_{|u|<\gamma} n^\beta |\varphi_{Z^n}(u) - 1| \to 0$ as $n \to \infty$ and

Proof. Recall that $\mathbb{P}(|\zeta_1^n| > x) = \mathbb{P}(|\theta_1| > xn^{\beta/\alpha}) = \mathcal{O}(n^{-\beta}x^{-\alpha})$ for $x \to +\infty$. Now, for every $0 < \varepsilon < a$, we have $\mathbb{E}[h(|\zeta_1^n|)] \leq \varepsilon + a\mathbb{P}(|\zeta_1^n| > \varepsilon) \to \varepsilon$ as $n \to \infty$ and since $\varepsilon > 0$ was chosen arbitrary, we obtain $\mathbb{E}[h(|\zeta_1^n|)] \to 0$ as $n \to \infty$. Therefore, we deduce the first claim since $n^\beta \mathbb{P}(|Z^n| > c) \;\leq\; n^\beta \mathbb{P}(|\zeta_1^n| > c/2) + n^\beta \mathbb{P}(\mathbb{E}[h(\zeta_1^n)] > c/2)$ for all $\theta > 0$ and the second summand equals zero for large enough n. Hence, we even obtain $\limsup_{n\to\infty} n^\beta \mathbb{P}(|Z^n| > c) \leq (2/c)^\alpha$.

For (b), let $\varepsilon > 0$ and note that h satisfies $|h(x) - h(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$. Choose n large enough such that $\mathbb{E}[h(|\zeta_1^n|)] \leq (a/2) \wedge \varepsilon$. We have $\mathbb{E}[h(Z^n)] = \mathbb{E}[h(Z^n) - h(Z^n + \mathbb{E}[h(\zeta_1^n)]) + \mathbb{E}[h(\zeta_1^n)]]$ and the quantity inside the outer expectation is equal to zero if $|Z^n| \leq a/2$. Therefore, by the Lipschitz continuity of h,

$$|n^\beta \mathbb{E}[h(Z^n)]| \;\leq\; n^\beta \mathbb{E}[(\varepsilon + \mathbb{E}[|h(\zeta_1^n|)])\,\mathbf{1}_{\{|Z^n|>a/2\}}] \;\leq\; 2\varepsilon n^\beta \mathbb{P}(|Z^n| > a/2)$$

[1]The finiteness comes from the fact that $1 - \sin(x)/x = \mathcal{O}(x^2)$ as $x \to 0$ and the property of the ν_n that $\int x^2 \wedge 1\,\nu_n(\mathrm{d}x) < \infty$.

and the result follows from (a) above and the fact that $\varepsilon > 0$ was chosen arbitrary.

Finally, we are going to prove (c). Let $\gamma > 0$. Set $C_\gamma := \max\{\sqrt{2\gamma}, 2\}$ and note that $|e^{iux} - 1| \leq C_\gamma(|x| \wedge 1)$ for all $x \in \mathbb{R}$ and $|u| \leq \gamma$ (which we obtain by Euler's identity and a simple application of the mean value theorem). Therefore,

$$\sup_{|u|<\gamma} |\varphi_{Z^n}(u) - 1| \;\leq\; \sup_{|u|<\gamma} \mathbb{E}[\,|e^{iuZ^n} - 1|\,] \;\leq\; C_\gamma \mathbb{E}[|Z^n| \wedge 1] \;\leq\; C_\gamma \mathbb{E}[|Z^n| \wedge a]$$

Let χ be such that $\mathbb{P}(\chi = 1) = \mathbb{P}(\chi = -1) = 1/2$ and $\chi \perp\!\!\!\perp \zeta_1^n$. Then, by multiplying Z^n with χ we symmetrise Z^n without affecting the inequality, i.e.

$$\begin{aligned}\sup_{|u|<\gamma} |\varphi_{Z^n}(u) - 1| &\leq C_\gamma \mathbb{E}[|\chi Z^n| \wedge a] \\ &\leq C_\gamma \big(\mathbb{E}[h(\chi Z^n)\, \mathbf{1}_{\{\chi Z^n \geq 0\}}] - \mathbb{E}[h(\chi Z^n)\, \mathbf{1}_{\{\chi Z^n < 0\}}]\big)\end{aligned}$$

