The Annals of Probability
2017, Vol. 45, No. 4, 2707–2765
DOI: 10.1214/16-AOP1123

GENERAL ROUGH INTEGRATION, LÉVY ROUGH PATHS AND A LÉVY–KINTCHINE-TYPE FORMULA

BY PETER K. FRIZ[*,†,1] AND ATUL SHEKHAR[‡,2]

TU Berlin[*], *WIAS*[†] *and ISI Bangalore*[‡]

We consider rough paths with jumps. In particular, the analogue of Lyons' extension theorem and rough integration are established in a jump setting, offering a pathwise view on stochastic integration against càdlàg processes. A class of Lévy rough paths is introduced and characterized by a sub-ellipticity condition on the left-invariant diffusion vector fields and a certain integrability property of the Carnot–Caratheodory norm with respect to the Lévy measure on the group, using Hunt's framework of Lie group valued Lévy processes. Examples of Lévy rough paths include a standard multidimensional Lévy process enhanced with a stochastic area as constructed by D. Williams, the pure area Poisson process and Brownian motion in a magnetic field. An explicit formula for the expected signature is given.

CONTENTS

Received January 2015; revised June 2016.

[1]Supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Starting Grant, agreement nr. 258237, and the DFG research unit FOR2402.

[2]Affiliated with TU Berlin when this project commenced. Supported by the Berlin Mathematical School.

MSC2010 subject classifications. 60H99.

Key words and phrases. Young integration, rough paths, Lévy processes, general theory of processes.

1. Introduction and background.

1.1. *Motivation and contribution of this paper.* An important aspect of general theory of stochastic processes [19, 40, 41] is its ability to deal with jumps. On the other hand, the (deterministic) theory of rough paths [14–16, 28, 30, 31] has been very successful in dealing with *continuous* stochastic processes (and more recently random fields arising from SPDEs, e.g., [14] for references). It is a natural question to what extent there is a "general" rough path theory which can handle jumps and ultimately offers a (rough)pathwise view on stochastic integration against càdlàg processes. In the spirit of Marcus canonical equations (e.g., [1, 22]) related questions were first raised by Williams [46] and we will comment in more detail in Section 1.2.9 on his work and the relation to ours. We can also mention the pathwise works of Mikosch–Norvaiša [35] and Simon [43], although their works assumes Young regularity of sample paths (q-variation, $q < 2$) and thereby does not cover the "rough" regime ($q > 2$) of interest for general processes.

Postponing the exact definition of *general* (by convention: càdlàg) rough path, let us start with a list of desirable properties and natural questions.

- An analogue of Lyons' fundamental *extension theorem* (see Section 1.2.5) should hold true. That is, any general geometric p-rough path $\mathbf{X}$ should admit canonically defined higher iterated integrals, thereby yielding a group-like element (the *signature* of $\mathbf{X}$).
- A general rough path $\mathbf{X}$ should allow the integration of 1-forms, and more general suitable *controlled rough paths* Y à la Gubinelli, leading to rough integrals of the form

$$\int f(X^-)\,d\mathbf{X} \quad \text{and} \quad \int Y^-\,d\mathbf{X}.$$

- Every semimartingale $X = X(\omega)$ with (rough path) Itô-lift $\mathbf{X}^I = \mathbf{X}^I(\omega)$, should give rise to a (random) rough integral that coincides under reasonable assumptions with the Itô-integral, so that a.s.

$$(\text{Itô}) \int f(X^-)\,dX = \int f(X^-)\,d\mathbf{X}^I.$$

- As model case for both semimartingales and jump Markov process, what is the precisely rough path nature of Lévy processes? In particular, it would be desirable to have a class of *Lévy rough paths* that captures natural (but noncanonical...) examples such as the pure area Poisson process or the Brownian rough path in a magnetic field?
- To what extent can we compute the *expected signature* of such processes? And what do we get from it?

In essence, we will give reasonable answers to all these points. We have not tried to push for maximal generality. For instance, in the spirit of Friz–Hairer ([14], Chapters 3–5), we develop general rough integration only in the level 2 setting, which is what matters most for probability. But that said, the required algebraic and geometric picture to handle the level N-case is still needed in this paper, notably when we discuss the extension theorem and signatures. For the most, we have chosen to work with (both canonically and non-canonically lifted) Lévy processes as model case for random càdlàg rough paths, this choice being similar to choosing Brownian motion over continuous semimartingales. (Let us also point to recent work of Chevyrev [6] who develops this theme from a random walk point of view.) In the final chapter, we discuss some extensions, notably to Markov jump diffusions and some simple Gaussian examples.

In his landmark paper ([30], page 220), Lyons gave a long and visionary list of advantages (to a probabilist) of constructing stochastic objects in a pathwise fashion: stochastic flows, differential equations with boundary conditions, Stroock–Varadhan support theorem, stochastic anlysis for non-semimartingales, numerical algorithms for SDEs, robust stochastic filtering, stochastic PDE with spatial roughness. Many other applications have been added to this list since. (We do not attempt to give references; an up-to-date bibliography with many applications of the (continuous) rough path theory can be found, e.g., in [14].) The present work lays in particular the foundation to revisit many of these problems, but now allowing for systematic treatment of jumps. We also note that integration against general rough paths can be considered as a generalization of the Föllmer integral [10] and, to some extent, Karandikar [21] (see also Soner et al. [44][3]), but free of implicit semimartingale features.

[3]There is much renewed interest in these theories from a model independent finance point of view, for example, [38] and references therein.

1.2. *Preliminaries*.

NOTATION. Throughout the article, $a \lesssim b$ means the existence of a positive constant C, which may depend on fixed parameters, such that $a \leq Cb$. We write $a \asymp b$ for a two-sided estimate of the form $a/C \leq b \leq Cb$.

1.2.1. *General Young integration* [8, 47]. We briefly review Young's integration theory, without making any continuity assumptions. Consider a path $X : [0, T] \to \mathbb{R}^d$ of finite p-variation, that is,

$$\|X\|_{p\text{-var};[0,T]} := \left(\sup_{\mathcal{P}} \sum_{[s,t]\in P} |X_{s,t}|^p\right)^{1/p} < \infty$$

with $X_{s,t} = X_t - X_s$ and sup (here and later on) taken over all for finite partitions $\mathcal{P}$ of $[0, T]$. As is well known, such paths are *regulated* in the sense of admitting left- and right-limits. In particular, $X_t^- := \lim_{s\uparrow t} X_s$ is cáglád and $X_t^+ := \lim_{s\downarrow t} X_s$ càdlàg (by convention: $X_0^- \equiv X_0$, $X_T^+ \equiv X_T$). Let us write $X \in W^p([0, T])$ for the space of càdlàg path of finite p -variation. A generic cáglád path of finite q-variation is then given by Y^- for $Y \in W^q([0, T])$. Any such pair (X, Y^-) has no common points of discontinuity on the same side of a point and the Young integral of Y^- against X,

$$\int_0^T Y^- \, dX \equiv \int_0^T Y_r^- \, dX_r \equiv \int_0^T Y_{s-} \, dX_s,$$

is well defined (see below) provided $1/p + 1/q > 1$ (or $p < 2$, in case $p = q$). We need the following.

DEFINITION 1. Assume $S = S(\mathcal{P})$ is defined on the partitions of $[0, T]$ and takes values in some normed space:

(i) Convergence in *Refinement Riemann–Stieltjes* (RRS) sense: we say (RRS) $\lim_{|\mathcal{P}|\to 0} S(\mathcal{P}) = L$ if for every $\varepsilon > 0$ there exists $\mathcal{P}_0$ such that for every "refinement" $\mathcal{P} \supset \mathcal{P}_0$ one has $|S(\mathcal{P}) - L| < \varepsilon$.

(ii) Convergence in *Mesh Riemann–Stieltjes* (MRS) sense: we say (MRS) $\lim_{|\mathcal{P}|\to 0} S(\mathcal{P}) = L$ if for every $\varepsilon > 0$ there exists $\delta > 0$ s.t. $\forall \mathcal{P}$ with mesh $|\mathcal{P}| < \delta$, one has $|S(\mathcal{P}) - L| < \varepsilon$.

THEOREM 2 (Young). *If $X \in W^p$ and $Y \in W^q$ with $\frac{1}{p} + \frac{1}{q} > 1$, then the Young integral is given by*

$$\int_0^T Y^- \, dX := \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_s^- X_{s,t} = \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t}, \tag{1.2.1}$$

where both limit exist in (*RRS*) *sense*. *Moreover*, *Young's inequality holds in either form*

$$\left|\int_s^t Y^- \, dX - Y_s^- X_{s,t}\right| \lesssim \|Y^-\|_{q\text{-}var;[s,t]} \|X\|_{p\text{-}var;[s,t]}, \tag{1.2.2}$$

$$\left|\int_s^t Y^- \, dX - Y_s X_{s,t}\right| \lesssim \|Y\|_{q\text{-}var;[s,t]} \|X\|_{p\text{-}var;[s,t]}. \tag{1.2.3}$$

At last, *if* X, Y *are continuous* (*so that in particular* $Y^- \equiv Y$), *the defining limit of the Young integral exists in* (*MRS*) *sense*.

Everything is well known here, although we could not find the equality of the limits in (1.2.1) pointed out explicitly in the literature. The reader can find the proof in Proposition 25.

1.2.2. *General Itô stochastic integration* [19, 40, 41]. Subject to the usual conditions, any semimartingale $X = X(\omega)$ may (and will) be taken with càdlàg sample paths. A classical result of Monroe [36] allows to write any (real-valued) martingale as a time-change of Brownian motion. As an easy consequence, semimartingales inherit a.s. finite 2^+ variation of sample paths from Brownian sample paths. See [24] for much more in this direction, notably a quantification of $\|X\|_{p\text{-var};[0,T]}$ for any $p > 2$ in terms of a BDG inequality. Let now Y be another (càdlàg) semimartingale, so that Y^- is previsible. The Itô integral of Y^- against X is then well defined, and one has the following classical Riemann–Stieltjes type description.

THEOREM 3 (Itô). *The Itô integral of* Y^- *against* X *has the presentation*, *with* $t_i^n = \frac{iT}{2^n}$,

$$\int_0^T Y^- \, dX = \lim_n \sum_i Y_{t_{i-1}^n}^- X_{t_{i-1}^n, t_i^n} = \lim_n \sum_i Y_{t_{i-1}^n} X_{t_{i-1}^n, t_i^n}, \tag{1.2.4}$$

where the limits exists in probability, *uniformly in* T *over compacts*.

Again, this is well known but perhaps the equality of the limits in (1.2.4) which the reader can find in Protter [40], Chapter 2, Theorem 21.

1.2.3. *Marcus canonical integration* [1, 2, 22, 33, 34]. Real (classical) particles do not jump, but may move at extreme speed. In this spirit, transform $X \in W^p([0,T])$ into $\tilde{X} \in C^{p\text{-var}}([0,\tilde{T}])$, by "stretching" time whenever

$$X_t - X_{t-} \equiv \Delta X_t \neq 0,$$

followed by replacing the jump by a straight line connecting X_{s-} with X_s, say

$$[0,1] \ni \theta \mapsto X_{t-} + \theta \Delta_t X.$$

Implemented in a (càdlàg) semimartingale context, this leads to *Marcus canonical integration*

$$\int_0^T f(X) \diamond dX$$
$$(1.2.5) \qquad := \int_0^T f(X_{t-})\, dX_t + \frac{1}{2}\int_0^T f'(X_{t-})\, d[X, X]_t^c$$
$$+ \sum_{t\in(0,T]} \Delta_t X \left\{\int_0^1 f(X_{t-} + \theta\Delta_t X) - f(X_{t-})\right\} d\theta.$$

(*Young canonical integration*, provided $p < 2$ and $f \in C^1$, is defined similarly, it suffices to omit the continuous quadratic variation term.) A useful consequence, for $f \in C^3(\mathbb{R}^d)$, say, is the chain rule

$$\int_0^t \partial_i f(X) \diamond dX^i = f(X_t) - f(X_0).$$

It is also possible to implement this idea in the context of SDEs,

$$(1.2.6) \qquad dZ_t = f(Z_t) \diamond dX_t$$

for $f : \mathbb{R}^d \to \mathbb{R}^{d\times k}$ where X is a semimartingale [22]. The precise meaning of this Marcus canonical equation is given by

$$\begin{aligned} Z_t &= Z_0 + \int_0^t f(Z_{s-})\, dX_s + \frac{1}{2}\int_0^t f' f(Z_s)\, d[X, X]_s^c \\ &\quad + \sum_{0<s\le t} \{\phi(f\Delta X_s, Z_{s-}) - Z_{s-} - f(Z_{s-})\Delta X_s\} \\ &= Z_0 + \int_0^t f(Z_{s-})\, dX_s + \frac{1}{2}\int_0^t f' f(Z_s)\, d[X, X]_s \\ &\quad + \sum_{0<s\le t} \left\{\phi(f\Delta X_s, Z_{s-}) - Z_{s-} - f(Z_{s-})\Delta X_s - f' f(Z_s)\frac{1}{2}(\Delta X_s)^{\otimes 2}\right\}, \end{aligned}$$

where $\phi(g, x)$ is the time 1 solution to $\dot{y} = g(y)$, $y(0) = x$. As one would expected from the aforementioned (first-order) chain rule, such SDEs respect the geometry.

THEOREM 4 ([22]). *If X is a càdlàg semimartingale and f and $f'f$ are globally Lipchitz, then solution to the Marcus canonical SDE* (1.2.6) *exists uniquely and it is a càdlàg semimartingale. Also, if M is manifold without boundary embedded in $\mathbb{R}^d$ and $\{f_i(x) : x \in M\}_{1\le i\le k}$ are vector fields on M, then*

$$\mathbb{P}(Z_0 \in M) = 1 \quad \Longrightarrow \quad \mathbb{P}(Z_t \in M\ \forall t \ge 0) = 1.$$

Flow properties of such SDEs were studied extensively by Kunita and coauthors; see, for example, [2].

1.2.4. *Continuous rough integration* [14, 16, 30]. Young integration of (continuous) paths has been the inspiration for the (continuous) rough integration, elements of which we now recall. Consider $p \in [2, 3)$ and $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^{p\text{-var}}([0, T])$ which in notation of [14] means validity of *Chen's relation*

$$\mathbb{X}_{s,u} = \mathbb{X}_{s,t} + \mathbb{X}_{t,u} + X_{s,t} \otimes X_{t,u} \tag{1.2.7}$$

and $\|\mathbf{X}\|_{p\text{-var}} := \|X\|_{p\text{-var}} + \|\mathbb{X}\|_{p/2\text{-var}}^{1/2} < \infty$, where

$$\|\mathbb{X}\|_{p/2\text{-var}} := \left(\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |\mathbb{X}_{s,t}|^{p/2}\right)^{2/p}.$$

For nice enough F (e.g., $F \in C^2$), both $Y_s := F(X_s)$ and $Y' := DF(X_s)$ are in $C^{p\text{-var}}$ and we have

$$\|R\|_{p/2\text{-var}} = \left(\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |R_{s,t}|^{p/2}\right)^{2/p} < \infty \\ \text{where } R_{s,t} := Y_{s,t} - Y'_s X_{s,t}. \tag{1.2.8}$$

THEOREM 5 (Lyons, Gubinelli). *Write $\mathcal{P}$ for finite partitions of* $[0, T]$. *Then*

$$\exists \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t} =: \int_0^T Y \, d\mathbf{X},$$

where the limit exists in (*MRS*) *sense*; *cf. Definition* 1.

Rough integration against $\mathbf{X}$ extends immediately to the integration of so-called *controlled rough paths*, that is, pairs (Y, Y'), with $Y, Y' \in C^{p\text{-var}}$ for which (1.2.8) holds. This gives meaning to a rough differential equation (RDE)

$$dY = f(Y) \, d\mathbf{X}$$

provided $f \in C^2$, say: A solution is simply a path Y such that $(Y, Y') := (Y, f(Y))$ satisfies (1.2.8) and such that the above RDE is satisfied in the (well-defined) integral sense, that is, for all $t \in [0, T]$,

$$Y_t - Y_0 = \int_0^t f(Y) \, d\mathbf{X}.$$

1.2.5. *Geometric rough paths and signatures* [15, 28, 30, 31]. A *geometric* rough path $\mathbf{X} = (X, \mathbb{X})$ is a rough path with $\mathrm{Sym}(\mathbb{X}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}$; and we write $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}_{\mathrm{g}}^{p\text{-var}}([0, T])$ accordingly. We work with generalized increments of the form $\mathbf{X}_{s,t} = (X_{s,t}, \mathbb{X}_{s,t})$ where we recall $X_{s,t} = X_t - X_s$ for path increments, while second-order increments $\mathbb{X}_{s,t}$ are determined from $(\mathbf{X}_{0,t})$ by Chen's relation

$$\mathbb{X}_{0,s} + X_{0,s} \otimes X_{s,t} + \mathbb{X}_{s,t} = \mathbb{X}_{0,t}.$$

Behind all this is the picture that $\mathbf{X}_{0,t} := (1, X_{0,t}, \mathbb{X}_{0,t})$ takes values in a Lie group $T_1^{(2)}(\mathbb{R}^d) \equiv \{1\} \oplus \mathbb{R}^d \oplus \mathbb{R}^{d\times d}$, embedded in the (truncated) tensor algebra $T^{(2)}(\mathbb{R}^d)$, and $\mathbf{X}_{s,t} = \mathbf{X}_{0,s}^{-1} \otimes \mathbf{X}_{0,t}$. From the usual power series in this tensor algebra, one defines for $a + b \in \mathbb{R}^d \oplus \mathbb{R}^{d\times d}$,

$$\log(1 + a + b) = a + b - \frac{1}{2} a \otimes a,$$

$$\exp(a + b) = 1 + a + b + \frac{1}{2} a \otimes a.$$

The linear space $\mathfrak{g}^{(2)}(\mathbb{R}^d) = \mathbb{R}^d \oplus \mathfrak{so}(d)$ is a Lie algebra under

$$\left[a + b, a' + b'\right] = a \otimes a' - a' \otimes a;$$

its exponential image $G^{(2)}(\mathbb{R}^d) := \exp(\mathfrak{g}^{(2)}(\mathbb{R}^d))$ is then a Lie (sub)group under

$$(1, a, b) \otimes \left(1, a', b'\right) = \left(1, a + a', b + a \otimes a' + b'\right).$$

At last, we recall that $G^{(2)}(\mathbb{R}^d)$ admits a so-called Carnot–Caratheodory norm (abbreviated as CC norm henceforth), with infimum taken over all curves $\gamma : [0, 1] \to \mathbb{R}^d$ of finite length L,

$$\begin{aligned}\|1 + a + b\|_{\mathrm{CC}} &:= \inf\left(L(\gamma) : \gamma_1 - \gamma_0 = a, \int_0^1 (\gamma_t - \gamma_0) \otimes d\gamma_t = b\right) \\ &\asymp |a| + |b|^{1/2} \\ &\asymp |a| + \left|\mathrm{Anti}(b)\right|^{1/2}.\end{aligned}$$

A left-invariant distance is induced by the group structure

$$d_{\mathrm{CC}}(g, h) = \left\|g^{-1} \otimes h\right\|_{\mathrm{CC}}$$

which turns $G^{(2)}(\mathbb{R}^d)$ into a Polish space. Geometric rough paths with roughness parameter $p \in [2, 3)$ are *precisely* classical paths of finite p-variation with values in this metric space.

PROPOSITION 6. *$\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}_g^{p\text{-var}}([0, T])$ iff $\mathbf{X} = (1, X, \mathbb{X}) \in C^{p\text{-var}}([0, T], G^{(2)}(\mathbb{R}^d))$. Moreover,*

$$\|\mathbf{X}\|_{p\text{-var}} \asymp \left(\sup_{\mathcal{P}} \sum_{[s,t]\in P} \|\mathbf{X}_{s,t}\|_{\mathrm{CC}}^p\right)^{1/p}.$$

The theory of geometric rough paths extends to all $p \geq 1$, and a geometric p-rough path is a path with values in $G^{([p])}(\mathbb{R}^d)$, the step-$[p]$ nilpotent Lie group with d generators, embedded in $T^{([p])}(\mathbb{R}^d)$, where

$$T^{(m)}(\mathbb{R}^d) = \bigoplus_{k=0}^{m} (\mathbb{R}^d)^{\otimes k} \subset \bigoplus_{k=0}^{\infty} (\mathbb{R}^d)^{\otimes k} \subset T\left((\mathbb{R}^d)\right)$$

(the last inclusion is strict, think polynomials versus power series) and again of finite p-variation with respect to the Carnot–Caratheodory distance (now defined on $G^{([p])}$).

THEOREM 7 (Lyons' extension). *Let $1 \le m := [p] \le p \le N < \infty$. A (continuous) geometric rough path $\mathbf{X}^{(m)} \in C^{p\text{-var}}([0,T], G^{(m)})$ admits an extension to a path $\mathbf{X}^{(N)}$ with values $G^{(N)} \subset T^{(N)}$, unique in the class of $G^{(N)}$-valued path starting from 1 and of finite p-variation with respect to CC metric on $G^{(N)}$. In fact,*

$$\|\mathbf{X}^{(N)}\|_{p\text{-var};[s,t]} \lesssim \|\mathbf{X}^{(m)}\|_{p\text{-var};[s,t]}.$$

REMARK 8. In view of this theorem, any $\mathbf{X} \in C^{p\text{-var}}([0,T], G^{(m)})$ may be regarded as $\mathbf{X} \in C^{p\text{-var}}([0,T], G^{(N)})$, any $N \ge m$, and there is no ambiguity in this notation.

