Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Branched Itô formula and natural Itô–Stratonovich isomorphism

Carlo Bellingeri[1], Emilio Ferrucci[2], Nikolas Tapia[3]

submitted: January 9, 2024

[1] Technische Universität Berlin
Str. des 17. Juni 136
10587 Berlin
Germany
E-Mail: bellinge@math.tu-berlin.de

[2] Mathematical Institute
University of Oxford
Andrew Wiles Building
Radcliffe Observatory Quarter (550)
Woodstock Road
Oxford
OX2 6GG
United Kingdom
E-Mail: emilio.rossiferrucci@maths.ox.ac.uk

[3] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: nikolasesteban.tapiamunoz@wias-berlin.de

No. 3083
Berlin 2024

2020 *Mathematics Subject Classification.* 60L20, 60L70, 16T30.

Key words and phrases. Itô formula, Itô-Stratonovich correction, rough differential equations, shuffle algebra, branched rough paths.

The first and third named authors are supported in part by the DFG Research Unit FOR2402. The second named author is supported by UKRI EPSRC Programme Grant EP/S026347/1. The third named author is also funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689). The authors would like thank Frederic Patras, Loïc Foissy and Kurusch Ebrahimi-Fard for fruitful conversations which took place at the conference "Structural Aspects of Signatures and Rough Paths" held at the Centre for Advanced Study (CAS) of Oslo, where a preliminary version of these results was presented. They would also like to thank the Mathematische Forschungsinstitut Oberwolfrach for their warm hospitality during the final stage of the writing of this manuscript.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Branched Itô formula and natural Itô–Stratonovich isomorphism

Carlo Bellingeri, Emilio Ferrucci, Nikolas Tapia

ABSTRACT. Branched rough paths define integration theories that may fail to satisfy the usual integration by parts identity. The intrinsically-defined projection of the Connes-Kreimer Hopf algebra onto its primitive elements defined by Broadhurst and Kreimer, and further studied by Foissy, allows us to view it as a commutative $\mathbf{B}_\infty$-algebra and thus to write an explicit change-of-variable formula for solutions to rough differential equations. This formula, which is realised by means of an explicit morphism from the Grossman-Larson Hopf algebra to the Hopf algebra of differential operators, restricts to the well-known Itô formula for semimartingales. We establish an isomorphism with the shuffle algebra over primitives, extending Hoffman's exponential for the quasi shuffle algebra, and in particular the usual Itô-Stratonovich correction formula for semimartingales. We place special emphasis on the one-dimensional case, in which certain rough path terms can be expressed as polynomials in the extended trace path, which for semimartingales restrict to the well-known Kailath-Segall polynomials. We end by describing an algebraic framework for generating examples of branched rough paths, and, motivated by the recent literature on stochastic processes, exhibit a few such examples above scalar $1/4$-fractional Brownian motion, two of which are "truly branched", i.e. not quasi-geometric.

CONTENTS

INTRODUCTION

One of the cornerstones of stochastic analysis is Itô's change-of-variable formula [30]. Given a continuous semimartingale Y and a smooth function φ, it tells us how to express $\varphi(Y)$ in terms of Itô integration against Y and Riemann-Stieltjes integration against the quadratic variation path $[Y]$:

$$\varphi(Y_t) - \varphi(Y_s) = \int_s^t \mathrm{D}\varphi(Y_u)\,\mathrm{d}Y_u + \frac{1}{2}\int_s^t \mathrm{D}^2\varphi(Y_u)\,\mathrm{d}[Y]_u \tag{0.1}$$

This result is arguably what elevates the status of Itô's theory to that of a "calculus", albeit one that does not satisfy the same identities as ordinary calculus, as exemplified by the above identity. The Itô formula provides the means to carry out many computations of interest in probability theory, such as those involving (conditional) expectations. For example, if Y satisfies the stochastic differential equation (SDE)

$$\mathrm{d}Y_t = \sigma(Y_t, t)\,\mathrm{d}W_t + \mu(Y_t, t)\,\mathrm{d}t \tag{0.2}$$