Note $\mathbb{E}[h(\chi Z^n)\, \mathbf{1}_{\{\chi Z^n \geq 0\}}] = \mathbb{E}[\{h(\chi Z^n) - h(\chi Z^n + \chi \mathbb{E}[h(\zeta_1^n)]) + \chi \mathbb{E}[h(\zeta_1^n)]\}\, \mathbf{1}_{\{\chi Z^n \geq 0\}}]$ since a simple calculation gives $\{\chi Z^n \geq 0\} \perp\!\!\!\perp \sigma(\zeta_1^n)$ and thus $\mathbb{E}[h(\chi \zeta_1^n)\, \mathbf{1}_{\{\chi Z^n \geq 0\}}] = \mathbb{E}[\chi \mathbb{E}[h(\zeta_1^n)]\, \mathbf{1}_{\{\chi Z^n \geq 0\}}]$. Now, proceeding similarly to (b), we find

$$|\mathbb{E}[h(\chi Z^n)\, \mathbf{1}_{\{\chi Z^n \geq 0\}}]| \;\leq\; 2\varepsilon \mathbb{P}(|\chi Z^n| > a/2) \;=\; 2\varepsilon \mathbb{P}(|Z^n| > a/2)$$

for n large enough such that $|\mathbb{E}[h(\zeta_1^n)]| \leq (a/2) \wedge \varepsilon$. Analogously, we obtain the same bound for $|\mathbb{E}[-h(\chi Z^n)\, \mathbf{1}_{\{\chi Z^n \geq 0\}}]|$. Hence, this gives us

$$\sup_{|u|<\gamma} n^\beta |\varphi_{Z^n}(u) - 1| \;\leq\; 4\varepsilon C_\gamma n^\beta \mathbb{P}(|Z^n| > a/2).$$

Due to (a) and since $\varepsilon > 0$ was chosen arbitrarily this yields (c). $\square$

Let $Y^n := \sum_{k=1}^{n^\beta} \zeta_k^n$ and recall $\mathcal{L}(Y^n) \to \mu$, where μ is the law of an α-stable Lévy process at time $t = 1$. Dnote the characteristics of μ by (b, c, ν). We can now prove Proposition 3.5.

Proof of Proposition 3.5. The proof follows very closely that of [25, Lem. VII.2.43], but has been simplified and adapted to our particular setting. Let $u \in \mathbb{R}$. Note that we can rewrite the characteristic function $\varphi_{X_1^n}$ by

$$\varphi_{Y^n}(u) \;=\; \mathbb{E}[\exp(iuY^n)] \;=\; \mathbb{E}[\exp(iu\,\zeta_1^n)]^{n^\beta} \;=\; \mathbb{E}[\exp(iu\,Z^n)]^{n^\beta}\; \exp\{iu\,n^\beta \mathbb{E}[h(\zeta_1^n)]\}$$

The idea is now to hope that the first factor tends to 1 and since we know that $\varphi_{Y^n} \to \varphi_\mu$ pointwise we can conclude that the second factor then has to converge to φ_μ. Finally, an application of Theorem A.1 with respect to the characteristics in the second factor should then yield the claim. However, at this point we do not have any knowledge on whether $\varphi_{Z^n}^{n^\beta}$ converges to 1 as $n \to \infty$. Yet, according to Lemma A.2 we do know about the asymptotic behaviour of $n^\beta |\varphi_{Z^n} - 1|$. Thus, we continue by adding a convenient factor,

$$\varphi_{Y^n}(u) \;=\; \Lambda_n(u)\, \exp\left\{iu\,n^\beta \mathbb{E}[h(\zeta_1^n)] + n^\beta \int e^{iuZ^n} - 1 - iu\,h(Z^n)\, \mathrm{d}\mathbb{P}\right\}$$