DEFINITION 9. Write $\pi_{(N)}$ respectively π_M for the projection $T((\mathbb{R}^d)) \to T^{(N)}(\mathbb{R}^d)$, respectively, $(\mathbb{R}^d)^{\otimes M}$. Call $g \in T((\mathbb{R}^d))$ group-like, if $\pi_{(N)}(g) \in G^{(N)}$ for all N. Consider a geometric rough path $\mathbf{X} \in C^{p\text{-var}}([0,T], G^{[p]})$. Then thanks to the extension theorem,

$$S(\mathbf{X})_{0,T} := \big(1, \pi_1(\mathbf{X}_{0,T}), \dots, \pi_m(\mathbf{X}_{0,T}), \pi_{m+1}(\mathbf{X}_{0,T}), \dots\big) \in T\big((\mathbb{R}^d)\big)$$

defines a group-like element called the signature of $\mathbf{X}$.

The signature solves a rough differential equation (RDE, ODE if $p = 1$) in the tensor algebra,

$$dS = S \otimes d\mathbf{X}, \qquad S_0 = 1. \tag{1.2.9}$$

To a significant extent, the signature determines the underlying path X, if of bounded variation; cf. [17]. (The rough path case was recently obtained in [5].) A basic, yet immensely useful fact is that multiplication in $T((\mathbb{R}^d))$, if restricted to group-like elements, can be linearized.

PROPOSITION 10 (Shuffle product formula). *Consider two multi-indices $v = (i_1, \dots, i_m)$, $w = (j_1, \dots, j_n)$*

$$\mathbf{X}^v \mathbf{X}^w = \sum \mathbf{X}^z,$$

where the (finite) sum runs over all shuffles z of v, w.

1.2.6. *Checking p-variation* [12, 23, 32]. While the continuous rough path theory can be formulated in either p-variation or ($1/p$-) Hölder metrics, the general setting crucially relies on p-variation. We quickly discuss some (known) methods to establish this type of regularity and comment on their respective importance to the jump setting:

(i) As in whenever $\gamma > p-1 > 0$, with $t_k^n = k2^{-n}T$, one has

$$\|X\|_{p\text{-var};[0,T]}^p \lesssim \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} |X_{t_k^n} - X_{t_{k-1}^n}|^p. \tag{1.2.10}$$

This estimate immediately gives finite expectation (and hence finiteness a.s.) of p-variation of Brownian motion, provided $p > 2$, using $\mathbb{E}[|B_{s,t}|^p] \lesssim |t-s|^{p/2}$. Unfortunately, this argument does not work for jump processes. Even for the standard Poisson process one only has $\mathbb{E}[|N_{s,t}|^p] \sim C_p|t-s|$ as $t-s \to 0$, so that the *expected value* of the right-hand side of (1.2.10) is infinity. (It can, however, be seen ([46], page 318) that this quantity is finite a.s.) An extension of (1.2.10) to rough path is

$$\|\mathbf{X}\|_{p\text{-var};[0,T]}^p \lesssim \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \{|X_{t_{k-1}^n,t_k^n}|^p + |\mathbb{X}_{t_{k-1}^n,t_k^n}|^{p/2}\}$$

and we note that for a *geometric* rough path $\mathbf{X} = (X, \mathbb{X})$, that is, when $\mathrm{Sym}(\mathbb{X}_{s,t}) = \frac{1}{2}X_{s,t} \otimes X_{s,t}$, we may replace $\mathbb{X}$ on the right-hand side by the area $\mathbb{A} = \mathrm{Anti}(\mathbb{X})$. Thanks to moment estimates for area increments one can then obtain rough path regularity for Lévy processes [46]; see Theorem 11 for a precise statement.

(ii) In [12], an embedding result $W^{\delta,q} \hookrightarrow C^{p\text{-var}}$ is shown, more precisely

$$\|X\|_{p\text{-var};[0,T]}^q \lesssim \int_0^T \int_0^T \frac{|X_t - X_s|^q}{|t-s|^{1+\delta q}} \, ds \, dt,$$

provided $1 < p < q < \infty$ with $\delta = 1/p \in (0,1)$. The extension to rough paths reads

$$\|\mathbf{X}\|_{p\text{-var};[0,T]}^q \lesssim \int_0^T \int_0^T \left\{ \frac{|X_{s,t}|^q}{|t-s|^{1+\delta q}} + \frac{|\mathbb{X}_{s,t}|^{q/2}}{|t-s|^{1+\delta q}} \right\} ds \, dt.$$

Since elements in $W^{\delta,q}$ are also α-Hölder, with $\alpha = \delta - 1/q > 0$, these embeddings are not suitable for noncontinuous paths. In fact, this suggests that a Besov-space based rough paths theory will also be unable to deal with jumps, as was also seen in [39].

(iii) In case of a strong Markov process X with values in some Polish space (E,d), a powerful criterion has been established by Manstavicius [32]. Define

$$\alpha(h,a) := \sup\{\mathbb{P}(d(X_t^{s,x}, x) \geq a)\}$$

with sup taken over all $x \in E$, and $s < t$ in $[0,T]$ with $t - s \leq h$. Under the assumption

$$\alpha(h,a) \lesssim \frac{h^{\beta}}{a^{\gamma}},$$

uniformly for h, a in a right neighbourhood of zero, the process X has finite p-variation for any $p > \gamma/\beta$. In the above Poisson example, noting $\mathbb{E}[|N_{s,t}|] = O(h)$ whenever $t - s \le h$, Chebychev inequality immediately gives $\alpha(h, a) \le h/a$, and we find finite p-variation, for any $p > 1$. (Of course $p = 1$ here, but one should not expect this borderline case from a general criterion.) The Manstavicius criterion will play an important role for us.

1.2.7. *Expected signatures* [7, 9, 17, 25]. Recall that for a smooth path $X : [0, T] \to \mathbb{R}^d$, its signature $\mathbf{S} = \mathbf{S}(X)$ is given by the group-like element

$$\left(1, \int_{0<t_1<T} dX_{t_1}, \int_{0<t_1<t_2<T} dX_{t_1} \otimes dX_{t_2}, \dots\right) \in T\big((\mathbb{R}^d)\big).$$

The signature solves an ODE in the tensor algebra,

$$d\mathbf{S} = \mathbf{S} \otimes dX, \qquad \mathbf{S}_0 = \mathbf{1}. \tag{1.2.11}$$

Generalizations to semimartingales are immediate, by interpretation of (1.2.11) as in the Itô, Stratonovich or Marcus stochastic differential equation. In the same spirit, X can be replaced by a generic (continuous) geometric rough path; (1.2.11) is then regarded as a (linear) rough differential equation.

Whenever $\mathbf{S}_T = \mathbf{S}_T(\omega)$ has sufficient integrability, we may consider the expected signature, that is,

$$\mathbb{E}\mathbf{S}_T \in T\big((\mathbb{R}^d)\big)$$

defined in the obvious componentwise fashion. To a significant extent, this object behaves like a moment generating function. In a recent work [7], it is shown that under some mild condition, the expected signature determines the law of $\mathbf{S}_T(\omega)$.

1.2.8. *Lévy processes* [1, 4, 18, 26, 42]. Recall that a d-dimensional Lévy process (X_t) is a stochastically continuous process such that (i) for all $0 < s < t < \infty$, the law of $X_t - X_s$ depends only on $t - s$; (ii) for all $t_1, \dots, t_k$ such that $0 < t_1 < \cdots < t_k$ the random variables $X_{t_{i+1}} - X_{t_i}$ are independent. The Lévy process can (and will) be taken with cádlág sample paths and are characterized by the Lévy triplet (a, b, K), where $a = (a^{i,j})$ is a positive semidefinite symmetric matrix, $b = (b^i)$ a vector and $K(dx)$ a Lévy measure on $\mathbb{R}^d$ [no mass at 0, integrates $\min(|x|^2, 1)$] so that

$$\begin{aligned} \mathbb{E}\big[e^{i\langle u, X_t\rangle}\big] = \exp\biggl(&-\frac{1}{2}\langle u, au\rangle + i\langle u, b\rangle \\ &+ \int_{\mathbb{R}^d} \big(e^{iuy} - 1 - iuy1_{\{|y|<1\}}\big)K(dy)\biggr). \end{aligned} \tag{1.2.12}$$

The Itô–Lévy decomposition asserts that any such Lévy process may be written as

$$X_t = \sigma B_t + bt + \int_{(0,t]\times\{|y|<1\}} y\tilde{N}(ds, dy) + \int_{(0,t]\times\{|y|\ge 1\}} yN(ds, dy), \tag{1.2.13}$$

where B is a d-dimensional Browbian motion, $\sigma\sigma^T = a$, and N (resp., $\tilde{N}$) is the Poisson random measure (resp., compensated PRM) with intensity $ds K(dy)$. A Markovian description of a Lévy process is given in terms of its generator

$$
\begin{aligned}
(\mathcal{L}f)(x) = {} & \frac{1}{2}\sum_{i,j=1}^{d} a^{i,j}\partial_i\partial_j f + \sum_{i=1}^{d} b^i \partial_i f \\
& + \int_{\mathbb{R}^d}\left(f(x+y) - f(x) - 1_{\{|y|<1\}}\sum_{i=1}^{d} y^i \partial_i f\right) K(dy).
\end{aligned}
\tag{1.2.14}
$$

By a classical result of Hunt [18], this characterisation extends to the Lévy process with values in a Lie group G, defined as above, but with $X_t - X_s$ replaced by $X_s^{-1}X_t$. Let $\{\mathfrak{u}_1, \dots, \mathfrak{u}_m\}$ be a basis of the Lie algebra $\mathfrak{g}$, thought of a left-invariant first-order differential operators. In the special case of exponential Lie groups, meaning that $\exp : \mathfrak{g} \to G$ is an analytical diffeomorphism [so that $g = \exp(x^i \mathfrak{u}_i)$ for all $g \in G$, with canonical coordinates $x^i = x^i(g)$ of the first kind] the generator reads

$$
\begin{aligned}
(\mathcal{L}f)(x) = {} & \frac{1}{2}\sum_{v,w=1}^{m} a^{v,w}\mathfrak{u}_v\mathfrak{u}_w f + \sum_{v=1}^{m} b^v \mathfrak{u}_v f \\
& + \int_{G}\left(f(xy) - f(x) - 1_{\{|y|<1\}}\sum_{v=1}^{m} y^v \mathfrak{u}_v f\right) K(dy).
\end{aligned}
\tag{1.2.15}
$$

As before, the Lévy triplet (a, b, K) consists of $(a^{v,w})$ positive semidefinite symmetric, $b = (b^v)$ and $K(dx)$ a Lévy measure on G [no mass at the unit element, integrates $\min(|x|^2, 1)$, with $|x|^2 := \sum_{i=v}^{m}(x^v)^2$].

1.2.9. *The work of D. Williams* [46]. Williams first considers the Young regime $p \in [1, 2)$ and shows that every $X \in W^p([0, T])$ may be turned into $\tilde{X} \in C^{p\text{-var}}([0, \tilde{T}])$, by replacing jumps by segments of straight lines (in the spirit of Marcus canonical equations, via some time change $[0, T] \to [0, \tilde{T}]$). Crucially, this can be done with a uniform estimate $\|\tilde{X}\|_{p\text{-var}} \lesssim \|X\|_{p\text{-var}}$. In the rough regime $p \geq 2$, Williams considers a generic d-dimensional Lévy process X enhanced with stochastic area

$$\mathbb{A}_{s,t} := \text{Anti}\int_{(s,t]} (X_r^- - X_s) \otimes dX_r,$$

where the stochastic integration is understood in Itô sense. On a technical level, his main results ([46], pages 310–320) are summarized in the following.

THEOREM 11 (Williams). *Assume X is a d-dimensional Lévy process X with triplet (a, b, K)*:

(i) *Assume K has compact support. Then*

$$\mathbf{E}\big[|\mathbb{A}_{s,t}|^2\big] \lesssim |t-s|^2.$$

(ii) *For any $p > 2$, with* sup *taken over all partitions of* $[0, T]$,

$$\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |\mathbb{A}_{s,t}|^{p/2} < \infty \qquad a.s.$$

Clearly, $(X, \mathbb{A})(\omega)$ is all the information one needs to have a (in our terminology) càdlàg geometric p-rough path $\mathbf{X} = \mathbf{X}(\omega)$, for any $p \in (2, 3)$. However, Williams does not discuss rough integration, nor does he give meaning (in the sense of an integral equation) to a rough differential equations driven by càdlàg p-rough paths. Instead he constructs, again in the spirit of Marcus, $\tilde{\mathbf{X}} \in \mathcal{C}^{p\text{-var}}([0, \tilde{T}])$, and then goes on to *define* a solution Y to an RDE driven by $\mathbf{X}(\omega)$ as reverse-time change of a (classical) RDE solution driven by the (continuous) geometric p-rough path $\tilde{\mathbf{X}}$. While this construction is of appealing simplicity, the time-change depends in a complicated way on the jumps of $X(\omega)$ and the absence of quantitative estimates makes any local analysis of so-defined RDE solution difficult (starting with the identification of Y as solution to the corresponding Marcus canonical equation). We shall not rely on any of Williams's result, although his ideas will be visible at various places in this paper. A simplified proof of Theorem 11 will be given below.

2. Rough paths in presence of jumps: Deterministic theory.

2.1. *General rough paths*: *Definition and first examples*. The following definitions are fundamental.

DEFINITION 12. Fix $p \in [2, 3)$. We say that $\mathbf{X} = (X, \mathbb{X})$ is a *general* (*càdlàg*) *rough path over* $\mathbb{R}^d$ if:

(i) Chen's relation holds, that is, for all $s \le u \le t$, $\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}$;

(ii) the following map

$$[0, T] \ni t \mapsto \mathbf{X}_{0,t} = (X_{0,t}, \mathbb{X}_{0,t}) \in \mathbb{R}^d \oplus \mathbb{R}^{d\times d}$$

is càdlàg;

(iii) p-variation regularity in rough path sense holds, that is,

$$\|X\|_{p\text{-var};[0,T]} + \|\mathbb{X}\|^{1/2}_{p/2\text{-var};[0,T]} < \infty.$$

We then write

$$\mathbf{X} \in \mathcal{W}^p = \mathcal{W}^p([0, T], \mathbb{R}^d).$$

DEFINITION 13. We call $\mathbf{X} \in \mathcal{W}^p$ *geometric if it takes values in* $G^{(2)}(\mathbb{R}^d)$, in symbols $\mathbf{X} \in \mathcal{W}^p_{\mathrm{g}}$. If, in addition,

$$(\Delta_t X, \Delta_t \mathbb{A}) := \log \Delta_t \mathbf{X} \in \mathbb{R}^d \oplus \{0\} \subset \mathfrak{g}^{(2)}(\mathbb{R}^d)$$

we call $\mathbf{X}$ *Marcus-like*, *in symbols* $\mathbf{X} \in \mathcal{W}^p_{\mathrm{M}}$.

As in the case of (continuous) rough paths (cf. Section 1.2.5),

$$\mathcal{W}^p_{\mathrm{g}} := \mathcal{W}^p_{\mathrm{g}}([0,T], \mathbb{R}^d) = W^p([0,T], G^{(2)}(\mathbb{R}^d))$$

so that general geometric p-rough paths are precisely paths of finite p-variation in $G^{(2)}(\mathbb{R}^d)$ equipped with CC metric. We can generalize the definition to general $p \in [1, \infty)$ at the price of working in the step-$[p-]$ free nilpotent group,

$$\mathcal{W}^p_{\mathrm{g}} = W^p([0,T], G^{([p])}).$$

As a special case of Lyons' extension theorem (Theorem 7), for a given continuous path $X \in W^p$ for $p \in [1,2)$, there is a *unique* rough path $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{W}^p$. [Uniqueness is lost when $p \geq 2$, as seen by the perturbation $\bar{\mathbb{X}}_{s,t} = \mathbb{X}_{s,t} + a(t-s)$, for some matrix a.]

The situation is different in presence of jumps and Lyons' first theorem fails, even when $p = 1$. Essentially, this is due to the fact that there are nontrivial pure jump paths of finite q-variation with $q < 1$.

PROPOSITION 14 (Canonical lifts of paths in Young regime). *Let* $X \in W^p([0,T], \mathbb{R}^d)$ *be a càdlàg path of finite* p*-variation for* $p \in [1,2)$.

(i) *It is lifted to a* (*in general*, *nongeometric*[4]) *rough path* $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{W}^p$ *by enhancing* X *with*

$$\mathbb{X}_{s,t} = (\mathit{Young}) \int_{(s,t]} X_{s,r-} \otimes dX_r.$$

(ii) *It is lifted to a Marcus-like càdlàg rough path* $\mathbf{X}^{\mathrm{M}} = (X, \mathbb{X}^{\mathrm{M}}) \in \mathcal{W}^p_{\mathrm{M}}$ *by enhancing* X *with*

$$\mathbb{X}^{\mathrm{M}}_{s,t} = \mathbb{X}_{s,t} + \frac{1}{2} \sum_{r \in (s,t]} (\Delta_r X) \otimes (\Delta_r X).$$

PROOF. (i) Càdlàg regularity is clear, as is Chen's relation (perhaps noting that $\int_{(u,t]} dX = X_{u,t}$ since X is càdlàg). As an application of Young's inequality, using $p < 2$, it is easy to see that

$$|\mathbb{X}_{s,t}| \lesssim \|X\|^2_{p\text{-var};[s,t]}.$$

[4]Consider the scalar càdlàg path $X_t = \mathbb{I}_{[1,2]}(t)$ on $[0,2]$. Then $\mathbb{X}_{0,2} = 0$, and hence $\mathbb{X}_{0,2} \neq \frac{1}{2} X^2_{0,2}$.

Note that $\omega(s,t) := \|X\|^p_{p\text{-var};[s,t]}$ is super-additive, that is, for all $s < u < t$, $\omega(s,u) + \omega(u,t) \le \omega(s,t)$, so that

$$\sum_{[s,t]\in\mathcal{P}} |\mathbb{X}_{s,t}|^{\frac{p}{2}} \lesssim \sum_{[s,t]\in\mathcal{P}} \|X\|^p_{p\text{-var};[s,t]} \lesssim \|X\|^p_{p\text{-var};[0,T]}.$$

Taking sup over $\mathcal{P}$, $\mathbb{X}$ has $\frac{p}{2}$ variation and this concludes (i). We then note that

$$\left|\sum_{r\in(s,t]} (\Delta_r X) \otimes (\Delta_r X)\right|^{\frac{p}{2}} \le \left(\sum_{r\in(s,t]} |\Delta_r X|^2\right)^{\frac{p}{2}} \le \sum_{r\in(s,t]} |\Delta_r X|^p;$$

where we used $p \le 2$ in the last inequality. Since the jumps of X are p-summable, we immediately conclude that $\mathbb{X}^{\mathrm{M}}$ also is of finite $\frac{p}{2}$ variation. Also, from "integration by parts formula for sums", it can be easily checked that $\mathrm{Sym}(\mathbb{X}^{\mathrm{M}}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}$. The fact that $(X, \mathbb{X}^{\mathrm{M}})$ forms a Marcus-like rough path comes from the underlying idea of the Marcus integral replaces jumps by straight lines which do not create area. Precisely,

$$\lim_{s\uparrow t} \mathbb{X}^{\mathrm{M}}_{s,t} =: \Delta_t \mathbb{X}^{\mathrm{M}} = \frac{1}{2}(\Delta_t X)^{\otimes 2}$$

which is symmetric. Thus, $\Delta_t \mathbb{A} = \mathrm{Anti}(\Delta_t \mathbb{X}^{\mathrm{M}}) = 0$. This completes the proof for part (ii). □

Clearly, in the continuous case every geometric rough path is Marcus-like and so there is need to distinguish them. The situation is different with jumps and there are large classes of Marcus-like as well as non-Marcus-like geometric rough paths. We give some examples.

EXAMPLE 15. (a) Pure area jump rough path. Consider a $\mathfrak{so}(d)$-valued path (A_t) of finite 1-variation, started at $A_0 = 0$. Then

$$\mathbf{X}_{0,t} := \exp(A_t)$$

defines a geometric rough path, for any $p \ge 2$, that is, $\mathbf{X}(\omega) \in \mathcal{W}^p_{\mathrm{g}}$ but, unless A is continuous,

$$\mathbf{X}(\omega) \notin \mathcal{W}^p_{\mathrm{M}}.$$

(b) Pure area Poisson process. Consider an i.i.d. sequence of a $\mathfrak{so}(d)$-valued r.v. $(\mathfrak{a}^n(\omega))$ and a standard Poisson process N_t with rate $\lambda > 0$. Then, with probability one,

$$\mathbf{X}_{0,t}(\omega) := \exp\left(\sum_{n=1}^{N_t} \mathfrak{a}^n(\omega)\right)$$

yields a geometric, non-Marcus-like càdlàg rough path for any $p \ge 2$.

(c) Pure area rough path. Fix $\mathfrak{a} \in \mathfrak{so}(d)$. Then

$$\mathbf{X}_{0,t} := \exp(\mathfrak{a}t),$$

yields a geometric rough path, $\mathbf{X} \in \mathcal{C}_g^p([0,T],\mathbb{R}^d)$, above the trivial path $X \equiv 0$, for any $p \in [2,3)$.

(d) Brownian rough path in magnetic field. Write

$$\mathbf{B}^{\mathrm{S}}_{s,t} = \left(B_{s,t}, \int_s^t B_{s,r} \otimes \circ dB_r\right)$$

for the canonical Brownian rough path based on iterated Stratonovich integration. If one considers the (zero-mass) limit of a physical Brownian particle, with nonzero charge, in a constant magnetic field [11] one finds the (noncanonical) Brownian rough path

$$\mathbf{B}^{\mathrm{m}}_{0,t} := \mathbf{B}^{\mathrm{S}}{}_{0,t} + (0, \mathfrak{a}t),$$

for some $\mathfrak{a} \in \mathfrak{so}(d)$. This yields a continuous, noncanonical geometric rough path lift of Brownian motion. More precisely, $\mathbf{B}^{\mathrm{m}} \in \mathcal{C}_g^p([0,T],\mathbb{R}^d)$ a.s., for any $p \in (2,3)$.