DOI 10.20347/WIAS.PREPRINT.3083 Berlin 2024

where W is a (multidimensional) Brownian motion—roughly speaking, this means $Y_{t+\mathrm{d}t} - Y_t$ is normally distributed with mean $\mu(Y_t, t)$ and variance $\sigma(Y_t, t)\sigma(Y_t, t)^\top$ at each time t—the Itô formula can be written as

$$\begin{aligned} \varphi(Y_t) - \varphi(Y_s) &= \int_s^t \mathrm{D}\varphi(Y_u)\sigma(Y_u, u)\,\mathrm{d}W_u + \int_s^t \mathcal{L}\varphi(Y_u, u)\,\mathrm{d}u \\ \text{with} \quad \mathcal{L}\varphi(x,u) &\coloneqq \mu(x,u) + \frac{1}{2}\sigma(x,u)\sigma(x,u)^\top\,, \end{aligned} \tag{0.3}$$

implying that, by the martingale property,

$$\mathbb{E}_s[\varphi(Y_t) - \varphi(Y_s)] = \int_s^t \mathbb{E}[\mathcal{L}\varphi(Y_u, u)]\,\mathrm{d}u$$

where $\mathbb{E}_s$ denotes expectation conditional on the information available at time s.

While the Itô integral is often preferred in probability and finance, thanks to its martingale-preserving properties, the Stratonovich integral is often preferred in physics and geometry, since it does satisfy the same identities as ordinary calculus (e.g. the Itô formula but without the second-order correction term). The two integrals are respectively defined by left endpoint and midpoint (or trapezoidal) Riemann-Stieltjes approximations:

$$\int H\mathrm{d}Y \eqqcolon L^2\lim \sum_{[s,t]} H_s Y_{s,t}, \qquad \int H \circ \mathrm{d}Y \eqqcolon L^2\lim \sum_{[s,t]} H_{\frac{s+t}{2}} Y_{s,t}$$

where $Y_{s,t} \coloneqq Y_t - Y_s$, H is an adapted and continuous integrand, and the limits are taken in L^2 along a partition with vanishing mesh size.

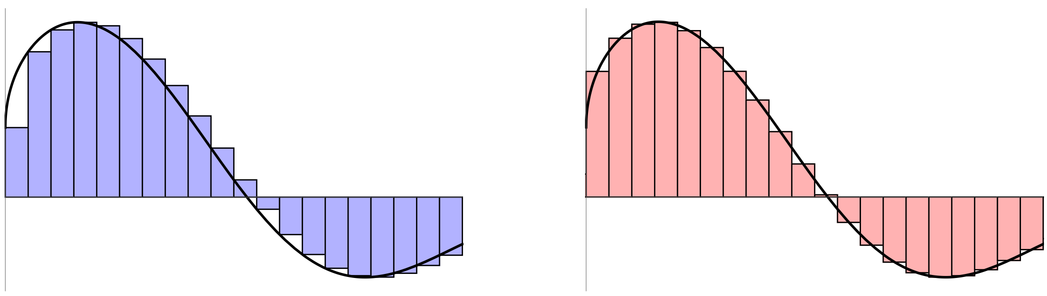


FIGURE 1. Visual representation for Riemann-Stieltjes sum approximations to Itô (left) and Stratonovich (right) integrals.

Despite the Itô and Stratonovich integrals being different, they are related by the identity $\int H \circ \mathrm{d}Y = \int H\mathrm{d}Y + \frac{1}{2}[H, Y]$, the bracket denoting the quadratic covariation between H and Y. In particular, this means that if Y satisfies the SDE (0.2), it equivalently satisfies the Stratonovich SDE

$$\mathrm{d}Y_t = \sigma(Y_t, t) \circ \mathrm{d}W_t + \left[\mu(Y_t, t) - \frac{1}{2}\mathrm{D}\sigma(Y_t, t) \cdot \sigma(Y_t, t)\right]\mathrm{d}t \tag{0.4}$$

This relates two a priori-distinct notions of stochastic differential equation, each with different advantages and drawbacks, at the sole cost of modifying the coefficients.

While Itô and Stratonovich SDEs constitute the main examples of random dynamical systems perturbed by instantaneous noise, they are by no means the only ones that can be conceived. A much more versatile notion of *controlled differential equation* involves considering a driving process $X \in C([0,T], V)$ (with V a finite-dimensional vector space) and an equation

$$\mathrm{d}Y_t = F(Y_t)\,\mathrm{d}X_t \tag{0.5}$$

with $F \in C^\infty(\mathcal{L}(V, W))$, $\mathcal{L}$ denoting linear maps. When X is smooth, of course, on may substitute $\mathrm{d}X_t = \dot{X}\mathrm{d}t$ above, but otherwise the equation must be understood differently. When X is of bounded ρ-variation with $\rho < 2$, such equation was first considered by Lyons [35], using Young's notion of integral [59]. When X is only of bounded $2 < \rho$-variation, however, Lyons's fundamental insight [38] is that the path X must be augmented by additional information to give the equation unique, well-posed and robust meaning. Initially, this superstructure—a *rough path*— involved postulating iterated integrals $\int_{s<u_1<\ldots<u_n<t} \mathrm{d}X_{u_1} \otimes \cdots \otimes \mathrm{d}X_{u_n}$ for all $n \le \lfloor \rho \rfloor$: this is usually done, when X is random, via probabilistic notions of convergence, subject to certain constraints, including a first-order integration by parts identity. Rough path theory is able to handle a wide variety of signals, including the special case of Stratonovich SDEs.