where Λ_n is defined as

$$\begin{aligned}\Lambda_n(u) \;&=\; \mathbb{E}[\exp(iu\,Z^n)]^{n^\beta} \exp\left\{-n^\beta \mathbb{E}\left[e^{iuZ^n} - 1 - iu\,h(Z^n)\right]\right\} \\ &=\; \big(\mathbb{E}[\exp(iu\,Z^n) - 1] + 1\big)^{n^\beta}\; e^{-n^\beta \mathbb{E}[\exp(iu\,Z^n) - 1]} \exp\left\{-iu\,n^\beta \mathbb{E}\left[h(Z^n)\right]\right\} \\ &=\; (\lambda_n + 1)^{n^\beta}\; e^{-n^\beta \lambda_n}\; \exp\left\{-iu\,n^\beta \mathbb{E}\left[h(Z^n)\right]\right\}\end{aligned}$$

and$\lambda_n := \mathbb{E}[\exp(iu\,Z^n) - 1]$. By Lemma A.2(b), $\exp\{-iu\,n^\beta \mathbb{E}[h(Z^n)]\} \to 1$ as $n \to \infty$. Hence, it only remains to show $(\lambda_n + 1)^{n^\beta}\; e^{-n^\beta \lambda_n} \to 1$: by Lemma A.2(c), $n^\beta |\lambda_n| \to 0$ as $n \to \infty$. Without loss of generality, let $n \geq 1$ large enough such that $|\lambda_n| \leq 1/2$. Note that the principal branch log of the complex logarithm has Taylor expansion $\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$ for all $|x| < 1$ and it is elementary to show that this gives ground to the inequality $|\log(1+x) - x| \leq 2|x|^2$ for all $|x| \leq 1/2$. Thus,

$$|n^\beta(\log(1+\lambda_n) - \lambda_n)| \;\leq\; 2n^\beta|\lambda_n|^2 \quad \xrightarrow[n\to\infty]{} \; 0$$

Finally, observe that therefore

$$(\lambda_n + 1)^{n^\beta} \, e^{-n^\beta \lambda_n} \;=\; \exp\{n^\beta \log(1+\lambda_n) - n^\beta \lambda_n\} \;\to\; 1$$

Thus, we obtain that $\Lambda_n(u) \to 1$ and eventually as $\varphi_{Y^n}(u) \to \varphi_\mu(u)$, this gives us

$$\begin{aligned}\exp\{\psi_{n^\beta \mathbb{E}[h(\zeta_1^n)]\,,\,0\,,\,n^\beta\,\mathbb{P}\circ(Z^n)^{-1}}(u)\} &= \exp\Big\{iu\, n^\beta \mathbb{E}[h(\zeta_1^n)] + n^\beta\!\int e^{iuZ^n} - 1 - iu\, h(Z^n)\;\mathrm{d}\mathbb{P}\Big\} \\ &\xrightarrow[n\to\infty]{} \;\; \varphi_\mu(u) = \exp\{\psi_{b,c,\nu}(u)\}\end{aligned}$$

which, by Theorem A.1, implies in particular $n^\beta \mathbb{E}[h(\zeta_1^n)] \to b$. □

Acknowledgments. The research of FW was funded by the EPSRC grant EP/S023925/1.

References

[1] Y. AÏT-SAHALIA and J. JACOD. "From tick data to semimartingales". In: *Ann. Appl. Probab.* 30.6 (2020), pp. 2740–2768.

[2] F. AVRAM and M. S. TAQQU. "Probability bounds for M-Skorohod oscillations". In: *Stoch. Process. Their Appl.* 33.1 (1989), pp. 63–72.

[3] F. AVRAM and M. S. TAQQU. "Weak convergence of sums of moving averages in the α-stable domain of attraction". In: *Ann. Probab.* 20.1 (1992), pp. 483–503.

[4] B. BAEUMER, M. M. MEERSCHAERT, and E. NANE. "Brownian subordinators and fractional Cauchy problems". In: *Trans. Amer. Math. Soc.* 361.7 (2009), pp. 3915–3930.

[5] B. v. BAHR and C.-G. ESSEEN. "Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$". In: *Ann. Math. Stat.* 36.1 (1965), pp. 299–303.