As is well known in rough path theory, it is not trivial to construct suitable $\mathbb{X}$ given some (irregular) path X, and most interesting constructions are of stochastic nature. At the same time, X does not determine $\mathbb{X}$, as was seen in the above examples. That said, once in possession of a (càdlàg) rough path, there are immediate ways to obtain further rough paths, of which we mention in particular perturbation of $\mathbb{X}$ by increments of some $p/2$-variation path, and, second, subordination of $(X, \mathbb{X})$ by some increasing (càdlàg) path. For instance, in a stochastic setting, any time change of the (canonical) Brownian rough path, by some Lévy subordinator for instance, will yield a general random rough path, corresponding to the (càdlàg) rough path associated to a specific semimartingale.

For Brownian motion, as for (general) semimartingales, there are two "canonical" candidates for $\mathbb{X}$, obtained by Itô- and Marcus-canonical (= Stratonovich in absence of jumps) integration, respectively. We have the following.

PROPOSITION 16. *Consider a d-dimensional* (*càdlàg*) *semimartingale X and let $p \in (2,3)$. Then the following three statements are equivalent*:

(i) $\mathbf{X}^{\mathrm{I}}(\omega) \in \mathcal{W}^p$ *a.s. where* $\mathbf{X}^{\mathrm{I}} = (X, \mathbb{X}^{\mathrm{I}})$ *and*

$$\mathbb{X}^{\mathrm{I}}_{s,t} := \int_s^t X_{s,r-} \otimes dX_r \qquad (\underline{I}t\hat{o}).$$

(ii) $\mathbf{X}^{\mathrm{M}}(\omega) \in \mathcal{W}^p_{\mathrm{M}}$ ($\subset \mathcal{W}^p_{\mathrm{g}}$) *a.s. where* $\mathbf{X}^{\mathrm{M}} = (X, \mathbb{X}^{\mathrm{M}})$ *and*

$$\mathbb{X}^{\mathrm{M}}_{s,t} := \int_s^t X_{s,r-} \diamond \otimes dX_r \qquad (\underline{M}arcus).$$

(iii) *The stochastic area* (*identical for both Itô- and Marcus lift*)

$$\mathbb{A}_{s,t} := \mathrm{Anti}(\mathbb{X}^{\mathrm{I}}_{s,t}) = \mathrm{Anti}(\mathbb{X}^{\mathrm{M}}_{s,t})$$

has a.s. finite $p/2$-variation.

PROOF. Clearly,

$$\mathrm{Sym}(\mathbb{X}^{\mathrm{M}}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}$$

is of finite $p/2$-variation, a consequence of $X \in W^p$ a.s., for any $p > 2$. Note that $\mathbb{X}^{\mathrm{M}} - \mathbb{X}^{\mathrm{I}}$ is symmetric,

$$(\mathbb{X}^{\mathrm{M}}_{s,t})^{i,j} - (\mathbb{X}^{\mathrm{I}}_{s,t})^{i,j} = \frac{1}{2}[X^i, X^j]^c_{s,t} + \frac{1}{2} \sum_{r \in (s,t]} \Delta_r X^i \Delta_r X^j,$$

and is of finite $\frac{p}{2}$ variation as $[X^i, X^j]^c$ is of bounded variation, while

$$\Bigg| \sum_{r \in (s,t]} \Delta_r X^i \Delta_r X^j \Bigg|^{\frac{p}{2}} \le \Bigg| \frac{1}{2} \sum_{r \in (s,t]} |\Delta_r X|^2 \Bigg|^{\frac{p}{2}} \lesssim \sum_{r \in (s,t]} |\Delta_r X|^p < \infty \qquad \text{a.s.}$$

because jumps of semimartingale is square summable, and thus $p \ge 2$ summable. □

We now give an elegant criterion for checking finite 2^+-variation of $G^{(2)}$-valued processes.

PROPOSITION 17. *Consider a $G^{(2)}(\mathbb{R}^d)$-valued strong Markov process $\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t = \exp(X_{s,t}, \mathbb{A}_{s,t})$. Assume*

$$\mathbb{E}|X_{s,t}|^2 \lesssim |t-s|,$$

$$\mathbb{E}|\mathbb{A}_{s,t}|^2 \lesssim |t-s|^2,$$

uniformly in $s, t \in [0, T]$. Then, for any $p > 2$,

$$\|X\|_{p\text{-}var} + \|\mathbb{A}\|_{p/2\text{-}var} < \infty \qquad a.s.$$

Equivalently, $\|\mathbf{X}\|_{p\text{-var}} < \infty$ a.s.

PROOF. Consider $s, t \in [0, T]$ with $|t-s| \le h$. Then

$$\mathbb{P}(|X_{s,t}| \ge a) \le \frac{1}{a^2} \mathbb{E}|X_{s,t}|^2 \lesssim \frac{h}{a^2},$$

$$\mathbb{P}(|\mathbb{A}_{s,t}|^{1/2} \ge a) = \mathbb{P}(|\mathbb{A}_{s,t}| \ge a^2) \le \frac{1}{a^2} \mathbb{E}|\mathbb{A}_{s,t}|$$

$$\le \frac{1}{a^2} (E|\mathbb{A}_{s,t}|^2)^{1/2} = \frac{h}{a^2}.$$

From properties of the Carnot–Caratheodory metric, $d_{\mathrm{CC}}(\mathbf{X}_s, \mathbf{X}_t) \asymp |X_{s,t}| + |\mathbb{A}_{s,t}|^{1/2}$ and the above estimates yield

$$\mathbb{P}\big(d(\mathbf{X}_s, \mathbf{X}_t) \geq a\big) \lesssim \frac{h}{a^2}.$$

Applying the result of Manstavicius (cf. Section 1.2.6) with $\beta = 1$, $\gamma = 2$ we obtain a.s. finite p-variation of $\mathbf{X}$, any $p > \gamma/\beta = 2$, with respect to d_{CC} and the statement follows. $\square$

As will be detailed in Section 3.1.1, this criterion, combined with the expected signature of a d-dimensional Lévy process, provides an immediate way to recover Williams's rough path regularity result on Lévy process (Theorem 11) and then significantly larger classes of jump diffusions. With the confidence that there exists large classes of random càdlàg rough paths, we continue to develop the deterministic theory.

2.2. *The minimal jump extension of càdlàg rough paths*. In view of Theorem 7, it is natural to ask for such extension theorem for càdlàg rough paths. (For continuous paths in Young regime, extension is uniquely given by n-fold iterated Young integrals.) However, in presence of jumps the uniqueness part of Lyons' extension theorem fails, as already seen by elementary examples of finite variation paths.

EXAMPLE 18. Let $p = 1$, $N = 2$ and consider the trivial path $X \equiv 0 \in W^1([0, 1], \mathbb{R}^d)$, identified with $\mathbf{X} \equiv (1, 0) \in W^1([0, 1], G^{(1)})$. Consider a nontrivial $\mathfrak{so}(d)$-valued càdlàg path $a(t)$, of pure finite jump type, that is,

$$a_{0,t} = \sum_{\substack{s \in (0,t] \\ \text{(finite)}}} \Delta a_s.$$

Then two possible lifts of $\mathbf{X}$ are given by

$$\mathbf{X}^{(2)} \equiv (1, 0, 0), \qquad \tilde{\mathbf{X}}^{(2)}_t \equiv (1, 0, a_t - a_0) \in \mathcal{W}^{1\text{-var}}_{\mathrm{g}} = W^1\big([0, 1], G^{(2)}\big).$$

We can generalize this example as follows.

EXAMPLE 19. Again $p = 1$, $N = 2$ and consider $X \in W^{1\text{-var}}$. Then

$$\mathbf{X}^{(2)}_t := \big(1, X_t, \mathbb{X}^{\mathrm{M}}_t\big) \in \mathcal{W}^{1\text{-var}}_{\mathrm{g}}$$

and another choice is given by

$$\tilde{\mathbf{X}}^{(2)}_t \equiv \big(1, X_t, \mathbb{X}^{M}_t + a_t - a_0\big) \in \mathcal{W}^{1\text{-var}}_{\mathrm{g}},$$

whenever, $a_t \in \mathfrak{so}(d)$ is piecewise constant, with finitely many jumps $\Delta a_t \neq 0$.

Note that, among all such lifts $\tilde{\mathbf{X}}_t^{(2)}$, the $\mathbf{X}_t^{(2)}$ is minimal in the sense that $\log^{(2)} \Delta\mathbf{X}_t^{(2)}$ has no 2-tensor component, and in fact,

$$\log^{(2)} \Delta\mathbf{X}_t^{(2)} = \Delta X_t.$$

We have the following far-reaching extension of this example. Note that we consider $\mathfrak{g}^n \supset \mathfrak{g}^m$ in the obvious way whenever $n \geq m$.

THEOREM 20 (Minimal jump extension). *Let $1 \leq p < \infty$ and $\mathbb{N} \ni n > m := [p]$. A càdlàg rough path $\mathbf{X}^{(m)} \in \mathcal{W}_g^p = W^p([0,T], G^{(m)})$ admits an extension to a path $\mathbf{X}^{(n)}$ of with values $G^{(n)} \subset T^{(n)}$, unique in the class of $G^{(n)}$-valued path starting from 1 and of finite p-variation with respect to CC metric on $G^{(n)}$ subject to the additional constraint*

$$\log^{(n)} \Delta\mathbf{X}_t^{(n)} = \log^{(m)} \Delta\mathbf{X}_t^{(m)}. \tag{2.2.1}$$

For the proof, we will adopt the Marcus/Williams idea of introducing an artificial additional time interval at each jump times of $\mathbf{X}^m$, during which the jump will be suitably traversed. Since $\mathbf{X}^m$ may have countably infinite many jumps, we number the jumps as follows. Let t_1 is such that

$$\|\Delta_{t_1}\mathbf{X}^{(m)}\|_{\mathrm{CC}} = \sup_{t\in[0,T]}\{\|\Delta_t\mathbf{X}^{(m)}\|_{\mathrm{CC}}\}.$$

Similarly, define t_2 with

$$\|\Delta_{t_2}\mathbf{X}^{(m)}\|_{\mathrm{CC}} = \sup_{t\in[0,T],t\neq t_1}\{\|\Delta_t\mathbf{X}^{(m)}\|_{\mathrm{CC}}\}$$

and so on. Note that the suprema are always attained and if $\|\Delta_t\mathbf{X}^{(m)}\|_{\mathrm{CC}} \neq 0$, then $t = t_k$ for some k. Indeed, it readily follows from the càdlàg (or regulated) property that for any $\epsilon > 0$, there are only finitely many jumps with $\|\Delta_t\mathbf{X}^{(m)}\|_{\mathrm{CC}} > \epsilon$.

Choose any sequence $\delta_k > 0$ such that $\sum_k \delta_k < \infty$. Starting from t_1, we recursively introduce an interval of length δ_k at t_k, during which the jump $\Delta_{t_k}\mathbf{X}^{(m)}$ is traversed suitably, to get a continuous curve $\tilde{\mathbf{X}}^{(m)}$ on the (finite) interval $[0, \tilde{T}]$ where

$$\tilde{T} = T + \sum_{k=1}^{\infty} \delta_k < \infty.$$

Taking motivation from simple examples, in order to get minimal jump extensions, we choose the "best possible" curve traversing the jump, so that it does not create additional terms in $\log^{(n)} \Delta\mathbf{X}_t^{(n)}$. If $[a, a+\delta_k] \subset [0, \tilde{T}]$ is the jump segment corresponding to the kth jump, define

$$\gamma_t^k = \mathbf{X}_{t_k-}^{(m)} \otimes \exp^{(m)}\left(\frac{t-a}{\delta_k}\log^{(m)} \Delta_{t_k}\mathbf{X}^{(m)}\right).$$

LEMMA 21. *$\gamma^k : [a, a+\delta_k] \to G^{(m)}$ is a continuous path of finite p variation w.r.t. the CC metric and we have the bound*

$$\|\gamma^k\|^p_{p\text{-var};[a,a+\delta_k]} \lesssim \|\Delta_{t_k}\mathbf{X}^{(m)}\|^p. \tag{2.2.2}$$

PROOF. Omit k. Without loss of generality we can assume that $\gamma_t = x \otimes \exp^{(m)}(t \log^{(m)}(x^{-1} \otimes y))$ for $t \in [0, 1]$ and $x, y \in G^{(m)}$. Thus by Campbell–Baker–Hausdorff formula

$$\gamma_{s,t} = \exp^{(m)}((t-s)\log^{(m)}(x^{-1} \otimes y))$$

Since $p \geq m$, it is easy to check that for $z \in \mathfrak{g}^m$ and $\lambda \in [0, 1]$,

$$\|\exp^{(m)}(\lambda z)\|^p \lesssim \lambda \|\exp^{(m)}(z)\|^p$$

So,

$$\|\gamma_{s,t}\|^p \lesssim (t-s)\|x^{-1} \otimes y\|^p$$

which proves the claim. □

LEMMA 22. *The curve $\tilde{\mathbf{X}}^{(m)} : [0, \tilde{T}] \to G^{(m)}$ constructed as above from $\mathbf{X}^{(m)} \in W^p([0,T], G^{(m)})$ is a continuous path of finite p variation w.r.t. the CC metric and we have the bound*

$$\|\tilde{\mathbf{X}}^{(m)}\|_{p\text{-var};[0,\tilde{T}]} \lesssim \|\mathbf{X}^{(m)}\|_{p\text{-var};[0,T]}. \tag{2.2.3}$$

PROOF. For simpler notation, omit m and write $\tilde{\mathbf{X}}, \mathbf{X}$. The curve $\tilde{\mathbf{X}}$ is continuous by construction. To see the estimate, introduce $\omega(s,t) = \|\mathbf{X}\|^p_{p\text{-var},[s,t]}$ with the notation $\omega(s, t-) := \|\mathbf{X}\|^p_{p\text{-var},[s,t)}$. Note that $\omega(s,t)$ is super-additive. Call t' the pre-image of $t \in [0, \tilde{T}]$ under the time change from $[0,T] \to [0,\tilde{T}]$. Note that $[0, \tilde{T}]$ contains (possibly countably many) jump segments I_n of the form $[a, a+\delta_k)$. Let us agree that point in these jump segments are "red" and all remaining points are "blue." Note that jump segments correspond to one point in the pre-image. For $0 \leq s < t \leq \tilde{T}$, there are following possibilities:

- Both s, t are blue, in which case $\|\tilde{\mathbf{X}}_{s,t}\|^p = \|\mathbf{X}_{s',t'}\|^p \leq \|\mathbf{X}\|^p_{p\text{-var},[s',t']} = \omega(s',t')$.
- Both s, t are red and in same jump segment $[a, a+\delta_k)$, in which case
$$\|\tilde{\mathbf{X}}_{s,t}\|^p \leq \|\gamma\|^p_{p\text{-var};[a,a+\delta_k]}.$$
- Both s, t are red but in different jump segment $s \in [a, a+\delta_k)$ and $t \in [b, b+\delta_l)$, in which case $s' = (a+\delta_k)'$, $t' = (b+\delta_l)'$ and
$$\begin{aligned}\|\tilde{\mathbf{X}}_{s,t}\|^p &\leq 3^{p-1}(\|\tilde{\mathbf{X}}_{s,a+\delta_k}\|^p + \|\tilde{\mathbf{X}}_{a+\delta_k,b}\|^p + \|\tilde{\mathbf{X}}_{b,t}\|^p)\\ &\leq 3^{p-1}(\|\gamma^k\|^p_{p\text{-var};[a,a+\delta_k]} + \|\mathbf{X}_{s',t'-}\|^p + \|\gamma^l\|^p_{p\text{-var},[b,b+\delta_l]})\\ &\leq 3^{p-1}(\|\gamma^k\|^p_{p\text{-var};[a,a+\delta_k]} + \omega(s',t'-) + \|\gamma^l\|^p_{p\text{-var},[b,b+\delta_l]}).\end{aligned}$$

- s is blue and $t \in [a, a+\delta_k)$ is red, in which case

$$\begin{aligned}\|\tilde{\mathbf{X}}_{s,t}\|^p &\le 2^{p-1}\big(\|\tilde{\mathbf{X}}_{s,a}\|^p + \|\tilde{\mathbf{X}}_{a,t}\|^p\big)\\ &\le 2^{p-1}\big(\omega(s', t'-) + \|\gamma^k\|^p_{p\text{-var};[a,a+\delta_k]}\big).\end{aligned}$$

- $s \in [a, a+\delta_k)$ is red and t is blue, then

$$\begin{aligned}\|\tilde{\mathbf{X}}_{s,t}\|^p &\le 2^{p-1}\big(\|\tilde{\mathbf{X}}_{s,a+\delta_k}\|^p + \|\tilde{\mathbf{X}}_{a+\delta_k,t}\|^p\big)\\ &\le 2^{p-1}\big(\|\gamma^k\|^p_{p\text{-var};[a,a+\delta_k]} + \omega(s', t')\big).\end{aligned}$$

In any case, by using Lemma 21, we see that

$$\|\tilde{\mathbf{X}}_{s,t}\|^p \lesssim \omega(s', t') + \|\Delta_{s'}\mathbf{X}\|^p + \|\Delta_{t'}\mathbf{X}\|^p$$

which implies for any partition $\mathcal{P}$ of $[0, \tilde{T}]$,

$$\begin{aligned}\sum_{[s,t]\in\mathcal{P}} \|\tilde{\mathbf{X}}_{s,t}\|^p &\lesssim \sum_{[s',t']} \omega(s', t') + \|\Delta_{s'}\mathbf{X}\|^p + \|\Delta_{t'}\mathbf{X}\|^p\\ &\lesssim \omega(0, T) + \sum_{0<s\le T} \|\Delta_s\mathbf{X}\|^p.\end{aligned}$$

Finally, note that

$$\sum_{0<s\le T} \|\Delta_s\mathbf{X}\|^p \le \|\mathbf{X}\|^p_{p\text{-var};[0,T]}$$

which proves the claim. □

PROOF OF THEOREM 20. Since $\tilde{\mathbf{X}}^{(m)}$ is continuous path of finite p-variation on $[0, \tilde{T}]$, from Theorem 7, it admits an extension $\tilde{\mathbf{X}}^{(n)}$ taking values in $G^{(n)}$ starting from 1 for all $n > m$. We emphasize that $S = \tilde{\mathbf{X}}^{(n)}$ can be obtained as linear RDE solution to

$$dS = S \otimes d\mathbf{X}^{(m)},\ S_0 = 1 \in T^{(n)}. \tag{2.2.4}$$

We claim that for each jump segment $[a, a+\delta_k]$,

$$\tilde{\mathbf{X}}^{(n)}_{a,a+\delta_k} = \exp^{(n)}\big(\log^{(m)}\big(\Delta_{t_k}\mathbf{X}^{(m)}\big)\big)$$

which amounts to proving that if $\gamma_t = x \otimes \exp^{(m)}(t \log^{(m)}(x^{-1} \otimes y))$ for $t \in [0, 1]$ and $x, y \in G^{(m)}$, then its extension $\gamma^{(n)}$ to $G^{(n)}$ satisfies

$$\gamma^{(n)}_{0,1} = \exp^{(n)}\big(\log^{(m)}\big(x^{-1} \otimes y\big)\big)$$

Again by the Campbell–Baker–Hausdorff formula,

$$\gamma_{s,t} = \exp^m\big((t-s)\log^m\big(x^{-1} \otimes y\big)\big)$$

Since $\gamma^{(n)}_{s,t} := \exp^{(n)}((t-s)\log^{(m)}(x^{-1}\otimes y))$ is clearly an extension of $\gamma_{s,t}$ from $G^{(m)}$ to $G^{(n)}$, by uniqueness of Theorem 7,

$$\gamma^{(n)}_{0,1} = \exp^{(n)}\big(\log^{(m)}(x^{-1}\otimes y)\big)$$

which finishes the existence part of Theorem 20.
For uniqueness, without loss of generality, assume $n = m+1$. Let $\mathbf{Z}^{(n)}_t = \mathbf{X}^{(m)}_t + M_t$ and $\mathbf{Y}^{(n)}_t = \mathbf{X}^{(m)}_t + N_t$ are two extensions of $\mathbf{X}^{(m)}_t$ as prescribed of Theorem 20, where $M_t, N_t \in (\mathbb{R}^d)^{\otimes n}$.

Consider

$$S_t = \mathbf{Z}^{(n)}_t \otimes \{\mathbf{Y}^{(n)}_t\}^{-1} = \big(\mathbf{X}^{(m)}_t + M_t\big)\otimes\big(\mathbf{X}^{(m)}_t + N_t\big)^{-1} = 1 + M_t - N_t,$$

where the last equality is due to truncation in the (truncated) tensor product. This in particular implies S_t is in centre of the group $G^{(n)}$ (actually group T^n_1), and thus so is $S_s^{-1}\otimes S_t$. So, by using symmetry and sub-additivity of CC norm,

$$\|S_s^{-1}\otimes S_t\| = \|\mathbf{Y}^{(n)}_s \otimes \mathbf{Z}^{(n)}_{s,t} \otimes \{\mathbf{Y}^{(n)}_t\}^{-1}\| = \|\mathbf{Z}^{(n)}_{s,t}\otimes\{\mathbf{Y}^{(n)}_{s,t}\}^{-1}\| \le \|\mathbf{Z}^{(n)}_{s,t}\| + \|\mathbf{Y}^{(n)}_{s,t}\|$$

which implies S_t is of finite p-variation. Also,

$$\Delta_t S = \mathbf{Y}^{(n)}_{t-}\otimes \Delta_t \mathbf{Z}^{(n)} \otimes \mathbf{Y}^{(n)}_t.$$

Since $\Delta_t\mathbf{Z}^{(n)} = \Delta_t\mathbf{Y}^{(n)}$, we see that $\log^{(n)}\Delta_t S = 0$, that is, S_t is continuous. Thus, $M - N$ is a continuous path in $(\mathbb{R}^d)^{\otimes n}$ with finite $\frac{p}{n} < 1$ variation, which implies $M_t = N_t$ completing the proof. □

REMARK 23. In the proof of uniqueness of minimal jump extension, we did not use the structure of group $G^{(n)}$. The fact that the minimal jump extension takes value in $G^{(n)}$ follows by construction. That said, if $\mathbf{Z}^{(n)}$ and $\mathbf{Y}^{(n)}$ are two extensions of $\mathbf{X}^{(m)}$ taking values in $T^{(n)}(\mathbb{R}^d)$, of finite p-variation w.r.t. norm

$$\|1+g\| := |g^1| + |g^2|^{\frac{1}{2}} + \cdots + |g^n|^{\frac{1}{n}}$$

and

$$\Delta_t\mathbf{Z}^{(n)} = \Delta_t\mathbf{Y}^{(n)} = \exp^{(n)}\big(\log^{(m)}\big(\Delta_t\mathbf{X}^{(m)}\big)\big)$$

the same argument as above implies

$$\mathbf{Z}^{(n)}_t = \mathbf{Y}^{(n)}_t.$$

We are now able to define the *signature of a càdlàg rough path*, generalising Definition 9. First, in view of Theorem 20, any $\mathbf{X}\in W^p([0,T],G^{(m)})$ may be regarded—via its minimal jump extension—as $\mathbf{X}\in W^p([0,T],G^{(N)})$, any $N\ge m$, and (as earlier, cf. remark before Definition 9) there is no ambiguity in this notation.