Lyons's theory was extended by Gubinelli in [25], who relaxed the algebraic requirements on the rough path, making it possible for the resulting integral not to satisfy the same identities as ordinary calculus. This involved requiring that the rough path not only contain the information representing linear iterated integrals, but also integrals of products (of integrals of products…and so on). Such an object is readily viewed as a functional on the Connes–Kreimer Hopf algebra $\mathcal{H}_{\mathsf{CK}}$, introduced at the end of the 20th century in the context of renormalization in Quantum Field Theory. Unterberger [57] also saw this connection independently around the same time. *Branched rough paths* are general enough to reproduce Itô's theory of integration, and arguably provide the most general framework for what a theory of path integration might look like. What was initially missing from Gubinelli's theory was a change-of-variable formula akin to (0.1), and an explanation of the relationship between branched and the previously-defined geometric rough paths. These two questions were addressed in Kelly's PhD thesis [31], and the latter jointly with Hairer [26]: for the former, an Itô-type formula is possible, at the cost of enlarging the original branched rough path; for the latter, every equation driven by a branched rough path can be expressed in terms of a geometric one, again, at the cost of enlarging (in a different way) the original path and choosing a geometric rough path above it. These two results, while of great significance, are vulnerable to the objection that the objects they consider—the extension of the branched rough path needed for the Itô formula and the associated geometric rough path containing the same information—are chosen in a highly non-canonical fashion. In contrast, the correction needed for the usual Itô formula, namely the quadratic variation $[Y]$, is defined intrinsically given the path Y (in fact even algebraically, as the defect in the integration by parts formula: $[Y]^{\alpha\beta} = Y^\alpha Y^\beta - \int Y^\alpha \mathrm{d}Y^\beta - \int Y^\beta \mathrm{d}Y^\alpha$). This was addressed, in relation to the second point above, by Boedihardjo and Chevyrev: in [4] they utilise results from Foissy [17] and Chapoton [9] to observe that the mapping from branched to geometric rough paths can be achieved in an entirely algebraic fashion, thus avoiding the troublesome choice of the rough path lift. The isomorphism of which they show the existence is still, however, not explicitly identified and computed, and many (indeed infinitely many) such choices exist, each corresponding to a free algebraic basis of the Gorssman–Larson Hopf algebra $\mathcal{H}_{\mathsf{GL}}$.

Our contributions are as follows. In Section 1 we view the decorated versions of the Connes-Kreimer and Grossman-Larson Hopf algebras as functors $\underline{\mathrm{Vec}} \to \underline{\mathrm{Hopf}}$. We use Foissy's result that $\mathcal{H}_{\mathsf{CK}}(U)$ (from now on $\mathcal{H}_{\mathsf{CK}}$, omitting the vector space U of decorations, argument of the functor) is cofree to think of $\mathcal{H}_{\mathsf{CK}}$ as a $\mathbf{B}_\infty$-algebras [19], and revisit Hoffman's duality result with $\mathcal{H}_{\mathsf{GL}}$ in this framework. A thorough review of the Oudom-Guin result about pre-Lie algebras [45] and its implications in the study of differential operators concludes the first section.

In Section 2.1 we consider branched rough paths defined on $\mathcal{H}_{\mathsf{CK}}$ and identify primitives in $\mathcal{H}_{\mathsf{CK}}$—call the space of these $\mathcal{P}$—as those elements which index paths: these form the "intrinsic correction terms" to the usual integration-by-parts identity. Integration against these is what makes the change-of-variable formula, Theorem 2.12 possible, the first of the two main results of this paper. Here we consider an equation similar in spirit to (0.5), but driven by X together with its correction terms: in and of itself, this type of equation is more general that those considered by previous authors. The Itô formula is achieved by lifting F to a Hopf algebra homomorphism $\mathbf{F}$ from $\mathcal{H}_{\mathsf{GL}}$ to the Hopf algebra of differential operators over the state space of the solution, which yields our "rough generator" (similar to (0.3)) to be integrated. When restricted to equations "without drift" (i.e. only driven by the original X and not its corrections), this yields a construction which can be compared to Kelly's, but does not rely on non-canonical rough path lifts and admits a more precise description of its coefficients as the Oudom-Guin extension of a pre-Lie morphism. Our Itô formula is particularly tractable in the special instance of quasi-geometric rough paths—those rough paths defined on Hoffman's quasi-shuffle algebra [27]—in which case it extends previous work of one of the authors [3]. Writing the quasi-shuffle algebra as a quotient of $\mathcal{H}_{\mathsf{CK}}$, and leveraging once again the algebra introduced previously to identify a section of this quotient map, leads to an explicit criterion for when a branched rough path is (quasi-)geometric, filling a notable gap in the literature.