[6] P. BECKER-KERN, M. M. MEERSCHAERT, and H.-P. SCHEFFLER. "Limit theorems for coupled continuous time random walks". In: *Ann. Probab.* 32.1B (2004), pp. 730–756.

[7] Q. BERGER. "Notes on Random Walks in the Cauchy Domain of Attraction". In: *Probab. Theory Relat. Fields* 175 (2019), pp. 1–44.

[8] P. BILLINGSLEY. *Convergence of probability measures.* 2nd ed. John Wiley & Sons, 1999.

[9] M. BURR. "Weak convergence of stochastic integrals driven by continuous-time random walks". In: *arXiv: 1110.0216* (2011).

[10] P. CARR and L. WU. "Time-changed Lévy processes and option pricing". In: *J. Financ. Econ.* 71.1 (2004), pp. 113–141.

[11] C. CONSTANTINESCU, R. LOEFFEN, and P. PATIE. "First passage times over stochastic boundaries for subdiffusive processes". In: *Trans. Amer. Math. Soc.* 375.3 (2022), pp. 1629–1652.

[12] D. DUFFIE and P. E. PROTTER. "From discrete- to continuous-time finance: weak convergence of the financial gain process". In: *Math. Finance* 2.1 (1992), pp. 1–15.

[13] R. DURRETT. *Probability: theory and examples.* 4th ed. Cambridge University Press, 2010.

[14] I. ELIAZAR and J. KLAFTER. "Correlation cascades of Lévy-driven random processes". In: *Physica A: Statistical Mechanics and its Applications* 376 (2007), pp. 1–26.

[15] I. ELIAZAR and J. KLAFTER. “Fractal Lévy correlation cascades”. In: *J. Phys. A. Math. Theor.* 40.16 (2007), F307–F314.

[16] W. FELLER. *An Introduction to Probability Theory and Its Applications.* 2nd ed. Vol. 2. New York: John Wiley & Sons, 1991.

[17] J. GAJDA and M. MAGDZIARZ. “Fractional Fokker-Planck equation with tempered α-stable waiting times: Langevin picture and computer simulation”. In: *Phys. Rev. E (3)* 82.1 (2010), pp. 011117, 6.

[18] S. GEISS. *Sharp convex generalizations of stochastic Gronwall inequalities.* 2023. arXiv: 2112.05047.

[19] G. GERMANO et al. “Stochastic calculus for uncoupled continuous-time random walks”. In: *Phys. Rev. E (3)* 79.6 (2009), pp. 066102, 12.

[20] M. HAHN, K. KOBAYASHI, and S. UMAROV. “Fokker-Planck-Kolmogorov equations associated with time-changed fractional Brownian motion”. In: *Proc. Amer. Math. Soc.* 139.2 (2011), pp. 691–705.

[21] M. HAHN, K. KOBAYASHI, and S. UMAROV. “SDEs driven by a time-changed Lévy process and their associated time-fractional order pseudo-differential equations”. In: *J. Theor. Probab.* 25.1 (2012), pp. 262–279.

[22] M. HAHN et al. “On time-changed Gaussian processes and their associated Fokker-Planck-Kolmogorov equations”. In: *Electron. Commun. Probab.* 16 (2011), pp. 150–164.

[23] B. I. HENRY and P. STRAKA. “Lagging and leading coupled continuous time random walks, renewal times and their joint limits”. In: *Stoch. Process. Their Appl.* 121.2 (2011), pp. 324–336.

[24] S. HOTTOVY, A. McDANIEL, and J. WEHR. “A small delay and correlation time limit of stochastic differential delay equations with state-dependent colored noise”. In: *J. Stat. Phys.* 175.1 (2019), pp. 19–46.

[25] J. JACOD and A. N. SHIRYAEV. *Limit theorems for stochastic processes.* 2nd ed. Springer, 2002.

[26] A. JACQUIER and L. TORRICELLI. “Anomalous diffusions in option prices: connecting trade duration and the volatility term structure”. In: *SIAM J. Financ. Math.* 11.4 (2020), pp. 1137–1167.