DEFINITION 24. Consider a càdlàg rough path $\mathbf{X} \in \mathcal{W}_g^p = W^p([0,T], G^{(m)})$. Then, thanks to Theorem 20,

$$S(\mathbf{X})_{0,T} := \big(1, \pi_1(\mathbf{X}_{0,T}), \dots, \pi_m(\mathbf{X}_{0,T}), \pi_{m+1}(\mathbf{X}_{0,T}), \dots\big) \in T\big((\mathbb{R}^d)\big)$$

defines a group-like element, called the signature of $\mathbf{X}$.

2.3. *Rough integration with jumps.* In this section, we will define rough integration for càdlàg rough paths in the spirit of [16, 47] and apply this for pathwise understanding of a stochastic integral. We restrict ourselves to case $p < 3$. For $p \in [1,2)$, Young integration theory is well established and the interesting case is for $p \in [2,3)$. Recall the meaning of convergence in (RRS) sense; cf. Definition 1. We shall distinguish the following Riemann sum approximations:

$$S(\mathcal{P}) := \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t}$$

and

$$S'(\mathcal{P}) := \sum_{[s,t]\in\mathcal{P}} Y_{s-} X_{s,t}.$$

In fact, if X and Y are regulated paths of finite p-variation for $p < 2$, then

$$C := (\mathrm{RRS}) \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t}$$

exist if either Y is càdlàg or Y is cáglád (left continuous with right limit) and X is càdlàg. (This can be easily verified by carefully reviewing the proof of existence of the Young integral as in [8], noting that we have restricted ourselves to left point evaluation in Riemann sums.) In particular, if X, Y are càdlàg paths then

$$C_1 := (\mathrm{RRS}) \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t}$$

and

$$C_2 := (\mathrm{RRS}) \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} Y_{s-} X_{s,t}$$

both exist. We now see that in the Young case both limits are equal.

PROPOSITION 25. *If X and Y are càdlàg paths of finite p-variation for $p < 2$, then $C_1 = C_2$.*

PROOF. For each $\epsilon > 0$,

$$\begin{aligned} S(\mathcal{P}) - S'(\mathcal{P}) &= \sum_{[s,t]\in\mathcal{P}} \Delta Y_s X_{s,t} \\ &= \sum_{[s,t]\in\mathcal{P}} \Delta Y_s 1_{|\Delta Y_s|>\epsilon} X_{s,t} + \sum_{[s,t]\in\mathcal{P}} \Delta Y_s 1_{|\Delta Y_s|\le\epsilon} X_{s,t} \end{aligned}$$

and since there are finitely many jumps of size bigger than ϵ and X is right continuous,

$$\lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} \Delta Y_s 1_{|\Delta Y_s|>\epsilon} X_{s,t} = 0.$$

On the other hand,

$$\begin{aligned}\left|\sum_{[s,t]\in\mathcal{P}} \Delta Y_s 1_{|\Delta Y_s|\le\epsilon} X_{s,t}\right|^2 &\le \sum_{[s,t]\in\mathcal{P}} \bigl(|\Delta Y_s|^2 1_{|\Delta Y_s|\le\epsilon}\bigr) \sum_{[s,t]\in\mathcal{P}} |X_{s,t}|^2 \\ &\le \varepsilon^{2-p} \sum_{[s,t]\in\mathcal{P}} |\Delta Y_s|^p \sum_{[s,t]\in\mathcal{P}} |X_{s,t}|^2 \\ &\le \varepsilon^{2-p} \|Y\|_{p\text{-var}}^p \|X\|_{2\text{-var}}^2,\end{aligned}$$

where we used $p < 2$ in the step. It thus follows that

$$\lim_{\epsilon\to 0} \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} \Delta Y_s 1_{|\Delta Y_s|\le\epsilon} X_{s,t} = 0$$

which proves the claim. □

One fundamental difference between continuous and càdlàg is absence of uniform continuity which implies small oscillation of a path in a small time interval. This becomes crucial in the construction of the integral, as also can be seen in construction of the Young integral (see [8]) when the integrator and integrand are assumed to have no common discontinuity on the same side of a point. This guarantees at least one of them to have small oscillation on small time intervals.

DEFINITION 26. A pair of functions $(X_{s,t}, Y_{s,t})$ defined for $\{0 \le s \le t \le T\}$ is called compatible if for all $\epsilon > 0$, there exists a partition $\tau = \{0 = t_0 < t_1 < \cdots < t_n = T\}$ such that for all $0 \le i \le n-1$,

$$\operatorname{Osc}\bigl(X, [t_i, t_{i+1}]\bigr) \wedge \operatorname{Osc}\bigl(Y, [t_i, t_{i+1}]\bigr) \le \epsilon,$$

where $\operatorname{Osc}(Z, [s,t]) := \sup\{|Z_{u,v}| | s \le u \le v \le t\}$.

PROPOSITION 27. *If X is a càdlàg path and Y is a càglàd path, then (X, Y) is a compatible pair.*

PROOF. See [8]. □

Let $\mathbf{X} = (X, \mathbb{X})$ be càdlàg rough path in the sense of Definition 12. For the purpose of rough integration, we will use a different enhancement

$$\tilde{\mathbb{X}}_{s,t} = \mathbb{X}_{s,t} + \Delta_s X \otimes X_{s,t}.$$

Note clearly that $\tilde{\mathbb{X}}$ is also of finite $\frac{p}{2}$ variation, $\tilde{\mathbb{X}}_{0,t}$ is a càdlàg path and for $s \le u \le t$,

$$\tilde{\mathbb{X}}_{s,t} - \tilde{\mathbb{X}}_{s,u} - \tilde{\mathbb{X}}_{u,t} = X^-_{s,u} \otimes X_{u,t}, \tag{2.3.1}$$

where $X^-_t := X_{t-}$, $X^-_0 = X_0 = 0$.

LEMMA 28. *For any $\epsilon > 0$, there exists a partition $\tau = \{0 = t_0 < t_1 \cdots < t_n = T\}$ such that for all $0 \le i \le n-1$,*

$$\operatorname{Osc}(\tilde{\mathbb{X}}, (t_i, t_{i+1})) \le \epsilon.$$

PROOF. Since $\tilde{\mathbb{X}}_{0,t}$ is càdlàg, from (2.3.1), it follows that for each $y \in (0, T)$, there exists a $\delta_y > 0$ such that

$$\operatorname{Osc}(\tilde{\mathbb{X}}, (y - \delta_y, y)) \le \epsilon \quad \text{and} \quad \operatorname{Osc}(\tilde{\mathbb{X}}, (y, y + \delta_y)) \le \epsilon.$$

Similarly, there exist δ_0 and δ_T such that $\operatorname{Osc}(\tilde{\mathbb{X}}, (0, \delta_0)) \le \epsilon$ and $\operatorname{Osc}(\tilde{\mathbb{X}}, (T - \delta_T, T)) \le \epsilon$. Now a family of open sets

$$[0, \delta_0), \qquad (y - \delta_y, y + \delta_y), \qquad \dots, \qquad (T - \delta_T, T]$$

form a open cover of interval $[0, T]$, so it has a finite subcover $[0, \delta_0), (y_1 - \delta_{y_1}, y_1 + \delta_{y_1}), \dots, (y_n - \delta_{y_n}, y_n + \delta_{y_n}), (T - \delta_T, T]$. Without loss of generality, we can assume that each interval in the finite subcover is the first interval that intersects its previous one and the claim follows by choosing

$$\begin{aligned} &t_0 = 0, \qquad t_1 \in (y_1 - \delta_{y_1}, \delta_0), \qquad t_2 = y_1, \\ &t_3 \in (y_2 - \delta_{y_2}, y_1 + \delta_{y_1}), \qquad \dots, \qquad t_{2n+1} = T. \end{aligned}$$ □

LEMMA 29. *For any càglàd path Y, the pair $(Y, \tilde{\mathbb{X}})$ is a compatible pair.*

PROOF. Choose a partition τ such that for all $[s, t] \in \tau$,

$$\operatorname{Osc}(Z, (s, t)) \le \epsilon$$

for $Z = Y, X, X^-$ and (due to Lemma 28) $\tilde{\mathbb{X}}$. We refine the partition τ by adding a common continuity point of Y_t, X_t, X^-_t and $\tilde{\mathbb{X}}_{0,t}$ in each interval (s, t). Note that such common continuity points will exist because a regulated path can have only countably many discontinuities. With this choice of partition, we observe that on every odd numbered $[s, t] \in \tau$,

$$\operatorname{Osc}(\tilde{\mathbb{X}}, [s, t]) \le \epsilon$$

and on every even numbered $[s, t] \in \tau$,

$$\operatorname{Osc}(Y, [s, t]) \le \epsilon.$$ □

DEFINITION 30. Given $X \in W^p$, a pair of càdlàg paths (Y, Y') of finite p-variation is called a controlled rough path if $R_{s,t} = Y_{s,t} - Y'_s X_{s,t}$ has finite $\frac{p}{2}$-variation.

It is easy to see that for f twice continuously differentiable

$$\big(Y_t, Y'_t\big) := \big(f(X_t), f'(X_t)\big)$$

is a controlled rough path in the above sense. Also,

$$\tilde{R}_{s,t} := Y^-_{s,t} - Y'_{s-} X^-_{s,t}$$

is of finite $\frac{p}{2}$-variation and pair $(\tilde{R}, X)$ is a compatible pair. For brevity, the variation-type norm will now be written as $\|\cdot\|_p$ rather than $\|\cdot\|_{p\text{-var}}$.

THEOREM 31. *Fix regularity $p \in [2, 3)$ and let $\mathbf{X} = (X, \mathbb{X})$ be a càdlàg rough path and (Y, Y') a controlled rough path. Then*

$$\int_0^T Y_{r-}\, d\mathbf{X}_r := \lim_{|\mathcal{P}|\to 0} S(\mathcal{P}) = \lim_{|\mathcal{P}|\to 0} S'(\mathcal{P}),$$

where both limits exist in (RRS) sense, as introduced in Definition 1, and

$$S(\mathcal{P}) := \sum_{[s,t]\in\mathcal{P}} Y_{s-} X_{s,t} + Y'_{s-} \tilde{\mathbb{X}}_{s,t} = \sum_{[s,t]\in\mathcal{P}} Y_{s-} X_{s,t} + Y'_{s-} (\mathbb{X}_{s,t} + \Delta_s X \otimes X_{s,t}),$$

$$S'(\mathcal{P}) = \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t}.$$

Furthermore, we have the following rough path estimates: there exists a constant C depending only on p such that

$$\left|\int_s^t Y_{r-}\, d\mathbf{X}_r - Y_{s-} X_{s,t} - Y'_{s-} \tilde{\mathbb{X}}_{s,t}\right| \le C\big(\|\tilde{R}\|_{\frac{p}{2},[s,t]} \|X\|_{p,[s,t]} + \|Y'^-\|_{p,[s,t]} \|\tilde{\mathbb{X}}\|_{\frac{p}{2},[s,t]}\big), \tag{2.3.2}$$

$$\left|\int_s^t Y_{r-}\, d\mathbf{X}_r - Y_s X_{s,t} - Y'_s \mathbb{X}_{s,t}\right| \le C\big(\|R\|_{\frac{p}{2},[s,t]} \|X\|_{p,[s,t]} + \|Y'\|_{p,[s,t]} \|\mathbb{X}\|_{\frac{p}{2},[s,t]}\big). \tag{2.3.3}$$

PROOF. We first consider the approximations given by $S(\mathcal{P})$. We first note that if ω is a super-additive function defined on intervals, that is, for all $s \le u \le t$

$$\omega[s, u] + \omega[u, t] \le \omega[s, t]$$

then, for any partition $\mathcal{P}$ of $[s, t]$ into $r \ge 2$ intervals, there exist intervals $[u_-, u]$ and $[u, u_+]$ such that

$$\omega[u_-, u_+] \le \frac{2}{r-1}\omega[s, t]. \tag{2.3.4}$$

Also, we can immediately verify that for Z of finite p-variation,

$$[s,t] \mapsto \|Z\|_{p,[s,t]}^{p}$$

defines a super-additive function. Moreover, if ω_1 and ω_2 are two positive super-additive functions, then for $\alpha, \beta \geq 0$, $\alpha+\beta \geq 1$,

$$[s,t] \mapsto (\omega_1[s,t])^{\alpha}(\omega_2[s,t])^{\beta} \tag{2.3.5}$$

is also a super-additive function.

Now, it is enough to prove that for any $\epsilon > 0$, there exist a partition $\mathcal{P}_0$ such that for all refinement partitions $\mathcal{P}$ of $\mathcal{P}_0$,

$$|S(\mathcal{P}) - S(\mathcal{P}_0)| \lesssim \epsilon.$$

We shall consider, at first, an arbitrary partition τ and then later on make a choice—related to the notion of compatible pair—which identifies it as the desired partition $\mathcal{P}_0$. To this end, consider $[s,t] \in \tau$ and also a refinement $\mathcal{P}$ of τ. Call $\mathcal{P}_{s,t}$ the resulting partition of $[s,t]$, induced by the refinement points of $\mathcal{P}$. Pick $p' \in (p,3)$ and note that (take $\alpha = 2/3$, $\beta = 1/3$ above)

$$\omega[s,t] := \|\tilde{R}\|_{\frac{p'}{2},[s,t]}^{\frac{p'}{3}} \|X\|_{p',[s,t]}^{\frac{p'}{3}} + \|Y'^{-}\|_{p',[s,t]}^{\frac{p'}{3}} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[s,t]}^{\frac{p'}{3}}$$

is a super-additive and there exist $u_- < u < u_+ \in \mathcal{P}_{s,t}$ such that (2.3.4) holds. Using (2.3.1),

$$\begin{aligned}
&|S(\mathcal{P}_{s,t}) - S(\mathcal{P}_{s,t} \setminus u)| \\
&\quad = |\tilde{R}_{u_-,u} X_{u,u_+} + Y'^{-}_{u_-,u} \tilde{\mathbb{X}}_{u,u_+}| \\
&\quad \leq \|\tilde{R}\|_{\frac{p'}{2},[u_-,u_+]} \|X\|_{p',[u_-,u_+]} + \|Y'^{-}\|_{p',[u_-,u_+]} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[u_-,u_+]} \\
&\quad \leq \big(\|\tilde{R}\|_{\frac{p'}{2},[u_-,u_+]}^{\frac{p'}{3}} \|X\|_{p',[u_-,u_+]}^{\frac{p'}{3}} + \|Y'^{-}\|_{p',[u_-,u_+]}^{\frac{p'}{3}} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[u_-,u_+]}^{\frac{p'}{3}}\big)^{\frac{3}{p'}} \\
&\quad \leq \frac{C}{(r-1)^{\frac{3}{p'}}} \big(\|\tilde{R}\|_{\frac{p'}{2},[s,t]}^{\frac{p'}{3}} \|X\|_{p',[s,t]}^{\frac{p'}{3}} + \|Y'^{-}\|_{p',[s,t]}^{\frac{p'}{3}} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[s,t]}^{\frac{p'}{3}}\big)^{\frac{3}{p'}} \\
&\quad \leq \frac{C}{(r-1)^{\frac{3}{p'}}} \big(\|\tilde{R}\|_{\frac{p'}{2},[s,t]} \|X\|_{p',[s,t]} + \|Y'^{-}\|_{p',[s,t]} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[s,t]}\big),
\end{aligned}$$

where C denotes a constant that may depend on p, p'. Iterating this, since $p' < 3$, we get that

$$\begin{aligned}
&|S(\mathcal{P}_{s,t}) - Y_{s-} X_{s,t} + Y'_{s-} \tilde{\mathbb{X}}_{s,t}| \\
&\qquad \leq C\big(\|\tilde{R}\|_{\frac{p'}{2},[s,t]} \|X\|_{p',[s,t]} + \|Y'^{-}\|_{p',[s,t]} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[s,t]}\big).
\end{aligned}$$

Thus,

$$\left|S(\mathcal{P}) - S(\tau)\right| \le C \sum_{[s,t]\in\tau} \|\tilde{R}\|_{\frac{p'}{2},[s,t]} \|X\|_{p',[s,t]} + \|Y'^{-}\|_{p',[s,t]} \|\tilde{\mathbb{X}}\|_{\frac{p'}{2},[s,t]}.$$

Note that $(\tilde{R}, X)$ and $(Y'^{-}, \tilde{\mathbb{X}})$ are compatible pairs. As a consequence, we can take τ such that for any interval $[s,t]$ in τ

$$\begin{aligned}\|\tilde{R}\|_{\frac{p'}{2},[s,t]} \|X\|_{p',[s,t]} &\le \operatorname{Osc}(\tilde{R},[s,t]) \operatorname{Osc}(X,[s,t])^{1-\frac{p}{p'}} \|\tilde{R}\|^{\frac{p}{p'}}_{\frac{p'}{2},[s,t]} \|X\|^{\frac{p}{p'}}_{p',[s,t]} \\ &\le \epsilon \times \|\tilde{R}\|^{\frac{p}{p'}}_{\frac{p'}{2},[s,t]} \|X\|^{\frac{p}{p'}}_{p',[s,t]}.\end{aligned}$$

It follows that

$$\left|S(\mathcal{P}) - S(\tau)\right| \le C\epsilon \sum_{[s,t]\in\tau} \|\tilde{R}\|^{\frac{p}{p'}}_{\frac{p}{2},[s,t]} \|X\|^{\frac{p}{p'}}_{p,[s,t]} + \|Y'^{-}\|^{\frac{p}{p'}}_{p,[s,t]} \|\tilde{\mathbb{X}}\|^{\frac{p}{p'}}_{\frac{p}{2},[s,t]}$$

and since the term under the summation sign is super-additive—take $\alpha = 1/p'$, $\beta = 2/p'$ in (2.3.5) to see this—we get

$$\left|S(\mathcal{P}) - S(\tau)\right| \le C\epsilon$$

which indeed identifies τ as the desired partition $\mathcal{P}_0$. The estimate (2.3.2) follows immediately as a by-product of the analysis above. It remains to deal with the case of Riemann sum approximations

$$S'(\mathcal{P}) = \sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t}.$$

To this end, consider the difference

$$S'(\mathcal{P}) - S(\mathcal{P}) = \sum_{[s,t]\in\mathcal{P}} R_{s-,s} X_{s,t} + \Delta Y'_s \mathbb{X}_{s,t}$$

and then use arguments similar as those in the proof of Proposition 25 to see that

$$\text{(RRS)} \quad \lim_{|\mathcal{P}|\to 0} \big(S'(\mathcal{P}) - S(\mathcal{P})\big) = 0.$$

The rest is then clear. □

As an immediate corollary of (2.3.2) and (2.3.3), we have the following.

COROLLARY 32. *For a controlled rough path* (Y, Y'),

$$(Z_t, Z'_t) := \left(\int_0^t Y_{r-}\, d\mathbf{X}_r, Y_t\right)$$

is also a controlled rough path.

COROLLARY 33. *If* (Y, Y') *is a controlled rough path and* $Z_t = \int_0^t Y_{r-}\, d\mathbf{X}_r$, *then*

$$\Delta_t Z = \lim_{s \uparrow t} \int_s^t Y_{r-}\, d\mathbf{X}_r = Y_{t-} \Delta_t X + Y'_{t-} \Delta_t \mathbb{X},$$

where $\Delta_t \mathbb{X} = \lim_{s \uparrow t} \mathbb{X}_{s,t}$.

Though we avoid to write down the long expression for the bounds of $\|Z\|_p$, $\|Z'\|_p$ and $\|R^Z\|_{\frac{p}{2}}$, it can be easily derived from (2.3.3). The important point here is that we can again, for Z taking values in a suitable space, readily define

$$\int_0^t Z_{r-}\, d\mathbf{X}_r.$$

The rough integral defined above is also compatible with the Young integral. If X is a finite p-variation path for $p < 2$, we can construct càdlàg rough path $\mathbf{X}$ by

$$\mathbb{X}_{s,t} := \int_s^t (X_{r-} - X_s) \otimes dX_r,$$

where the right-hand side is understood as a Young integral.