[27] A. JAKUBOWSKI. “Convergence in various topologies for stochastic integrals driven by semimartingales”. In: *Ann. Probab.* 24.4 (1996), pp. 2141–2153.

[28] A. JAKUBOWSKI, J. MEMIN, and G. PAGES. “Convergence en loi des suites d’intégrales stochastiques sur l’espace D^1 de Skorokhod”. In: *Probab. Theory Relat. Fields.* 81 (1989), pp. 111–137.

[29] S. JIN and K. KOBAYASHI. “Strong approximation of stochastic differential equations driven by a time-changed Brownian motion with time-space-dependent coefficients”. In: *J. Math. Anal. Appl.* 476.2 (2019), pp. 619–636.

[30] S. JIN and K. KOBAYASHI. “Strong approximation of time-changed stochastic differential equations involving drifts with random and non-random integrators”. In: *BIT* 61.3 (2021), pp. 829–857.

[31] E. JUM and K. KOBAYASHI. “A strong and weak approximation scheme for stochastic differential equations driven by a time-changed Brownian motion”. In: *Probab. Math. Statist.* 36.2 (2016), pp. 201–220.

[32] A. JURLEWICZ et al. “Fractional governing equations for coupled random walks”. In: *Comput. Math. Appl.* 64.10 (2012), pp. 3021–3036.

[33] T. KAWATA. *Fourier analysis in probability theory.* Probability and mathematical statistics : a series of monographs and textbooks. Academic Press, 1972.

[34] I. KIM, K.-H. KIM, and S. LIM. “A Sobolev space theory for stochastic partial differential equations with time-fractional derivatives”. In: *Ann. Probab.* 47.4 (2019), pp. 2087–2139.

[35] K.-H. KIM and D. PARK. “A Sobolev Space Theory for Time-Fractional Stochastic Partial Differential Equations Driven by Lévy Processes”. In: *Journal of Theoretical Probability* Online first articles (2023), pp. 1–50.

[36] K. KOBAYASHI. “Stochastic calculus for a time-changed semimartingale and the associated stochastic differential equations”. In: *J. Theoret. Probab.* 24.3 (2011), pp. 789–820.

[37] T. G. KURTZ and P. E. PROTTER. "Weak limit theorems for stochastic integrals and stochastic differential equations". In: *Ann. Probab.* 19.3 (1991), pp. 1035–1070.

[38] N. N. LEONENKO et al. "Correlation structure of time-changed Lévy processes". In: *Commun. Appl. Ind. Math.* 6.1 (2014), e–483, 22.

[39] M. MAGDZIARZ. "Stochastic representation of subdiffusion processes with time-dependent drift". In: *Stochastic Process. Appl.* 119.10 (2009), pp. 3238–3252.

[40] M. MAGDZIARZ and R. L. SCHILLING. "Asymptotic properties of Brownian motion delayed by inverse subordinators". In: *Proc. Amer. Math. Soc.* 143.10 (2015), pp. 4485–4501.

[41] M. MAGDZIARZ and T. ZORAWIK. "Stochastic representation of a fractional subdiffusion equation. The case of infinitely divisible waiting times, Lévy noise and space-time-dependent coefficients". In: *Proc. Amer. Math. Soc.* 144.4 (2016), pp. 1767–1778.

[42] M. M. MEERSCHAERT, E. NANE, and P. VELLAISAMY. "Fractional Cauchy problems on bounded domains". In: *Ann. Probab.* 37.3 (2009), pp. 979–1007.

[43] M. M. MEERSCHAERT, E. NANE, and Y. XIAO. "Correlated continuous time random walks". In: *Stat. Probab. Lett.* 79.9 (2009), pp. 1194–1202.

[44] M. M. MEERSCHAERT and H.-P. SCHEFFLER. "Limit theorems for continuous-time random walks with infinite mean waiting times". In: *J. Appl. Probab.* 41 (2004), pp. 623–638.