PROPOSITION 34. *If* X, Y *are a càdlàg path of finite* p *and* q *variation respectively with* $\frac{1}{p} + \frac{1}{q} > 1$, *then for any* $\theta > 0$ *with* $\frac{1}{p} + \frac{1}{q} \geq \frac{1}{\theta}$,

$$Z_{s,t} := \int_s^t (Y_{r-} - Y_s)\, dX_r$$

has finite θ *variation. In particular,* $\mathbb{X}$ *has finite* $\frac{p}{2}$ *variation.*

PROOF. From Young's inequality,

$$|Z_{s,t}|^\theta \leq C \|X\|^\theta_{p,[s,t]} \|Y\|^\theta_{q,[s,t]}.$$

If $\frac{1}{p} + \frac{1}{q} \geq \frac{1}{\theta}$, the right-hand side is super-additive, which implies $\|Z\|_\theta < \infty$. $\square$

THEOREM 35. *If* X, Y *are càdlàg paths of finite* p*-variation for* $p < 2$, *then*

$$\int_0^t Y_{r-}\, d\mathbf{X}_r = \int_0^t Y_{r-}\, dX_r.$$

PROOF. The difference between Riemann sum approximation of corresponding integrals can be written as

$$S(\mathcal{P}) = \sum_{[s,t] \in \mathcal{P}} Y'_{s-} \tilde{\mathbb{X}}_{s,t}.$$

Choose $p < p' < 2$. From Young's inequality,

$$|\tilde{\mathbb{X}}_{s,t}| = \left|\int_s^t (X_{r-} - X_{s-}) \otimes dX_r\right| \leq C\|X^-\|_{p',[s,t]}\|X\|_{p',[s,t]},$$

(X^-, X) is a compatible pair, which implies for properly chosen $\mathcal{P}$,

$$|S(\mathcal{P})| \leq C\epsilon \sum_{[s,t]\in\mathcal{P}} \|X^-\|_{p,[s,t]}^{\frac{p}{p'}}\|X\|_{p,[s,t]}^{\frac{p}{p'}}.$$

Noting again that the term under summation sign is super-additive,

$$\text{(RRS)} \lim_{|\mathcal{P}|\to 0} S(\mathcal{P}) = 0. \qquad \square$$

REMARK 36 (Marcus canoncial rough integration). Given a càdlàg geometric rough path $\mathbf{X} = (X, \mathbb{X})$, and suitable f, say in $\in C^2$, we may regard $f(X)$ as a controlled rough path and so Theorem 31 yields a rough integral of the form

$$\int_0^T f(X_{t-})\, d\mathbf{X}_t.$$

We can adapt Marcus canonical integration to the geometric rough path case, taking into account the possibility of jumps in the area. As this will be discussed below in full detail, in the more involved setting of differential equations, we shall be brief: log-linear interpolation between jumps motivates the definition

$$\begin{aligned}
&\int_0^T f(X) \diamond d\mathbf{X} \\
&\quad := \int_0^T f(X_{t-})\, d\mathbf{X}_t \\
&\qquad + \sum_{t\in(0,T]} \Delta X_t \left\{\int_0^1 f(X_{t-} + \theta\Delta X_t) - f(X_{t-}) - \frac{1}{2} f'(X_{t-})(\Delta X_t)\right\} d\theta \\
&\qquad + \sum_{t\in(0,T]} \Delta_t \mathbb{A}\left\{\int_0^1 f'(X_{t-} + \theta\Delta X_t) - f'(X_{t-})\right\} d\theta.
\end{aligned}$$

Note that the expression in the first (resp., second) set of curly brackets is of order $O(|\Delta X|^2)$ [resp., $O(|\Delta X|)$] while $|\Delta X|$ (resp. $|\mathbb{A}|$) is summable of order p (resp. $p/2$), for $p < 3$, which is enough to guarantee that all sums are absolutely convergent. Let us make the link to Marcus canonical (stochastic) integration. To this end, assume the (geometric) rough path $\mathbf{X} = (X, \mathbb{X})$ is written as sum of a (in general, nongeometric: think Itô) rough path, say $\mathbf{X}^I = (X, \mathbb{X}^I)$, plus a pure second-level, finite 1-variation càdlàg perturbation of the form $(0, \frac{1}{2}\Gamma)$.

Then

$$\int_0^T f(X)\diamond d\mathbf{X}$$

$$:= \int_0^T f(X_{t-})\,d\mathbf{X}_t^I + \int_0^T f'(X_{t-})\,d\Gamma_t^c + \sum_{t\in(0,T]} f'(X_{t-})\Delta\Gamma_t$$

$$+ \sum_{t\in(0,T]} \Delta X_t\left\{\int_0^1 f(X_{t-}+\theta\Delta X_t) - f(X_{t-}) - \frac{1}{2}f'(X_{t-})(\Delta X_t)\right\}d\theta$$

$$+ \sum_{t\in(0,T]} \Delta_t\mathbb{A}\left\{\int_0^1 f'(X_{t-}+\theta\Delta X_t) - f'(X_{t-})\right\}d\theta.$$

Assuming furthermore $\mathbf{X}$ to be Marcus-like (so that $\Delta_t\mathbb{A}\equiv 0$) and also $\Delta\Gamma_t \equiv \Delta X_t^{\otimes 2}$, which is exactly what happens when X is a semimartingale and $\Gamma\equiv[X,X]$ its quadratic (co)variation, we recover the precise form of (1.2.5).

2.4. *Rough differential equations with jumps*. In the case of *continuous* RDEs, the difference between nongeometric (Itô-type) and geometric situations, is entirely captured in one's choice of the second-order information $\mathbb{X}$, so that both cases are handled with the *same* notion of (continuous) RDE solution. In the jump setting, the situation is different and a geometric notion of RDE solution requires additional terms in the equation in the spirit of Marcus's canonical (stochastic) equations [1, 22, 33, 34]. We now define both solution concepts for RDEs with jumps. (Of course, they coincide in absence of jumps, i.e. when $(\Delta X_s, \Delta\mathbb{X}_s)\equiv(0,0)$.)

DEFINITION 37. (i) For suitable $f=(f_1,\dots,f_d)$ and a càdlàg geometric p-rough path $\mathbf{X}=(X,\mathbb{X})\in\mathcal{W}_g^p$, call a path Z [better: controlled rough path $(Z,f(Z))$] solution to the *rough canonical equation*

$$dZ_t = f(Z_t)\diamond d\mathbf{X}_t$$

if, by definition,

$$Z_t = Z_0 + \int_0^t f(Z_{s-})\,d\mathbf{X}_s + \sum_{0<s\le t}\left\{\phi\left(f\Delta X_s + \frac{1}{2}[f,f]\Delta\mathbb{X}_s; Z_{s-}\right)\right.$$

$$\left. - Z_{s-} - f(Z_{s-})\Delta X_s - f'f(Z_{s-})\Delta\mathbb{X}_s\right\},$$

where, as in Section 1.2.3, $\phi(g,x)$ is the time 1 solution to $\dot{y}=g(y)$, $y(0)=x$. When $\mathbf{X}$ is Marcus-like, that is, $\mathbf{X}\in\mathcal{W}^p_{\mathrm{M}}$ this becomes[5]

$$Z_t = Z_0 + \int_0^t f(Z_{s-})\,d\mathbf{X}_s + \sum_{0<s\le t}\biggl\{\phi(f\Delta X_s, Z_{s-}) - Z_{s-} - f(Z_{s-})\Delta X_s - f'f(Z_{s-})\frac{1}{2}(\Delta X_s)^{\otimes 2}\biggr\}.$$

(ii) For suitable f and a càdlàg p-rough path call a path Z [or better: controlled rough path $(Z,f(Z))$] solution to the (*general*) *rough differential equation*

$$dZ_t = f(Z_{t-})\,d\mathbf{X}_t$$

if, by definition,

$$Z_t = Z_0 + \int_0^t f(Z_{s-})\,d\mathbf{X}_s.$$

We shall not consider the solution type (ii) further here.

THEOREM 38. *Fix initial data Z_0. Then Z is a solution to $dZ_t = f(Z_t)\diamond d\mathbf{X}_t$ if and only if $\tilde{Z}$ is a solution to the* (*continuous*) *RDE*

$$d\tilde{Z}_t = f(\tilde{Z}_t)\,d\tilde{\mathbf{X}}_t,$$

where $\tilde{\mathbf{X}}\in\mathcal{C}^p_g$ is constructed from $\mathbf{X}\in\mathcal{W}^p_g$ as in Theorem 20.

PROOF. (i) We illustrate the idea by considering X of finite 1-variation, with *one* jump at $\tau\in[0,T]$. This jump time becomes an interval $\tilde{I}=[a,a+\delta]\subset[0,\tilde{T}]=[0,T+\delta]$ in the stretched time scale. Now

$$\tilde{Z}_{0,\tilde{T}} \approx \sum_{[s,t]\in\mathcal{P}} f(\tilde{Z}_s)\tilde{X}_{s,t}$$

in the sense of (MRS) convergence, as $|\mathcal{P}|\to 0$. In particular, noting that $\tilde{X}_{s,t} = \frac{(t-s)}{\delta}\Delta X_\tau$ whenever $[s,t]\subset[a,a+\delta]$

$$\tilde{Z}_{a,a+\delta} = \lim_{|\tilde{\mathcal{P}}|\to 0}\sum_{[s,t]\in\tilde{\mathcal{P}}} f(\tilde{Z}_s)\tilde{X}_{s,t} = \frac{1}{\delta}\int_a^{a+\delta} f(\tilde{Z}_r)\Delta X_\tau\,dr$$

$$\Longrightarrow\quad \tilde{Z}_{a,a+\delta} = \phi(f\Delta X_\tau, \tilde{Z}_a) - \tilde{Z}_a.$$

On the other hand, by refinement of $\mathcal{P}$, we may insist that the end-points of $\tilde{I}$ are contained in $\mathcal{P}$ which thus has the form

$$\mathcal{P} = \mathcal{P}_1\cup\tilde{\mathcal{P}}\cup\mathcal{P}_2$$

[5]Note cancellation of $[f,f]\in\mathfrak{so}(d)$ with $\Delta\mathbb{X}_s = (\Delta X_s)^{\otimes 2}/2\in\mathrm{Sym}(d)$.

and so

$$\tilde{Z}_{0,\tilde{T}} \approx \sum_{[s,t]\in\mathcal{P}_1} f(\tilde{Z}_s)\tilde{X}_{s,t} + \sum_{[s,t]\in\tilde{\mathcal{P}}} f(\tilde{Z}_s)\tilde{X}_{s,t} + \sum_{[s,t]\in\mathcal{P}_2} f(\tilde{Z}_s)\tilde{X}_{s,t}$$

from which we learn, by sending $|\tilde{\mathcal{P}}| \to 0$, that

$$\tilde{Z}_{0,\tilde{T}} \approx \sum_{[s,t]\in\mathcal{P}_1} f(\tilde{Z}_s)\tilde{X}_{s,t} + \phi(f\Delta X_\tau, \tilde{Z}_a) - \tilde{Z}_a + \sum_{[s,t]\in\mathcal{P}_2} f(\tilde{Z}_s)\tilde{X}_{s,t}.$$

We now switch back to the original time scale. Of course, $Z \equiv \tilde{Z}$ on $[0,\tau)$ while $Z_t = \tilde{Z}_{t+\delta}$ for $t \in [\tau, T]$ and in particular

$$\begin{aligned} Z_{0,T} &= \tilde{Z}_{0,\tilde{T}}, \\ Z_{\tau-} &= \tilde{Z}_a, \\ Z_\tau &= \tilde{Z}_{a+\delta}. \end{aligned}$$

But then, with $\mathcal{P}_1'$ and $\mathcal{P}_2'$ partitions of $[0,\tau]$ and $[\tau, T]$, respectively,

$$\begin{aligned} Z_{0,T} &\approx \sum_{\substack{[s',t']\in\mathcal{P}_1' \\ t'<\tau}} f(Z_{s'})X_{s',t'} + \sum_{[s',t']\in\mathcal{P}_2'} f(Z_{s'})X_{s',t'} + \phi(f\Delta X_\tau, Z_{\tau-}) - Z_{\tau-} \\ &\approx \sum_{[s',t']\in\mathcal{P}'} f(Z_{s'})X_{s',t'} + \phi(f\Delta X_\tau, \tilde{Z}_a) \\ &\quad + \{\phi(f\Delta X_\tau, Z_{\tau-}) - Z_{\tau-} - f(Z_{\tau-})\Delta X_\tau\} \end{aligned}$$

since $f(Z_{s'})X_{s',\tau} \to f(Z_{\tau-})\Delta X_\tau$ as $|\mathcal{P}'| \to 0$, with $[s',\tau] \in \mathcal{P}'$. By passing to the (RRS) limit, find

$$Z_{0,T} = \int_0^T f(Z_s^-)\,dX + \{\phi(f\Delta X_\tau, Z_{\tau-}) - Z_{\tau-} - f(Z_{\tau-})\Delta X_\tau\}.$$

This argument extends to countable many jumps. We want to show that

$$\begin{aligned} Z_T &= Z_0 + \int_0^T f(Z_{s-})\,dX_s + \sum_{0<s\le T}\{\cdots\} \\ &= Z_0 + (\text{RRS}) \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} f(Z_s)X_{s,t} + \lim_{\eta\downarrow 0} \sum_{\substack{s\in(0,T]: \\ |\Delta X_s|>\eta}} \{\cdots\}. \end{aligned}$$

What we know is (MRS)-convergence of the time-changed problem. That is, given $\varepsilon > 0$, there exists δ s.t. $|\mathcal{P}| < \delta$ implies

$$\tilde{Z}_{0,\tilde{T}} \approx_\varepsilon \sum_{[s,t]\in\mathcal{P}} f(\tilde{Z}_s)\tilde{X}_{s,t},$$

where $a \approx_\varepsilon b$ means $|a-b| \le \varepsilon$. For fixed $\eta > 0$, include all (but only finitely many, say N) points $s \in (0,T]$: $|\Delta X_s| > \eta$ in $\mathcal{P}$, giving rise to $(\tilde{\mathcal{P}}_j : 1 \le j \le N)$. Sending the mesh of these to zero gives, as before,

$$Z_{0,T} \approx_\varepsilon \sum_{[s',t'] \in \mathcal{P}'} f(Z_{s'}) X_{s',t'} + \sum_{\substack{s \in (0,T]: \\ |\Delta X_s| > \eta}} \{\phi(f \Delta X_s, Z_{s-}) - Z_{s-} - f(Z_{s-}) \Delta X_s\}.$$

In fact, due to summability of $\sum_{s \in (0,T]} \{\cdots\}$, we can pick $\eta > 0$ such that

$$Z_{0,T} \approx_{2\varepsilon} \sum_{[s',t'] \in \mathcal{P}'} f(Z_{s'}) X_{s',t'} + \sum_{s \in (0,T]:} \{\phi(f \Delta X_s, Z_{s-}) - Z_{s-} - f(Z_{s-}) \Delta X_s\}$$

and this is good enough to take the (RRS) lim as $|\mathcal{P}'| \to 0$.

(ii) We now consider the case of $\mathbf{X} = (1, X, \mathbb{X}) = \exp(X + A) \in \mathcal{W}_g^p$, again starting with *one* jump at $\tau \in [0,T]$. As above, the jump time becomes an interval $[a, a+\delta] \subset [0, \tilde{T}] = [0, T+\delta]$ in the stretched time scale. Now

$$\tilde{Z}_{0,\tilde{T}} \approx \sum_{[s,t] \in \mathcal{P}} f(\tilde{Z}_s) \tilde{X}_{s,t} + f' f(\tilde{Z}_s) \tilde{\mathbb{X}}_{s,t}$$

in the sense of (MRS) convergence, as $|\mathcal{P}| \to 0$. Recall, by the very construction of $\tilde{\mathbf{X}} \in \mathcal{C}_g^p$, whenever $[s,t] \subset [a, a+\delta]$

$$\begin{aligned}
\tilde{\mathbf{X}}_t &= \exp^{(2)}\left(\log^{(2)} \mathbf{X}_{\tau-} + \frac{t-a}{\delta} \log^{(2)} \Delta \mathbf{X}_\tau\right), \\
\tilde{\mathbf{X}}_{s,t} &= \exp^{(2)}\left(\frac{t-s}{\delta} \log^{(2)} \Delta \mathbf{X}_\tau\right) \\
&= 1 + \frac{t-s}{\delta} \Delta X_\tau \\
&\quad + \left(\frac{t-s}{\delta} \Delta A_\tau + \frac{1}{2}\left(\frac{t-s}{\delta} \Delta X_\tau\right)^{\otimes 2}\right) \\
&\equiv 1 + \tilde{X}_{s,t} + \tilde{\mathbb{X}}_{s,t}.
\end{aligned}$$

We want to compute $\tilde{Z}_{a,a+\delta}$ and by reparametrisation we may take $[a, a+\delta] = [0,1]$ without loss of generality. Then $\tilde{Z} = \int_0^\cdot f(\tilde{Z})\, d\tilde{\mathbf{X}}$ on the unit interval is equivalent to, for $s, t \in [0,1]$,

$$\tilde{Z}_{s,t} = f(\tilde{Z}_s) \Delta X_\tau (t-s) + f' f(\tilde{Z}_s) \Delta A_\tau (t-s) + o(t-s).$$

Division by $t - s$, and taking $t \downarrow s$, shows that on $\tilde{Z}$ solves on $[0,1]$ the ordinary differential equation

$$d\tilde{Z} = \big(f(\tilde{Z}_s) \Delta X_\tau + f' f(\tilde{Z}_s) \Delta A_\tau\big)\, dt.$$

And it follows that

$$\tilde{Z}_{a,a+\delta} = \phi\left(f \Delta X_s + \frac{1}{2}[f, f] \Delta \mathbb{X}_s; \tilde{Z}_a\right) - \tilde{Z}_a.$$

It is then straightforward to adapt the subsequent steps of (i) to this setting. $\square$

Theorem 38 above settles the problem, already hinted to in Section 1.2.9, of giving honest meaning in the sense of integral equations to RDEs with jumps (in the spirit of Marcus). Having such explicit meaning is clearly important for any subsequent analysis. More specifically, the following corollary will be crucial in order to compute the expected signature of Lévy rough paths; cf. Theorem 53.

COROLLARY 39. *For a càdlàg rough path* $\mathbf{X} = 1 + X + \mathbb{X} = \exp(X + \mathbb{A}) \in \mathcal{W}_g^p$ *for* $p \in [2, 3)$, *the minimal jump extension* $\mathbf{X}^{(n)}$ *taking values in* $G^{(n)}(\mathbb{R}^d)$ *satisfies the Marcus-type differential equation*:

$$\mathbf{X}_t^{(n)} = 1 + \int_0^t \mathbf{X}_{r-}^{(n)} \otimes d\mathbf{X}_r + \sum_{0<s\le t} \mathbf{X}_{s-}^{(n)} \otimes \{\exp^{(n)}(\log^{(2)} \Delta\mathbf{X}_s) - \Delta\mathbf{X}_s\}, \tag{2.4.1}$$

where the integral is understood as a rough integral and summation term is well defined as an absolutely summable series.

PROOF. This follows from Theorem 38 and (2.2.4). □

2.5. *Rough path stability.* We briefly discuss stability of rough integration and rough differential equations. In the context of càdlàg rough integration, Section 2.3, it is a natural to estimate $Z^1 - Z^2$, in p-variation norm, where

$$Z^i = \int_0 Y^i \, d\mathbf{X}^i \qquad \text{for } i = 1, 2.$$

Now, the analysis presented in Section 2.3 adapts without difficulty to this situation. For instance, when $Y^i = F(X^i)$, one easily finds

$$|Z^1 - Z^2|_{p\text{-var}} \le C_{F,M}(|X_0^1 - X_0^2| + \|X^1 - X^2\|_{p\text{-var}} + \|\mathbb{X}^1 - \mathbb{X}^2\|_{p\text{-var}})$$

provided $F \in C^2$ and $|X_0^i| + \|X^i\|_{p\text{-var}} + \|\mathbb{X}^i\|_{p\text{-var}} \le M$. (The situation can be compared with [14], Section 4.4, where the analogous estimate is in the α-Hölder setting.)

The situation is somewhat different in the case of Marcus-type RDEs, $dY_t^i = f(Y_t^i) \diamond d\mathbf{X}_t^i$. In principle, the difference $Y^1 - Y^2$, in p-variation norm, is controlled, as above uniformly on bounded sets, by

$$\|\tilde{X}^1 - \tilde{X}^2\|_{p\text{-var}} + \|\tilde{\mathbb{X}}^1 - \tilde{\mathbb{X}}^2\|_{p\text{-var}},$$

where $\tilde{\mathbf{X}}^i = (\tilde{X}^i, \tilde{\mathbb{X}}^i) \in \mathcal{C}_g^p$ is constructed from $\mathbf{X}^i \in \mathcal{W}_g^p$ as in Theorem 20. (This follows immediately from the continuous rough paths theory.) However, since the time-change depends in a complicated way on the underlying jumps it seems unlikely that Wong–Zakai-type results such as those obtained in [22] are trivially recovered along these lines. Instead, we suspect the correct approach (left for subsequent work) is to use Theorem 38 to rewrite the Marcus-type RDE as an honest integral equation (in the sense of Definition 37), followed by the direct stability estimate of all involved quantities.

2.6. *Rough versus stochastic integration.* Consider a d-dimensional Lévy process X_t enhanced with

$$\mathbb{X}_{s,t} := (\text{Itô}) \int_{(s,t]} (X_{r-} - X_s) \otimes dX_r.$$

We show that rough integration against the Itô lift actually yields a standard stochastic integral in Itô sense. An immediate benefit, say when taking $Y = f(X)$ with $f \in C^2$, is the universality of the resulting stochastic integral, defined on a set of full measure simultaneously for all such integrands.

THEOREM 40. *Let X be a d-dimensional Lévy process, and consider adapted processes Y and Y' such that (Y, Y') is a controlled rough path. Then Itô and the rough integral coincide,*

$$\int_{(0,T]} Y_{s-}\, dX_s = \int_0^T Y_{s-}\, d\mathbf{X}_s \qquad a.s.$$

PROOF. By Theorem 31, there exist partitions $\mathcal{P}_n$ with

$$\left| S(\mathcal{P}_n) - \int_0^T Y_{s-}\, d\mathbf{X}_s \right| \le \frac{1}{n},$$

where

$$S(\mathcal{P}_n) := \sum_{[s,t]\in\mathcal{P}_n} Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t}.$$

Let $X_t = M_t + V_t$ be the Lévy–Itô decomposition with martingale M and bounded variation part V. Since (V^-, V), (V^-, M), (M^-, V) are compatible pairs, we can choose the corresponding τ_n for $\epsilon = \frac{1}{n}$ from their compatibility. Without loss of generality, we can assume $\mathcal{P}_{n-1} \cup \tau_n \cup D_n \subset \mathcal{P}_n$, where D_n is the nth dyadic partition. We know from general stochastic integration theory that, possibly along some subsequence, almost surely,

$$S'(\mathcal{P}_n) = \sum_{[s,t]\in\mathcal{P}_n} Y_s X_{s,t} \to \int_{(0,T]} Y_{s-}\, dX_s \qquad \text{as } n \to \infty$$

Thus, it suffices to prove that almost surely, along some subsequence,

$$S''(\mathcal{P}_n) = \sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{X}_{s,t} \to 0.$$

Now

$$\mathbb{X}_{s,t} = \mathbb{M}_{s,t} + \mathbb{V}_{s,t} + \int_{(s,t]} (M_{r-} - M_s) \otimes dV_r + \int_{(s,t]} (V_{r-} - V_s) \otimes dM_r.$$

Using a similar argument as in Theorem 35,

$$\sum_{[s,t]\in\mathcal{P}_n} Y'_s \biggl(\mathbb{V}_{s,t} + \int_{(s,t]} (M_{r-} - M_s) \otimes dV_r + \int_{(s,t]} (V_{r-} - V_s) \otimes dM_r\biggr) \to 0.$$

We are left to show that

$$\sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{M}_{s,t} \to 0.$$

By the very nature of Itô lift,

$$\mathrm{Sym}(\mathbb{M}_{s,t}) = \frac{1}{2} M_{s,t} \otimes M_{s,t} - \frac{1}{2}[M, M]_{s,t}$$

and it follows from standard (convergence to) quadratic variation results for semimartingale (due to Föllmer [10]) that one is left with

$$\sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{A}_{s,t} \to 0,$$

where $\mathbb{A}_{s,t} = \mathrm{Anti}(\mathbb{M}_{s,t})$. At this point, let us first assume that $|Y'|_\infty \le K$ uniformly in ω. We know from Theorem 11 (or Corollary 42)

$$\mathbb{E}\bigl[|\mathbb{A}_{s,t}|^2\bigr] \le C|t-s|^2 \tag{2.6.1}$$

and using the standard martingale argument (orthogonal increment property),

$$\mathbb{E}\biggl[\biggl|\sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{A}_{s,t}\biggr|^2\biggr] = \sum_{[s,t]\in\mathcal{P}_n} \mathbb{E}\bigl[|Y'_s \mathbb{A}_{s,t}|^2\bigr] \le K^2 C \sum_{[s,t]\in\mathcal{P}_n} |t-s|^2 = \mathcal{O}\bigl(|\mathcal{P}_n|\bigr)$$

which implies, along some subsequence, almost surely,

$$\sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{A}_{s,t} \to 0. \tag{2.6.2}$$

Finally, for unbounded Y', introduce stopping times

$$T_K = \inf\Bigl\{t \in [0,T] : \sup_{s\in[0,t]} |Y'_s| \ge K\Bigr\}.$$

Similarly, as in the previous case,

$$\mathbb{E}\biggl[\biggl|\sum_{[s,t]\in\mathcal{P}_n, s\le T_K} Y'_s \mathbb{A}_{s,t}\biggr|^2\biggr] = \mathcal{O}(\mathcal{P}_n).$$

Thus, almost surely on the event $\{T_K > T\}$

$$\sum_{[s,t]\in\mathcal{P}_n} Y'_s \mathbb{A}_{s,t} \to 0$$

and sending $K \to \infty$ completes the proof. $\square$

We remark that the identification of rough with stochastic integrals is by no means restricted to Lévy processes, and the method of proof here obviously applies to a semimartingale situation. As a preliminary remark, one can always drop the bounded variation part (and thereby gain integrability). Then, with finite p-variation rough path regularity, for some $p < 3$, of (Itô, by Proposition 16 equivalently: Stratonovich) lift (see Section 3.2.2), the proof proceeds along the same lines until the moment where one shows (2.6.2). For the argument then to go through, one only needs

$$\sum_{[s,t]\in\mathcal{P}} \mathbb{E}\big[|\mathbb{A}_{s,t}|^2\big] \to 0 \qquad \text{as } |\mathcal{P}| \to 0,$$

which follows from (2.6.1), an estimate which will be extended to general classes of Markov jump processes in Section 3.2.1. That said, we note that much less than (2.6.1) is necessary, and clearly this has to be exploited in a general semimartingale context.

3. Stochastic processes as rough paths and expected signatures.

3.1. *Lévy processes.*

3.1.1. *A Lévy–Kintchine formula and rough path regularity.* In this section, we assume (X_t) is a d-dimensional Lévy process with triplet (a, b, K). The main insight of this section is that the expected signature is well suited to study rough path regularity. More precisely, we consider the Marcus canonical signature $S = S(X)$, given as a solution to

$$dS = S \otimes \diamond dX,$$
$$S_0 = (1, 0, 0 \dots) \in T\big((\mathbb{R}^d)\big).$$

With $S_{s,t} = S_s^{-1} \otimes S_t$ as usual, this gives random group-like elements

$$S_{s,t} = \big(1, \mathbf{X}^1_{s,t}, \mathbf{X}^2_{s,t}, \dots\big) = \big(1, X_{s,t}, \mathbb{X}^{\mathrm{M}}_{s,t}, \dots\big)$$

and then the step-n signature of $X|_{[s,t]}$ by projection,

$$\mathbf{X}^{(n)}_{s,t} = \big(1, \mathbf{X}^1_{s,t}, \dots, \mathbf{X}^n_{s,t}\big) \in G^{(N)}\big(\mathbb{R}^d\big).$$

The expected signature is obtained by taking component-wise expectation and exists under a natural assumption on the tail behaviour of the Lévy measure

$K = K(dy)$. In fact, it takes "Lévy–Kintchine form" as detailed in the following theorem. We stress that fact that the expected signature contains significant information about *the process* $(X_t : 0 \le t \le T)$, where a classical moment generating function of X_T only carries information about the *random variable* X_T.

THEOREM 41 (Lévy–Kintchine formula). *If the measure* $K 1_{|y|\ge 1}$ *has moments up to order* N, *then*

$$\mathbb{E}[\mathbf{X}_{0,T}^{(N)}] = \exp(CT)$$

with the tensor algebra valued exponent

$$C = \left(0, b + \int_{|y|\ge 1} yK(dy), \frac{a}{2} + \int \frac{y^{\otimes 2}}{2!} K(dy), \dots, \int \frac{y^{\otimes N}}{N!} K(dy)\right) \in T^{(N)}(\mathbb{R}^d).$$

In particular, if $K I_{|y|\ge 1}$ *has finite moments of all orders, the expected signature is given by*

$$\mathbb{E}[S(X)_{0,T}] = \exp\left[T\left(b + \frac{1}{2}a + \int (\exp(y) - 1 - y1_{|y|<1})K(dy)\right)\right] \in T((\mathbb{R}^d)).$$

The proof is based on the Marcus SDE $dS = S \otimes \diamond dX$ in $T^{(N)}(\mathbb{R}^d)$, so that $\mathbf{X}_{0,T}^{(n)} = S$ and will be given in detail below. We note that Fawcett's formula [3, 9, 29] for the expected value of iterated Stratonovich integrals of d-dimensional Brownian motion (with covariance matrix $a = I$ in the aforementioned references)

$$\mathbb{E}[S(B)_{0,T}] = \mathbb{E}\left[\left(1, \int_{0<s<T} \circ dB, \int_{0<s<t<T} \circ dB \otimes \circ dB, \dots\right)\right] = \exp\left[\frac{T}{2}a\right]$$

is a special case of the above formula. Let us in fact give a (novel) elementary argument for the validity of Fawcett's formula. The form $\mathbb{E}[S(B)_{0,T}] = \exp(TC)$ for some $C \in T((\mathbb{R}^d))$ is actually an easy consequence of independent increments of Brownian motion. But Brownian scaling implies the k^{th} tensor level of $S(B)_{0,T}$ scales as $T^{k/2}$, which already implies that C must be a pure 2-tensor. The identification $C = a/2$ is then an immediate computation. Another instructive case which allows for an elementary proof is the case when X is a compound Poisson process, that is, $X_t = \sum_{i=1}^{N_t} J_i$ for some i.i.d. d-dimensional random variables J_i and N_t a Poisson process with intensity λ. In Lévy terminology, one has a triplet $(0, 0, K)$ where K is λ times the law of J_i. Since jumps are to be traversed along straight lines, Chen's rule implies

$$\mathbb{E}[S^N(X)_{0,1}|N_1 = n] = \mathbb{E}[\exp(J_1) \otimes \cdots \otimes \exp(J_n)] = \mathbb{E}[\exp(J_1)]^{\otimes n}$$

and thus

$$\mathbb{E}[S^N(X)_{0,1}] = \exp[(\lambda(\mathbb{E}\exp(J) - 1))]$$

which gives, with all integrations over $\mathbb{R}^d$,

$$C = \int (\exp(y) - 1) K(dy).$$

Before turning to the proof of Theorem 41, we give the following application. It relies on the fact that the expected signature allows to easily extract information about the stochastic area.

COROLLARY 42. *Let X be a d-dimensional Lévy process. Then, for any $p > 2$, a.s.*

$$(X, \mathbb{X}^{\mathrm{M}}) \in \mathcal{W}^p_{\mathrm{M}}([0,T], \mathbb{R}^d) \qquad a.s.$$

We call the resulting Marcus-like (geometric) rough path the Marcus lift of X.

PROOF. Without loss of generality, all jumps have size less than 1. [This amounts to drop a bounded variation term in the Itô–Lévy decomposition. This does not affect the p-variation sample path properties of X, nor—in view of basic Young (actually Riemann–Stieltjes) estimates—those of $\mathbb{X}^{\mathrm{M}}$.] We establish the desired rough path regularity as an application of Proposition 17, which requires us to show

$$\mathbb{E}|X_{s,t}|^2 \lesssim |t-s|,$$
$$\mathbb{E}|\mathbb{A}_{s,t}|^2 \lesssim |t-s|^2.$$

While the first estimate is immediate from the L^2-isometry of stochastic integrals against Poisson random measures (drift and the Brownian component obviously pose no problem), the second one is more subtle in nature and indeed fails—in the presence of jumps—when $\mathbb{A}$ is replaced by the full second level $\mathbb{X}^{\mathrm{M}}$. (To see this, take $d = 1$ so that $\mathbb{X}^{\mathrm{M}}_{s,t} = X^2_{s,t}/2$ and note that even for the standard Poisson process $\mathbb{E}|X_{s,t}|^4 \lesssim |t-s|$ but not $|t-s|^2$.)

It is clearly enough to consider $\mathbb{A}^{i,j}_{s,t}$ for indices $i \neq j$. It is enough to work with $S^4(X) =: \mathbf{X}$. Using the geometric nature of $\mathbf{X}$, by using the shuffle product formula,

$$\begin{aligned}(\mathbb{A}^{i,j}_{s,t})^2 &= \frac{1}{4}(\mathbb{X}^{i,j}_{s,t} - \mathbb{X}^{j,i}_{s,t})(\mathbb{X}^{i,j}_{s,t} - \mathbb{X}^{j,i}_{s,t}) \\ &= \mathbf{X}^{iijj}_{s,t} - \mathbf{X}^{ijji}_{s,t} - \mathbf{X}^{jiij}_{s,t} + \mathbf{X}^{jjii}_{s,t}.\end{aligned}$$

On the other hand,

$$\mathbb{E}\mathbf{X}_{s,t} = \exp[(t-s)C] = 1 + (t-s)C + O(t-s)^2$$

so that it is enough to check that $C^{iijj} - C^{ijji} - C^{jiij} + C^{jjii} = 0$. But this is obvious from the symmetry of

$$\pi_4 C = \frac{1}{4!}\int y^{\otimes 4} K(dy). \qquad \square$$

We now give the proof of the Lévy–Kintchine formula for the expected signature of Lévy processes. We first state some lemmas required.

The following lemma, a generalization of [40], Chapter 1, Theorem 38, is surely well known but since we could not find a precise reference we include the short proof.

LEMMA 43. *Let F_s be a càglàd adapted process with $\sup_{0<s\le t}\mathbb{E}[|F_s|] < \infty$ and g be a measurable function with $|g(x)| \le C|x|^k$ for some $C > 0$, $k \ge 2$ and $g \in L^1(K)$. Then*

$$\mathbb{E}\Big[\sum_{0<s\le t} F_s g(\Delta X_s)\Big] = \int_0^t \mathbb{E}[F_s]\,ds \int_{\mathbb{R}^d} g(x)K(dx).$$

PROOF. At first, we prove the following:

$$\mathbb{E}\Big[\sum_{0<s\le t} |F_s|\big|g(\Delta X_s)\big|\Big] \le t\|g\|_1 \sup_{0<s\le t}\mathbb{E}\big[|F_s|\big]. \tag{3.1.1}$$

To this end, w.l.o.g., we can assume g vanishes in a neighbourhood of zero. The general case follows by an application of Fatou's lemma. Also, it is easy to check the inequality when F_s is a simple predictable process. For general F_s, we choose a sequence of simple predictable process $F_s^n \to F_s$ pointwise. Since there are only finitely many jumps away from zero, we see that

$$\sum_{0<s\le t} \big|F_s^n\big|\big|g(\Delta X_s)\big| \to \sum_{0<s\le t} |F_s|\big|g(\Delta X_s)\big| \qquad \text{a.s.}$$

and the claim follows again by Fatou's lemma.

Now, define $\bar g = \int_{\mathbb{R}^d\setminus 0} g(x)K(dx)$ and $M_t = \sum_{0<s\le t} g(\Delta X_s) - t\bar g$. Then it is easy to check that M_t is a martingale. Also,

$$N_t := \int_{(0,t]} F_s\,dM_s = \sum_{0<s\le t} F_s g(\Delta X_s) - \bar g \int_0^t F_s\,ds$$

is a local martingale. From (3.1.1), $\mathbb{E}[\sup_{0<s\le t}|N_s|] < \infty$. So, N_t is a martingale, which thereby implies that $\mathbb{E}[N_t] = 0$, completing the proof. $\square$

LEMMA 44. *If the measure $K\mathbb{I}_{|y|\ge 1}$ has moments up to order N, then with $S_t = S^N(X)_{0,t}$,*

$$\mathbb{E}\Big[\sup_{0<s\le t}|S_s|\Big] < \infty.$$

PROOF. We will prove it by induction on N. For $N = 1$, $S_t = \mathbf{1} + X_t$, and the claim follows from the classical result that $\mathbb{E}[\sup_{0<s\le t} |X_s|] < \infty$ iff $K I_{|y|\ge 1}$ has the finite first moment. Now, note that

$$S_t = 1 + \int_0^t \pi_{N,N-1}(S_{r-}) \otimes dX_r + \frac{1}{2}\int_0^t \pi_{N,N-1}(S_r) \otimes a\,dr + \sum_{0<s\le t} \pi_{N,N-1}(S_{r-}) \otimes \{e^{\Delta X_s} - \Delta X_s - 1\},$$

where $\pi_{N,N-1} : T_1^N(\mathbb{R}^d) \to T_1^{N-1}(\mathbb{R}^d)$ is the projection map. From the induction hypothesis and Lemma 43, the last two terms on the right-hand side has finite expectation in supremum norm. Using Lévy–Itô decomposition,

$$\int_0^t \pi_{N,N-1}(S_{r-}) \otimes dX_r = \int_0^t \pi_{N,N-1}(S_{r-}) \otimes dM_r + \int_0^t \pi_{N,N-1}(S_{r-}) \otimes b\,dr + \sum_{0<s\le t} \pi_{N,N-1}(S_{r-}) \otimes \Delta X_s 1_{|\Delta X_s|\ge 1},$$

where M is the martingale. Again by the induction hypothesis and Lemma 43, the last two terms are of finite expectation in supremum norm. Finally,

$$L_t = \int_0^t \pi_{N,N-1}(S_{r-}) \otimes dM_r$$

is a local martingale. By the Burkholder–Davis–Gundy inequality and noting that

$$[M]_t = at + \sum_{0<s\le t} (\Delta X_s)^2 1_{|\Delta X_s|<1}$$

we see that

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{0<s\le t} |L_s|\Big] \\
&\quad\lesssim \mathbb{E}\Big[\Big\{\int_0^t \big|\pi_{N,N-1}(S_{r-})\big|^2 d[M]_r\Big\}^{\frac{1}{2}}\Big] \\
&\quad\lesssim \mathbb{E}\Big[\Big\{\int_0^t \big|\pi_{N,N-1}(S_{r-})\big|^2 dr\Big\}^{\frac{1}{2}}\Big] \\
&\qquad + \mathbb{E}\Big[\Big\{\sum_{0<r\le t} \big|\pi_{N,N-1}(S_{r-})\big|^2 |\Delta X_r|^2 1_{|\Delta X_r|<1}\Big\}^{\frac{1}{2}}\Big] \\
&\quad\lesssim \mathbb{E}\Big[\sup_{r\le t}\big|\pi_{N,N-1}(S_{r-})\big|\Big] \\
&\qquad + \mathbb{E}\Big[\Big\{\sup_{r\le t}\big|\pi_{N,N-1}(S_{r-})\big|\Big\}^{\frac{1}{2}}\Big\{\sum_{0<r\le t} \big|\pi_{N,N-1}(S_{r-})\big| |\Delta X_r|^2 1_{|\Delta X_r|<1}\Big\}^{\frac{1}{2}}\Big]
\end{aligned}$$

$$\lesssim \mathbb{E}\Big[\sup_{r\le t}\big|\pi_{N,N-1}(S_{r-})\big|\Big] + \mathbb{E}\Big[\sup_{r\le t}\big|\pi_{N,N-1}(S_{r-})\big|\Big] + \mathbb{E}\Big[\sum_{0<r\le t}\big|\pi_{N,N-1}(S_{r-})\big||\Delta X_r|^2 1_{|\Delta X_r|<1}\Big],$$

where in the last line, we have used $\sqrt{ab} \lesssim a+b$. Again by the induction hypothesis and Lemma 43, we conclude that

$$\mathbb{E}\Big[\sup_{0<s\le t}|L_s|\Big] < \infty$$

completing the proof. □

PROOF OF THEOREM 41. As before,

$$S_t = 1 + \int_0^t S_{r-}\otimes dM_r + \int_0^t S_r \otimes \left(b + \frac{a}{2}\right) dr + \sum_{0<s\le t} S_{s-}\otimes \{e^{\Delta X_s} - \Delta X_s 1_{|\Delta X_s|<1} - 1\}.$$

By Lemma 44, $\int_0^t S_{r-}\otimes dM_r$ is indeed a martingale. Also note that S_t has a jump iff X_t has a jump, so that almost surely $S_{t-} = S_t$. Thanks to Lemma 43,

$$\mathbb{E}S_t = 1 + \int_0^t \mathbb{E}S_r \otimes \left(b + \frac{a}{2}\right) dr + \int_0^t \mathbb{E}S_r dr \otimes \int \left(e^y - y1_{|y|<1} - 1\right) K(dy)$$

and solving this linear ODE in $T_1^N(\mathbb{R}^d)$ completes the proof. □

3.1.2. *Lévy rough paths*. Corollary 42 tells us that the Marcus lift of some d-dimensional Lévy process X has sample paths of finite p-variation with respect to the CC norm on $G^{(2)}$, that is,

$$\mathbf{X}^{\mathrm{M}} := \left(1, X, \mathbb{X}^{\mathrm{M}}\right) \in W_{\mathrm{g}}^{p}\left([0,T], G^{(2)}(\mathbb{R}^d)\right).$$

It is clear from the nature of Marcus integration that $\mathbf{X}^{\mathrm{M}}_{s,t}$ is $\sigma(X_r : r < s \le t)$-measurable. It easily follows that $\mathbf{X}^{\mathrm{M}}$ is a Lie group-valued Lévy process, with values in the Lie group $G^{(2)}(\mathbb{R}^d)$, and in fact a Lévy rough path in the following sense.

DEFINITION 45. Let $p \in [2,3)$. A $G^{(2)}(\mathbb{R}^d)$-valued process $(\mathbf{X})$ with (càdlàg) rough sample paths $\mathbf{X}(\omega) \in W_{\mathrm{g}}^{p}$ a.s. (on any finite time horizon) is called *Lévy p-rough path* iff it has stationary independent left-increments [given by $\mathbf{X}_{s,t}(\omega) = \mathbf{X}_s^{-1}\otimes \mathbf{X}_t$].

Not every Lévy rough path arises as Marcus lift of some d-dimensional Lévy process. For instance, the *pure area Poisson process* and then *the noncanonical*

Brownian rough path ("*Brownian motion in a magnetic field*") from Example 15 plainly do not arise from iterated Marcus integration.

Given any Lévy rough path $\mathbf{X} = (1, X, \mathbb{X})$, it is clear that its projection $X = \pi_1(\mathbf{X})$ is a classical Lévy process on $\mathbb{R}^d$ which then admits, thanks to Corollary 42, a Lévy rough path lift $\mathbf{X}^{\mathrm{M}}$. This suggests the following terminology. We say that $\mathbf{X}$ is a *canonical Lévy rough path* if $\mathbf{X}$ and $\mathbf{X}^{\mathrm{M}}$ are indistinguishable, and call $\mathbf{X}$ a *noncanonical Lévy rough path* otherwise.

Let us also note that there are $G^{(2)}(\mathbb{R}^d)$-valued Lévy processes which are not *a Lévy p-rough path* in the sense of the above definition, for they may fail to have finite p-variation for $p \in [2, 3)$ (and thereby missing the in rough path theory crucial link between regularity and level of nilpotency, $[p] = 2$). To wit, *area-valued Brownian motion*

$$\mathbf{X}_t := \exp^{(2)}\big(B_t[e_1, e_2]\big) \in G^{(2)}\big(\mathbb{R}^d\big)$$

is plainly a $G^{(2)}(\mathbb{R}^d)$-valued Lévy process, but

$$\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} \|\mathbf{X}_{s,t}\|_{\mathrm{CC}}^p \sim \sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |B_{s,t}|^{p/2} < \infty$$

if and only if $p > 4$.