[45] M. M. MEERSCHAERT and A. SIKORSKII. *Stochastic Models for Fractional Calculus*. Berlin, Boston: De Gruyter, 2012.

[46] R. METZLER, E. BARKAI, and J. KLAFTER. "Anomalous diffusion and relaxation close to thermal equilibrium: a fractional Fokker-Planck equation approach". In: *Phys. Rev. Lett.* 82 (18 1999), pp. 3563–3567.

[47] R. METZLER and J. KLAFTER. "The random walk's guide to anomalous diffusion: a fractional dynamics approach". In: *Phys. Rep.* 339.1 (2000), p. 77.

[48] R. METZLER and J. KLAFTER. "The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics". In: *J. Phys. A* 37.31 (2004), R161–R208.

[49] R. METZLER et al. "Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking". In: *Phys. Chem. Chem. Phys.* 16 (44 2014), pp. 24128–24164.

[50] E. NANE and Y. NI. "Path stability of stochastic differential equations driven by time-changed Lévy noises". In: *ALEA Lat. Am. J. Probab. Math. Stat.* 15.1 (2018), pp. 479–507.

[51] E. NANE and Y. NI. "Stability of the solution of stochastic differential equation driven by time-changed Lévy noise". In: *Proc. Amer. Math. Soc.* 145.7 (2017), pp. 3085–3104.

[52] V. PAULAUSKAS and S. T. RACHEV. "Cointegrated processes with infinite variance innovations". In: *Ann. Appl. Probab.* 8.3 (1998), pp. 775–792.

[53] V. PAULAUSKAS, S. T. RACHEV, and F. J. FABOZZI. "Comment on "Weak convergence to a matrix stochastic integral with stable processes"". In: *Econom. Theory* 27.4 (2011), pp. 907–911.

[54] J. PAULSEN. "On Cramér-like asymptotics for risk processes with stochastic return on investments". In: *Ann. Appl. Probab.* 12.4 (2002), pp. 1247–1260.

[55] J. PAULSEN and H. K. GJESSING. "Ruin theory with stochastic return on investments". In: *Adv. Appl. Probab.* 29.4 (1997), pp. 965–985.

[56] R. REBOLLEDO. "Central limit theorems for local martingales". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 51.3 (1980), pp. 269–286.

[57] E. SCALAS, R. GORENFLO, and F. MAINDARDI. "Fractional calculus and continuous-time finance". In: *Phys. A: Stat. Mech. Appl.* 284.1 (2000), pp. 376–384.

[58] E. SCALAS and N. VILES. "A functional limit theorem for stochastic integrals driven by a time-changed symmetric α-stable Lévy process". In: *Stoch. Process. Their Appl.* 124 (2014), pp. 385–410.

[59] A. V. SKOROKHOD. "Limit theorems for stochastic processes". In: *Theory Probab. Its Appl.* 1.3 (1956), pp. 261–290.

[60] L. SLOMINSKI. “Stability of strong solutions of stochastic differential equations”. In: *Stochastic Process. Appl.* 31.2 (1989), pp. 173–202.

[61] A. SØJMARK and F. WUNDERLICH. *Functional CLTs for subordinated Lévy models in physics, finance, and econometrics.* 2023. arXiv: 2312.15119.

[62] A. SØJMARK and F. WUNDERLICH. *Weak convergence of stochastic integrals on Skorokhod space in Skorokhod's J1 and M1 topologies.* 2023. arXiv: 2309.12197.

[63] S. UMAROV, M. HAHN, and K. KOBAYASHI. *Beyond the triangle: Brownian motion, Ito Calculus, and Fokker–Planck equation — fractional generalizations.* World Scientific, 2018.

[64] W. WHITT. *Stochastic-process limits. An introduction to stochastic-process limits and their application to queues.* Ed. by Peter W. Glynn and Stephen M. Robinson. Springer Series in Operations Research. Springer, 2002.

[65] H. YAOZHONG, S.-E. A. MOHAMMED, and F. YAN. “Discrete-time approximations of stochastic delay equations: the Milstein scheme”. In: *Ann. Probab.* 32.1A (2004), pp. 265–314.