REMARK 46. One could define $G^4(\mathbb{R}^d)$-valued Lévy rough paths, with p-variation regularity where $p \in [4, 5)$, an example of which is given by area-valued Brownian motion. But then again not every $G^4(\mathbb{R}^d)$-valued Lévy process will be a $G^4(\mathbb{R}^d)$-valued Lévy rough path and so on. In what follows, we remain in the step 2 setting of Definition 45.

We now characterize Lévy rough paths among $G^{(2)}(\mathbb{R}^d)$-valued Lévy processes, themselves characterized by Hunt's theory of the Lie group-valued Lévy processes; cf. Section 1.2.8. To this end, let us recall $G^{(2)}(\mathbb{R}^d) = \exp(\mathfrak{g}^2(\mathbb{R}^d))$, where

$$\mathfrak{g}^{(2)}\big(\mathbb{R}^d\big) = \mathbb{R}^d \oplus \mathfrak{so}(d).$$

For $g \in G^{(2)}(\mathbb{R}^d)$, let $|g|$ be the Euclidean norm of $\log g \in \mathfrak{g}^{(2)}(\mathbb{R}^d)$. With respect to the canonical basis, any element in $\mathfrak{g}^{(2)}(\mathbb{R}^d)$ can be written as in coordinates as $(x^v)_{v\in J}$ where

$$J := \{i : 1 \le i \le d\} \cup \{jk : 1 \le j < k \le d\}.$$

Write also

$$I := \{i : 1 \le i \le d\}.$$

THEOREM 47. *Every $G^{(2)}(\mathbb{R}^d)$-valued Lévy process* $(\mathbf{X})$ *is characterized by a triplet* $(\mathbf{a}, \mathbf{b}, \mathbf{K})$ *with*

$$\mathbf{a} = (a^{v,w} : v, w \in J),$$
$$\mathbf{b} = (b^v : v \in J),$$
$$\mathbf{K} \in \mathcal{M}(G^{(2)}(\mathbb{R}^d)): \qquad \int_{G^{(2)}(\mathbb{R}^d)} (|g|^2 \wedge 1)\mathbf{K}(dg).$$

The projection $X := \pi_1(\mathbf{X})$ *is a standard* d*-dimensional Lévy process*, *with triplet*

$$(a, b, K) := ((a^{i,j} : i, j \in I), (b^k : k \in I), (\pi_1)_* \mathbf{K}),$$

where K *is the push forward of* $\mathbf{K}$ *under the projection map*. *Call* $(\mathbf{a}, \mathbf{b}, \mathbf{K})$ an enhanced Lévy triplet, *and* $\mathbf{X}$ an enhanced Lévy process.

PROOF. This is really a special case of Hunt's theory. Let us detail, however, an explicit construction which will be useful later on: every $G^{(2)}(\mathbb{R}^d)$-valued Lévy process $\mathbf{X}$ (started at 1) can be written in terms of a $\mathfrak{g}^{(2)}(\mathbb{R}^d)$-valued (standard) Lévy process (X, Z), started at 0, as

$$\mathbf{X}_t = \exp(X_t, \mathbb{A}_t + Z_t),$$

where $\mathbb{A}_t = \mathbb{A}_{0,t}$ is the stochastic area associated to X. Indeed, for $v, w \in J$, write $x = (x^v)$ for a generic element in $\mathfrak{g}^{(2)}$ and then

$$((a^{v,w}), (b^v), \mathfrak{K})$$

for the Lévy-triplet of (X, Z). Of course, X and Z are also [$\mathbb{R}^d$- and $\mathfrak{so}(d)$-valued] Lévy process with triplets

$$((a^{i,j}), (b^i), K) \quad \text{and} \quad ((a^{jk,lm}), (b^{jk}), \mathbb{K}),$$

respectively, where K and $\mathbb{K}$ are the image measures of $\mathfrak{K}$ under the obvious projection maps, onto $\mathbb{R}^d$ and $\mathfrak{so}(d)$, respectively. Define also the image measure under exp, that is, $\mathbf{K} = \exp_* \mathbb{K}$. It is then easy to see that $\mathbf{X}$ is a Lévy process in the sense of Hunt (cf. Section 1.2.8) with triplet $(\mathbf{a}, \mathbf{b}, \mathbf{K})$. Conversely, given $(\mathbf{a}, \mathbf{b}, \mathbf{K})$, one constructs a $\mathfrak{g}^{(2)}(\mathbb{R}^d)$-valued Lévy process (X, Z) with triplet $((a^{v,w}), (b^v), \log_* \mathbf{K})$ and easily checks that the $\exp(X, \mathbb{A} + Z)$ is the desired $G^{(2)}$-valued Lévy process. □

Recall that the definition of the Carnot–Caratheodory (CC) norm on $G^{(2)}(\mathbb{R}^d)$ from Section 1.2.5. The definition below should be compared with the classical definition of the Blumenthal–Getoor (BG) index.

DEFINITION 48. Given a Lévy measure $\mathbf{K}$ on the Lie group $G^{(2)}(\mathbb{R}^d)$, call

$$\beta := \inf\left\{q > 0 : \int_{G^{(2)}(\mathbb{R}^d)} (\|g\|_{\mathrm{CC}}^q \wedge 1)\mathbf{K}(dg) < \infty\right\}$$

the Carnot–Caratheodory Blumenthal–Getoor (CCBG) index.

Unlike the classical BG index, the CCBG index is not restricted to $[0, 2]$.

LEMMA 49. *The CCBG index takes values in* $[0, 4]$.

PROOF. Set $\log(g) = x + a \in \mathbb{R}^d \oplus \mathfrak{so}(d)$. Then

$$\|g\|_{\mathrm{CC}}^q \asymp \sum_i |x^i|^q + \sum_{j<k} |a^{jk}|^{q/2}.$$

By the very nature of $\mathbf{K}$, it integrates $|x^i|^2$ and $|a^{jk}|^2$, and hence $\beta \le 4$. [The definition of CC Blumenthal–Getoor extends immediately to $G^{(N)}(\mathbb{R}^d)$, in which case $\beta \le 2N$.] $\square$

THEOREM 50. *Consider a* $G^{(2)}(\mathbb{R}^d)$*-valued Lévy process* $\mathbf{X}$ *with enhanced triplet* $(\mathbf{a}, \mathbf{b}, \mathbf{K})$. *Assume*:

(i) *the sub-ellipticity condition*

$$a^{v,w} \equiv 0 \qquad \textit{unless } v, w \in I = \{i : 1 \le i \le d\};$$

(ii) *the following bound on the CCBG index*:

$$\beta < 3.$$

Let $p \in (2, 3)$. *Then a.s.* $\mathbf{X}$ *is a Lévy* p*-rough path if* $p > \beta$ *and this condition is sharp*.

PROOF. Set $\log(g) = x + a \in \mathbb{R}^d \oplus \mathfrak{so}(d)$. Then

$$\|g\|_{\mathrm{CC}}^{2\rho} \asymp \sum_i |x^i|^{2\rho} + \sum_{j<k} |a^{jk}|^{\rho}.$$

Let K denote the image measure of $\mathbf{K}$ under the projection map $g \mapsto x \in \mathbb{R}^d$. Let also $\mathbb{K}$ denote the image measure under the map $g \mapsto a \in \mathfrak{so}(d)$. Since $\mathbf{K}$ is a Lévy measure on $G^{(2)}(\mathbb{R}^d)$, we know that

$$\int_{\mathfrak{so}(d)} (|a|^\rho \wedge 1)\mathbb{K}(da) < \infty, \tag{3.1.2}$$

whenever $\beta < 2\rho < 3$. We now show that $\mathbf{X}$ enjoys p-variation. We have seen in the proof of Theorem 47 that any such Lévy process can be written as

$$\log \mathbf{X} = (X, \mathbb{A} + Z),$$

where X is a d-dimensional Lévy process with triplet

$$((a^{i,j}), (b^i), K)$$

with $\mathfrak{so}(d)$-valued area $\mathbb{A} = \mathbb{A}_{s,t}$ and a $\mathfrak{so}(d)$-valued Lévy process Z with triplet

$$(0, (b^{jk}), \mathbb{K}).$$

We know that $\mathbb{E}[|X_{s,t}|^2] \lesssim |t-s|$ and $\mathbb{E}[|A_{s,t}|^2] \lesssim |t-s|^2$ and so, for $|t-s| \le h$,

$$\mathbb{P}\big(|X_{s,t}| > a\big) \le \frac{h}{a^2},$$
$$\mathbb{P}\big(|A_{s,t}|^{1/2} > a\big) \le \frac{h}{a^2}.$$

On the other hand,

$$\mathbb{P}\big(|Z_{s,t}|^{1/2} > a\big) \le \frac{1}{a^{2\rho}}\mathbb{E}\big(|Z_{s,t}|^{\rho}\big) \sim \frac{h}{a^{2\rho}} \int_{\mathfrak{so}(d)} \big(|a|^{\rho} \wedge 1\big)\mathbb{K}(da)$$

and so

$$\mathbb{P}\big(\|\mathbf{X}_{s,t}\|_{\mathrm{CC}} > a\big) \lesssim \frac{h}{a^{2\rho\vee 2}}.$$

It then follows from Manstavicius' criterion (cf. Section 1.2.6) applied with $\beta = 1$, $\gamma = 2\rho \vee 2$, that $\mathbf{X}$ has indeed p-variation, for any $p > 2\rho \vee 2$, and by taking the infimum, for all $p > \beta \vee 2$.

It remains to be seen that the conditions are sharp. Indeed, if the sup-ellipticity condition is violated, say if $a^{v,w} \neq 0$ for some $v = jk$, say, this means (Brownian) diffusity (and hence finite 2^+- but not 2-variation) in direction $[e_j, e_k] \in \mathfrak{so}(d)$. As a consequence, $\mathbf{X}$ has 4^+-variation (but not 4-variation), in particular, it fails to have p-variation for some $p \in [2, 3)$. Similarly, if one considers an α-stable process in direction $[e_j, e_k]$, with well-known finite α^+- but not α-variation, we see that the condition $p > \beta$ cannot be weakened. □

3.1.3. *Expected signatures for Lévy rough paths*. Let us return to Theorem 41, where we computed, subject to suitable integrability assumptions of the Lévy measure, the expected signature of a Lévy process, lifted by means of "Marcus" iterated integrals. There we found that the expected signature over $[0, T]$ takes the Lévy–Kintchine form

$$\mathbb{E}[\mathbf{X}_{0,T}] = \exp\left\{T\left(b + \frac{a}{2} + \int_{\mathbb{R}^d} \big(\exp(y) - 1 - y\mathbb{I}_{|y|<1}\big)K(dy)\right)\right\}$$

for some symmetric, positive semidefinite matrix a, a vector b and a Lévy measure K, provided $K\mathbb{I}_{|y|\ge 1}$ has moments of all orders. In absence of a drift b and jumps, the formula degenerate to Fawcett's form, that is,

$$\exp\left(T\frac{a}{2}\right)$$

for a *symmetric* 2-tensor a. Let us present two examples of Lévy rough paths, for which the expected signature is computable and *different* from the above form.

EXAMPLE 51. We return to the noncanonical Brownian rough path $\mathbf{B}^{\mathrm{m}}$, the zero-mass limit of physcial Brownian motion in a magnetic field, as discussed in Example 15. The signature $S = S^{\mathrm{m}}$ is then given by Lyons' extension theorem applied to $\mathbf{B}^{\mathrm{m}}$, or equivalently, by solving the following rough differential equation:

$$dS_t = S_t \otimes d\mathbf{B}_t^{\mathrm{m}}(\omega), \qquad S_0 = 1.$$

In [11], it was noted that the expected signature takes the Fawcett form,

$$\mathbb{E}[S_{0,T}^{\mathrm{m}}] = \exp\left\{T \frac{\tilde{a}}{2}\right\},$$

but now for a not necessarily symmetric 2-tensor $\tilde{a}$, the antisymmetric part of which depends on the charge of the particle and the strength of the magnetic field.

EXAMPLE 52. Consider the *pure area Poisson process* from Example 15. Fix some $\mathfrak{a} \in \mathfrak{so}(d)$ and let (N_t) be standard Poisson process, and rate $\lambda > 0$. We set

$$\mathbf{X}_t := \bigotimes_{i=1}^{N_t} \exp^{(2)}(\mathfrak{a}) \in G^{(2)}(\mathbb{R}^d);$$

noting that the underlying path is trivial, $X = \pi_1(\mathbf{X}) \equiv 0$ and clearly $\mathbf{X}$ is a non-Marcus–Lévy p-rough path, for any $p \geq 2$. The signature of $\mathbf{X}$ is by definition the minimal jump extension of $\mathbf{X}$ as provided by Theorem 20. We leave it as an easy exercise to the reader to see that the signature S is given by

$$S_t = \bigotimes_{i=1}^{N_t} \exp(\mathfrak{a}) \in T((\mathbb{R}^d)).$$

With due attention to the fact that computations take place in the (noncommutative) tensor algebra, we then compute explicitly

$$\begin{aligned}
\mathbb{E}S_T &= \sum_{k\geq 0} e^{\mathfrak{a}k} e^{-\lambda T} (\lambda T)^k / k! \\
&= e^{-\lambda T} \sum_{k\geq 0} (\lambda T e^{\mathfrak{a}})^k / k! \\
&= \exp[\lambda T (e^{\mathfrak{a}} - 1)].
\end{aligned}$$

Note that the jump is not described by a Lévy-measure on $\mathbb{R}^d$ but rather by a Dirac measure on $G^{(2)}$, assigning unit mass to $\exp \mathfrak{a} \in G^{(2)}$.

We now give a general result that covers all these examples. Indeed, Example 51 is precisely the case of $\tilde{a} = a + 2\mathfrak{b}$ with antisymmetric $\mathfrak{b} = (b^{j,k}) \neq 0$, and symmetric $a = (a^{i,j})$. As for example (ii), everything is trivial but $\mathbf{K}$, which assigns unit mass to the element $\exp \mathfrak{a}$.

THEOREM 53. *Consider a Lévy rough path* $\mathbf{X}$ *with enhanced triplet* $(\mathbf{a}, \mathbf{b}, \mathbf{K})$. *Assume that* $\mathbf{K}1_{\{|g|>1\}}$ *integrates all powers of* $|g| := |\log g|_{\mathbb{R}^d \oplus \mathfrak{so}(d)}$. *Then the signature of* $\mathbf{X}$, *by definition, the minimal jump extension of* $\mathbf{X}$ *as provided by Theorem* 20, *is given by*

$$\begin{aligned}\mathbb{E}S_{0,T} = \exp\Biggl[T\Biggl(\frac{1}{2}\sum_{i,j=1}^{d} a^{i,j} e_i \otimes e_j + \sum_{i=1}^{d} b^i e_i + \sum_{j<k} b^{j,k}[e_j, e_k] \\ + \int_{G^{(2)}} \{\exp(\log_{(2)} g) - g 1_{\{|g|<1\}}\}\mathbf{K}(dg)\Biggr)\Biggr]. \qquad (3.1.3)\end{aligned}$$

PROOF. We saw in Corollary 39 that S solves

$$S_t = 1 + \int_0^t S_{s-} \otimes d\mathbf{X}_s + \sum_{0<s\le t} S_{s-} \otimes \{\exp(\log^{(2)} \Delta\mathbf{X}_s) - \Delta\mathbf{X}_s\}.$$

With notation as in the proof of Theorem 50,

$$\mathbb{X}_{s,t} = \pi_2 \exp(X_{s,t} + \mathbb{A}_{s,t} + Z_t - Z_s) = \frac{1}{2} X_{s,t} \otimes X_{s,t} + \mathbb{A}_{s,t} + Z_{s,t},$$

$$\mathbb{X}^{\mathrm{I}}_{s,t} = \frac{1}{2}(X_{s,t} \otimes X_{s,t} - [X, X]_{s,t}) + Z_{s,t},$$

where we recall that (X, Z) is a $\mathbb{R}^d \oplus \mathfrak{so}(d)$ valued Lévy process. With $Z_{s,t} = Z_t - Z_s$, we note additivity of $\Xi := \mathbb{X} - \mathbb{X}^{\mathrm{I}}$ given by

$$\Xi_{s,t} := \frac{1}{2}[X, X]_{s,t} + Z_{s,t} = \frac{1}{2} a(t-s) + \frac{1}{2} \sum_{r\in(s,t]} |\Delta X_r|^{\otimes 2} + Z_{s,t}.$$

But then

$$\int_0^t S_{s-} \otimes d\mathbf{X}_s = \int_0^t S_{s-} \otimes d\mathbf{X}^{\mathrm{I}}_s + \int_0^t S_{s-} \otimes d\Xi$$

and so, thanks to Theorem 40 on consistency of Itô with rough integration, we can express S as solution to a proper Itô integral equation,

$$\begin{aligned}S_t &= 1 + \int_0^t S_{s-} \otimes dX_s + \int_0^t S_{s-} \otimes d\Xi + \sum_{0<s\le t} S_{s-} \otimes \{\exp(\log^{(2)} \Delta\mathbf{X}_s) - \Delta\mathbf{X}_s\} \\ &\equiv 1 + (1) + (2) + (3).\end{aligned}$$

Let M^X be the martingale part in the Itô–Lévy decomposition of X, and write also $N^{\mathbb{K}}$ for the Poisson random measure with intensity $ds\mathbb{K}(dy)$. Then, with $\mathfrak{b} \equiv \sum_{j<k} b^{jk}[e_j, e_k]$,

$$X_t = M^X_t + bt + \int_{(0,t]\times\{|y|+|a|\ge1\}} yN^{\mathbb{K}}(ds, d(y, \mathfrak{a})) \in \mathbb{R}^d,$$

$$Z_t = M_t^Z + \mathfrak{b}t + \int_{(0,t]\times\{|y|+|a|\geq 1\}} \mathfrak{a} N^{\mathbb{K}}(ds, d(y,\mathfrak{a})) \in \mathfrak{so}(d),$$
$$\Xi_t = \frac{1}{2}at + \frac{1}{2}\int_{(0,t]\times\{|y|\geq 1\}} y^{\otimes 2} N^{\mathbb{K}}(ds, d(y,\mathfrak{a})) + Z_t \in (\mathbb{R}^d)^{\otimes 2}.$$

Check (inductively) integrability of S_t and note that $\int S_{s-}\, dM_s$ has zero mean, for either martingale choice. It follows that

$$\Phi_t = 1 + \int_0^t \Phi_s \otimes (C_1 + C_2 + C_3)\, ds,$$

where

$$C_1 = b + \int_{g^2(\mathbb{R}^d)} y 1_{\{|y|+|a|>1\}} \mathbb{K}(y,\mathfrak{a}),$$
$$C_2 = \frac{1}{2}a + \frac{1}{2}\int_{g^2(\mathbb{R}^d)} y^{\otimes 2} 1_{\{|y|+|a|>1\}} \mathbb{K}(y,\mathfrak{a}) + \mathfrak{b} + \int_{g^2(\mathbb{R}^d)} \mathfrak{a} 1_{\{|y|+|a|>1\}} \mathbb{K}(y,\mathfrak{a}),$$
$$C_3 = \int_{G^{(2)}(\mathbb{R}^d)} \{\exp(\log^{(2)} g) - g\} \mathbf{K}(dg).$$

Recall $\mathbb{K} = \log_*^{(2)} \mathbf{K}$ so that the sum of the three integrals over $g^2(\mathbb{R}^d)$ is exactly

$$\int_{G^{(2)}} g 1_{\{|g|\geq 1\}} \mathbf{K}(dg),$$

where $|g| = |\log g| = |y| + |\mathfrak{a}|$. And it follows that

$$C_1 + C_2 + C_3 = \frac{1}{2}a + b + \mathfrak{b} + \int_{G^{(2)}(\mathbb{R}^d)} \{\exp(\log_{(2)} g) - g 1_{\{|g|<1\}}\} \mathbf{K}(dg)$$

which completes our proof. □

3.1.4. *The moment problem for random signatures.* Any Lévy rough path $\mathbf{X}(\omega)$ over some fixed time horizon $[0,T]$ determines, via the minimal jump extension theorem, a random group-like element, say $S_{0,T}(\omega) \in T((\mathbb{R}^d))$. What information does the expected signature really carry? This was first investigated by Fawcett [9], and more recently by Chevyrev–Lyons [7]. Using a criterion from [7], we can show the following.

THEOREM 54. *The law of $S_{0,T}(\omega)$ is uniquely determined from its expected signature whenever*

$$\forall \lambda > 0: \qquad \int_{y\in G^{(2)}:|y|>1} \exp(\lambda|y|) \mathbf{K}(dy) < \infty.$$

PROOF. As in [7], we need to show that $\exp(C)$, equivalently $C = (C^0, C^1, C^2, \dots) \in T((\mathbb{R}^d))$, has sufficiently fast decay as the tensor levels grow. In particular, only the jumps matter. More precisely, by a criterion put forward in [7] we need to show that

$$\sum \lambda^m C^m < \infty,$$

where (for $m \geq 3$),

$$C^m = \pi_m \left(\int_{G^{(2)}} (e_{(n)}^{\log_{(2)} g} - g) \mathbf{K}(dg) \right) \in (\mathbb{R}^d)^{\otimes m}.$$

We leave it as elementary exercise to see that this is implied by the exponential moment condition on $\mathbf{K}$. □

3.2. *Further classes of stochastic processes.*

3.2.1. *Markov jump diffusions.* Consider a d-dimensional strong Markov with generator

$$\begin{aligned}(\mathcal{L}f)(x) = \frac{1}{2} \sum_{i,j \in I} a^{i,j}(x) \partial_i \partial_j f + \sum_{i \in I} b^i(x) \partial_i f \\ + \int_{\mathbb{R}^d} \left\{ f(x+y) - f(x) - 1_{\{y \leq 1\}} \sum_{i \in I} y^i \partial_i f \right\} K(x, dy).\end{aligned}$$

Throughout, assume $a = \sigma\sigma^T$ and σ, b bounded Lipschitz, $K(x, \cdot)$ a Lévy measure, with uniformly integrable tails. Such a process can be constructed as jump diffusion [19]; the martingale problem is discussed in Stroock [45]. As was seen, even in the Lévy case, with a (constant) Lévy triplet (a, b, K), showing finite p-variation in rough path sense is nontrivial, the difficulty of course being the stochastic area

$$A_{s,t}(\omega) = \text{Anti} \int_{(s,t]} (X_r^- - X_s) \otimes dX \in \mathfrak{so}(d);$$

where stochastic integration is understood in the Itô sense. In this section, we will prove the following.

THEOREM 55. *With probability one, $X(\omega)$ lifts to a $G^{(2)}$-valued path, with increments given by*

$$\mathbf{X}_{s,t} := \exp^{(2)}(X_{s,t} + A_{s,t}) = \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$$

and $\mathbf{X}$ is a càdlàg Marcus-like, geometric p-rough path, for any $p > 2$.

Note the immediate consequences of this theorem: the minimal jump extension of the geometric rough $(X, \mathbb{X}^{\mathrm{M}})$ can be identified with the Marcus lift, stochastic integrals and differential equations driven by X can be understood deterministically as function of $\mathbf{X}(\omega)$ and are identified with corresponding rough integrals and canonical equations. As in the Lévy case discussed earlier, we base the proof on the expected signature and point out some Markovian aspects of independent interest. Namely, we exhibit the step-N Marcus lift as $G^{(N)}$-valued Markov process and compute its generator. To this end, recall (e.g., [15], Remark 7.43) the generating vector fields $U_i(g) = g \otimes e_i$ on $G^{(N)}$, with the property that

$$\mathrm{Lie}(U_1, \dots, U_d)|_g = \mathcal{T}_g G^{(N)}.$$

PROPOSITION 56. *Consider a d-dimensional Markov process (X) with generator as above and the Marcus canonical equation $dS = S \otimes \diamond dX$, started from*

$$1 \equiv (1, 0, \dots, 0) \in G^{(N)}(\mathbb{R}^d) \subset T^{(N)}(\mathbb{R}^d).$$

Then S takes values in $G^{(N)}(\mathbb{R}^d)$ and is Markov with the generator, for $f \in C_c^2$,

$$\begin{aligned}
(\mathcal{L}f)(x) &= (\mathcal{L}^{(N)} f)(x) \\
&= \frac{1}{2} \sum_{i,j \in I} a^{i,j}(\pi_1(x)) U_i U_j f + \sum_{i \in I} b^i(\pi_1(x)) U_i f \\
&\quad + \int_{\mathbb{R}^d} \Big\{ f(x \otimes Y) - f(x) - 1_{\{y \le 1\}} \sum_{i \in I} y^i U_i f \Big\} K(x, dy) \\
&\quad \textit{with } Y \equiv \exp^{(n)}(y).
\end{aligned}$$

PROOF (Sketch). Similar to the proof of Theorem 41. Write $X = M + V$ for the semimartingale decomposition of X. We have

$$dS = S \otimes \diamond dX = \sum_{i \in I} U_i(S) \diamond dX^i$$

and easily deduce an evolution equation for $f(S_t) = f(1)$. Taking the expected value leads to the form $(\mathcal{L}f)$. □

Since N was arbitrary, this leads to the expected signature. We note that in the (Lévy) case of x-independent characteristics, Φ does not depend on x in which case the PIDE reduces to the ODE $\partial_t \Phi = C \otimes \Phi$, which leads to the Lévy–Kinthchine form $\Phi(t) = \exp(Ct)$ obtained previously. We also note that the solution $\Phi = (1, \Phi^1, \Phi^2, \dots)$ to the PIDE system given in the next theorem can be iteratively constructed. In the absence of jumps, this system reduces to a system of PDEs derived by Ni Hao [27, 37].

THEOREM 57 (PIDE for expected signature). *Assume uniformly bounded jumps, σ, b bounded and Lipschitz, $a = \sigma\sigma^T$, the expected signature $\Phi(x,t) = E^x S_{0,t}$ exists. Set*

$$
\begin{aligned}
C(x) &:= \sum_{i\in I} b^i(x)e_i + \frac{1}{2}\sum_{i,j\in I} a^{i,j}(x)e_i \otimes e_j \\
&\quad + \int_{\mathbb{R}^d}\Bigl(Y - 1 - \mathbb{I}_{\{y\le 1\}}\sum_{i\in I} y^i e_i\Bigr)K(x,dy) \\
&\quad \text{with } Y = \exp(y) \in T\bigl((\mathbb{R}^d)\bigr).
\end{aligned}
$$

Then $\Phi(x,t)$ solves

$$
\begin{cases}
\partial_t \Phi = C \otimes \Phi + \mathcal{L}\Phi + \sum_{i,j\in I} a^{i,j}(\partial_j \Phi)(x)e_i \\
\qquad\qquad + \int_{\mathbb{R}^d}(Y-1)\otimes\bigl(\Phi(x\otimes Y) - \Phi(x)\bigr)K(x,dy), \\
\Phi(x,0) = 1.
\end{cases}
$$

PROOF. It is enough to establish this in $T^{(N)}(\mathbb{R}^d)$, for arbitrary integer N. We can see that

$$\mathbb{E}^x \mathbf{X}_t^{(N)} =: u(x,t),$$

for $x \in G^{(N)}(\mathbb{R}^d) \subset T^{(N)}(\mathbb{R}^d)$ is well defined, in view of the boundedness assumptions made on the coefficients, and then a (vector-valued, unique linear growth) solution to the backward equation

$$
\begin{aligned}
\partial_t u &= \mathcal{L}u, \\
u(x,0) &= x \in T^{(N)}\bigl(\mathbb{R}^d\bigr).
\end{aligned}
$$

It is then clear that

$$E^x \mathbf{X}_{0,t}^{(N)} = x^{-1} \otimes u(x,t) =: \Phi(x,t)$$

also satisfies a PDE. Indeed, noting the product rule for second-order partial-integro operators,

$$
\begin{aligned}
\bigl(\mathcal{L}[fg]\bigr)(x) &= \bigl((\mathcal{L}[f])g\bigr)(x) + \bigl(f\mathcal{L}[g]\bigr)(x) + \Gamma(f,g), \\
\Gamma(f,g) &= \sum_{i,j\in I} a^{i,j}(U_i f U_j g)(x) \\
&\quad + \int_{G^{(2)}} \bigl(f(x\otimes Y) - f(x)\bigr)\bigl(g(x\otimes Y) - g(x)\bigr)\mathbf{K}(dy)
\end{aligned}
$$

and also noting the action of U_v on $f(x) \equiv x$, namely $U_i f = x \otimes e_v$, we have

$$\begin{aligned}\mathcal{L}x &= x \otimes C \\ &:= x \otimes \Big\{ \sum_{v \in J} b^v \otimes e_v + \frac{1}{2} \sum_{i,j \in I} a^{i,j} e_i \otimes e_j \\ &\quad + \int_{G^{(2)}} \Big(Y - 1 - \mathbb{I}_{\{y \le 1\}} \sum_{v \in J} Y^v \otimes e_v \Big) \mathbf{K}(dy) \Big\}, \\ \Gamma(x, g) &= x \otimes \Big\{ \sum_{i,j \in I} a^{i,j} (U_j g)(x) e_i + \int_{G^{(2)}} (Y - 1)\big(g(x \otimes Y) - g(x)\big)\mathbf{K}(dy) \Big\}.\end{aligned}$$

As a consequence,

$$x \otimes \partial_t \Phi = \partial_t u = \mathcal{L}u = \mathcal{L}(x \otimes \Phi) = (\mathcal{L}x) \otimes \Phi + x \otimes \mathcal{L}[\Phi] + \Gamma(x, \Phi)$$

and hence

$$\begin{aligned}\partial_t \Phi = C \otimes \Phi + \Big\{ &\mathcal{L}[\Phi] + \sum_{i,j \in I} a^{i,j} (U_j \Phi)(x) e_i \\ &+ \int_{G^{(2)}} (Y - 1)\big(\Phi(x \otimes Y) - \Phi(x)\big)\mathbf{K}(dy) \Big\}.\end{aligned} \tag{3.2.1}$$

□

We can now show rough path regularity for general jump diffusions.

PROOF OF THEOREM 55. Only the p-variation statement requires a proof. The key remark is that the above PIDE implies

$$\Phi_t = 1 + (\partial_t|_{t=0}\phi)t + O(t^2) = 1 + Ct + O(t^2),$$

where our assumptions on a, b, K guarantee uniformity of the O-term in x. We can then argue exactly as in the proof of Corollary 42. □

3.2.2. *Semimartingales*. In [24], Lépingle established finite p-variation of general semimartingales, any $p > 2$, together with a powerful Burkholder–Davis–Gundy-type (BDG) estimates. For *continuous* semimartingales, the extension to the (Stratonovich–Marcus) rough path lift was obtained in [13] (see also [15], Chapter 14), but so far the general (discontinuous) case eluded us.[6] (By Proposition 16, it does not matter if one establishes finite p-variation in the rough path sense for the Itô or Marcus lift.)

As it is easy to explain, let us just point to the difficulty in extending Lépingle's result in the first place: he crucially relies on Monroe's result [36], stating

[6] UPDATE: Rough path p-variation of lifted general semimartingales, together with a BDG inequality, is established in forthcoming work by I. Chevyrev and the first named author.

that every (scalar) càdlàg semimartingale can be written as a time-changed scalar Brownian motion for a (càdlàg) family of stopping times (on a suitably extended probability space). This, however, fails to hold true in higher dimensions and not every (Marcus or Itô) lifted general semimartingale will be a (càdlàg) time-change of some enhanced Brownian motion ([15], Chapter 13) in which case the finite p-variation would be an immediate consequence of known facts about the enhanced Brownian motion (a.k.a. Brownian rough path) and invariance of p-variation under reparametrization.

A large class of general semimartingales for which finite p-variation (in the rough path sense, any $p > 2$) can easily be seen, consists of those with summable jumps. Following Kurtz et al. ([22], page 368), such a semimartingale, that is, with a jump replaced by straight lines over stretched time, may be interpreted as *a continuous semimartingale*. One can then apply [13, 15] and again appeal to invariance of p-variation under reparametrization to see that such (enhanced) semimartingales have a.s. p-rough sample paths, for any $p > 2$.

Another class of general semimartingales for which finite p-variation can easily be seen, consists of time-changed Lévy processes (a popular class of processes used in mathematical finance). Indeed, appealing once more to invariance of p-variation under reparametrization, the statement readily follows from the corresponding p-variation regularity of Lévy rough paths.

3.2.3. *Gaussian processes.* We start with a brief review of some aspects of the work of Jain and Monrad [20]. Given a (for the moment, scalar) zero-mean, separable Gaussian process on $[0, T]$, set $\sigma^2(s,t) = \mathbb{E}X_{s,t}^2 = |X_t - X_s|_{L^2}^2$ where $L^2 = L^2(\mathbb{P})$. We regard the process X as a Banach space valued path $[0, T] \to H = L^2(P)$ and assume finite 2ρ-variation, in the sense of Jain and Monrad's condition

$$F(T) := \sup_{\mathcal{P}} \sum_{[u,v]\in\mathcal{P}} \left|\sigma^2(u,v)\right|^\rho = \sup_{\mathcal{P}} \sum_{[u,v]\in\mathcal{P}} |X_t - X_s|_{L^2}^{2\rho} < \infty \tag{3.2.2}$$

with partitions $\mathcal{P}$ of $[0, T]$. It is elementary to see that p-variation paths can always be written as time-changed Hölder continuous paths with exponent $1/p$ (see, e.g., Lemma 4.3. in [8]). Applied to our setting, with $\alpha^* = 1/(2\rho)$, $\tilde{X} \in C^{\alpha^*\text{-Höl}}([0, F(T)], H)$ so that

$$\tilde{X} \circ F = X \in W^{2\rho}([0, T], H).$$

Now in view of the classical Kolmogorov criterion, and equivalence of moments for Gaussian random variables, knowing

$$|\tilde{X}_t - \tilde{X}_s|_{L^2} \le C|t - s|^{\alpha^*}$$

implies that $\tilde{X}$ (or a modification thereof) has a.s. α-Hölder samples paths, for any $\alpha < \alpha^*$. But then, trivially, $\tilde{X}$ has a.s. finite p-variation sample paths, for any $p >$

$1/\alpha = 2\rho$, and so does X by invariance of p-variation under reparametrization. (It should be noted that such X has only discontinuities at deterministic times, inherited from the jumps of F.) In a nutshell, this is one of the main results of Jain and Monrad [20], as summarized by Dudley–Norvaiša in [8], Theorem 5.3. We have the following extension to Gaussian rough paths.

THEOREM 58. *Consider a d-dimensional zero-mean, separable Gaussian process (X) with independent components. Let $\rho \in [1, 3/2)$ and assume*

$$\sup_{\mathcal{P},\mathcal{P}'} \sum_{\substack{[s,t]\in\mathcal{P} \\ [u,v]\in\mathcal{P}'}} \left|\mathbb{E}(X_{s,t} \otimes X_{u,v})\right|^{\rho} < \infty. \tag{3.2.3}$$

Then X has a càdlàg modification, denoted by the same letter, which lifts a.s. to a random geometric càdlàg rough path, with $\mathbb{A} = \mathrm{Anti}(\mathbb{X})$ given as L^2-limit of Riemann–Stieltjes approximations.

PROOF. In a setting of continuous Gaussian processes, condition (3.2.3), that is, finite ρ-variation of the covariance, is well known [14, 15]. It plainly implies the Jain–Monrad condition (3.2.2), for each component (X^i). With $F(t) := \sum_{i=1}^{d} F^i(t)$, we can then write

$$\tilde{X} \circ F = X$$

for some d-dimensional, zero mean (by Kolmogorov criterion: continuous) Gaussian process $\tilde{X}$, whose covariance also enjoys finite ρ-variation. We can now apply a standard (continuous) Gaussian rough path theory [14, 15] and construct a canonical geometric rough path lift of $\tilde{X}$; that is,

$$\tilde{\mathbf{X}} = (\tilde{X}, \tilde{\mathbb{X}}) \in \mathcal{C}^{\rho}$$

with probability 1. The desired geometric càdlàg rough path lift is then given by

$$(X, \mathbb{X}) = \mathbf{X} := \tilde{\mathbf{X}} \circ F.$$

The statement about L^2-convergence of Riemann–Stieltjes approximations follows immediately from the corresponding statements for $\mathrm{Anti}(\tilde{\mathbb{X}})$, as found in [14], Chapter 10.2. □

Acknowledgements. Both authors would like to thank the referees for their careful reading.

REFERENCES

[1] APPLEBAUM, D. (2004). *Lévy Processes and Stochastic Calculus*. *Cambridge Studies in Advanced Mathematics* **93**. Cambridge Univ. Press, Cambridge. MR2072890

[2] APPLEBAUM, D. and KUNITA, H. (1993). Lévy flows on manifolds and Lévy processes on Lie groups. *J. Math. Kyoto Univ.* **33** 1103–1123. MR1251218

[3] BAUDOIN, F. (2004). *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press, London. MR2154760

[4] BERTOIN, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge. MR1406564

[5] BOEDIHARDJO, H., GENG, X., LYONS, T. and YANG, D. (2016). The signature of a rough path: Uniqueness. *Adv. Math.* **293** 720–737. MR3474333

[6] CHEVYREV, I. (2015). Random walks and Lévy processes as rough paths. *Probab. Theory Relat. Fields*. To appear. Available at arXiv:1510.09066.

[7] CHEVYREV, I. and LYONS, T. (2016). Characteristic functions of measures on geometric rough paths. *Ann. Probab.* **44** 4049–4082. MR3572331

[8] DUDLEY, R. M. and NORVAIŠA, R. (1998). An Introduction to P-variation and Young Integrals: With Emphasis on Sample Functions of Stochastic Processes. Lecture Notes. MaPhySto, Dept. Mathematical Sciences, Univ. Aarhus.

[9] FAWCETT, T. A. (2002). Non-commutative harmonic analysis. Ph.D. thesis, Mathematical Institute, Univ. Oxford.

[10] FÖLLMER, H. (1981). Calcul d'Itô sans probabilités. In *Seminar on Probability*, *XV* (*Univ. Strasbourg*, *Strasbourg*, 1979/1980) (*French*). *Lecture Notes in Math*. **850** 143–150. Springer, Berlin. MR0622559

[11] FRIZ, P., GASSIAT, P. and LYONS, T. (2015). Physical Brownian motion in a magnetic field as a rough path. *Trans. Amer. Math. Soc.* **367** 7939–7955. MR3391905

[12] FRIZ, P. and VICTOIR, N. (2006). A variation embedding theorem and applications. *J. Funct. Anal.* **239** 631–637. MR2261341

[13] FRIZ, P. and VICTOIR, N. (2008). The Burkholder–Davis–Gundy inequality for enhanced martingales. In *Séminaire de Probabilités XLI*. *Lecture Notes in Math.* **1934** 421–438. Springer, Berlin. MR2483743

[14] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths*. *With an introduction to regularity structures*. Springer, Cham. MR3289027

[15] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669

[16] GUBINELLI, M. (2004). Controlling rough paths. *J. Funct. Anal.* **216** 86–140. MR2091358

[17] HAMBLY, B. and LYONS, T. (2010). Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math.* (2) **171** 109–167. MR2630037

[18] HUNT, G. A. (1956). Semi-groups of measures on Lie groups. *Trans. Amer. Math. Soc.* **81** 264–293. MR0079232

[19] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

[20] JAIN, N. C. and MONRAD, D. (1983). Gaussian measures in B_p. *Ann. Probab.* **11** 46–57. MR0682800

[21] KARANDIKAR, R. L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* **57** 11–18. MR1327950

[22] KURTZ, T. G., PARDOUX, É. and PROTTER, P. (1995). Stratonovich stochastic differential equations driven by general semimartingales. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 351–377. MR1324812

[23] LEDOUX, M., LYONS, T. and QIAN, Z. (2002). Lévy area of Wiener processes in Banach spaces. *Ann. Probab.* **30** 546–578. MR1905851

[24] LÉPINGLE, D. (1976). La variation d'ordre p des semi-martingales. *Z. Wahrsch. Verw. Gebiete* **36** 295–316. MR0420837

[25] LE JAN, Y. and QIAN, Z. (2013). Stratonovich's signatures of Brownian motion determine Brownian sample paths. *Probab. Theory Related Fields* **157** 209–223. MR3101845

[26] LIAO, M. (2004). *Lévy Processes in Lie Groups*. *Cambridge Tracts in Mathematics* **162**. Cambridge Univ. Press, Cambridge. MR2060091

[27] LYONS, T. and NI, H. (2015). Expected signature of Brownian motion up to the first exit time from a bounded domain. *Ann. Probab.* **43** 2729–2762. MR3395473

[28] LYONS, T. and QIAN, Z. (2002). *System Control and Rough Paths*. Oxford Univ. Press, Oxford. Oxford Science Publications. MR2036784

[29] LYONS, T. and VICTOIR, N. (2004). Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** 169–198. Stochastic analysis with applications to mathematical finance. MR2052260

[30] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. MR1654527

[31] LYONS, T. J., CARUANA, M. and LÉVY, T. (2007). *Differential Equations Driven by Rough Paths*. *Lecture Notes in Math.* **1908**. Springer, Berlin. MR2314753

[32] MANSTAVIČIUS, M. (2004). p-variation of strong Markov processes. *Ann. Probab.* **32** 2053–2066. MR2073185

[33] MARCUS, S. I. (1978). Modeling and analysis of stochastic differential equations driven by point processes. *IEEE Trans. Inform. Theory* **24** 164–172. MR0487784

[34] MARCUS, S. I. (1980/81). Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics* **4** 223–245. MR0605630

[35] MIKOSCH, T. and NORVAIŠA, R. (2000). Stochastic integral equations without probability. *Bernoulli* **6** 401–434. MR1762553

[36] MONROE, I. (1972). On embedding right continuous martingales in Brownian motion. *Ann. Math. Stat.* **43** 1293–1311. MR0343354

[37] NI, H. (2013). The Expected Signature of a Stochastic Process. Ph.D. thesis, Mathematical Institute, Univ. Oxford.

[38] PERKOWSKI, N. and PRÖMEL, D. J. (2016). Pathwise stochastic integrals for model free finance. *Bernoulli* **22** 2486–2520. MR3498035

[39] PRÖMEL, D. J. and TRABS, M. (2016). Rough differential equations driven by signals in Besov spaces. *J. Differential Equations* **260** 5202–5249. MR3448778

[40] PROTTER, P. E. (2005). *Stochastic Integration and Differential Equations*, 2nd ed., version 2.1, corrected third printing. *Stochastic Modelling and Applied Probability* **21**. Springer, Berlin. MR2273672

[41] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol.* 2. *Cambridge Mathematical Library*. Cambridge Univ. Press, Cambridge. Itô calculus, Reprint of the second (1994) edition. MR1780932

[42] SATO, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. Translated from the 1990 Japanese original, revised by the author. MR1739520

[43] SIMON, T. (2003). Small deviations in p-variation for multidimensional Lévy processes. *J. Math. Kyoto Univ.* **43** 523–565. MR2028666

[44] SONER, H. M., TOUZI, N. and ZHANG, J. (2011). Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.* **16** paper no. 67, 1844–1879. MR2842089

[45] STROOCK, D. W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrsch. Verw. Gebiete* **32** 209–244. MR0433614

[46] WILLIAMS, D. R. E. (2001). Path-wise solutions of stochastic differential equations driven by Lévy processes. *Rev. Mat. Iberoam.* **17** 295–329. MR1891200

[47] YOUNG, L. C. (1936). An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67** 251–282. MR1555421

Institut für Mathematik
Technische Universität Berlin
Strasse des 17. Juni 136
10623 Berlin
Germany
and
Weierstrass-Institut für Angewandte
Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
Germany
E-mail: friz@math.tu-berlin.de

Indian Statistical Institute
8th Mile, Mysore Road
Bangalore 560 059
Bangalore
India