In Section 3 we begin by introducing the Eulerian idempotent of a Hopf algebra, and use it to define an isomorphism $\mathrm{Log}\colon \mathcal{H}_{\mathsf{CK}} \to \sqcup\!\sqcup(\mathcal{P})$, the latter being the shuffle Hopf algebra over $\mathcal{P}$. This constitutes our second main result, Theorem 3.4. Since Log is defined solely in terms of operations intrinsic to the $\mathbf{B}_\infty$-algebra $\mathcal{H}_{\mathsf{CK}}$, it can be viewed as a natural transformation between the functors $\mathcal{H}_{\mathsf{CK}}$ and $\sqcup\!\sqcup(\mathcal{P})$. The proof that this is, indeed, bijective makes again use of the $\mathbf{B}_\infty$-structure, and explicit closed-form and recursive expressions for the inverse Exp are identified. Since the shuffle Hopf algebra is the one on which geometric rough paths are defined, this immediately yields a recipe for obtaining, given a branched rough path $\mathbf{X}$, a geometric one $\overline{\mathbf{X}}$ which carries the same information. By this we mean that an equation driven by $\mathbf{X}$ can be re-expressed as an equation driven by $\overline{\mathbf{X}}$ whose coefficients are explicitly computable in terms of pre-Lie products of coefficients obtained through the Hopf morphism $\mathbf{F}$ and the adjoint (under the $\mathcal{H}_{\mathsf{CK}}$-$\mathcal{H}_{\mathsf{GL}}$ duality) of Log. We proceed by showing how Exp extends Hoffman's exponential (after which it is named) between the quasi-shuffle and shuffle algebras, and further relate it to the arborified exponential introduced by Bruned, Curry and Ebrahimi-Fard [6]. Hoffman's original motivation to introduce this isomorphism was to study algebraic relations between multiple zeta values. We end the section with a comparison to the previous approaches of Hairer-Kelly and Boedihardjo-Chevyrev, providing further motivation for our work. More precisely, the dual of our isomorphism may be viewed as a special case of the latter, the general case of which does not however guarantee naturality.

In the final section, Section 4, we restrict our attention to the more tractable case in which the underlying vector space is one-dimensional (the "undecorated" setting). While there is not much to say about one-dimensional geometric rough paths (there is only one over each trace path, given by powers $X^n_{s,t}/n!$), one-dimensional branched rough paths are already interesting: their linear iterated integrals over a single primitive path component can be expressed as polynomials in the correction terms. This fact, which is proved using Eulerian and Dynkin idempotents, extends the well-known Kailath-Segall polynomials [51] to the branched setting. Our proof of such identities is new even in the classical semimartingale setting, in which such polynomials are the orthogonal polynomials for Gaussian processes (Hermite polynomials) and Poisson processes (Charlier polynomials), and are of great significance in stochastic analysis. We end the paper by laying out a general framework through which examples of branched rough paths may be constructed. Still working in the one-dimensional setting, and choosing a fractional Brownian motion with Hurst parameter $1/4$ as our trace path, we construct examples, motivated by recent results concerning Itô formulae in law. We consider a second-order correction term given by an independent Brownian motion, against which integrals must be defined in order to fix the definition of our rough path: two out of four such choices produce rough paths which are "truly branched", i.e. not quasi-geometric.

We believe that many of the modern challenges presented by stochastic analysis can be met by a careful and parsimonious treatment of the algebraic structures involved, with an emphasis on natural structure and functoriality, and look forward to further developments of the theory in this spirit, particularly in direction of SPDEs and regularity structures.

1. Algebraic preliminaries

1.1. The Connes-Kreimer Hopf algebra over a vector space.

We begin by introducing the Connes-Kreimer Hopf algebra over a vector space as an algebra of tensors. We assume the reader is familiar with the original construction [11] (see for example the survey [15]), as well as with the basics of Hopf algebras [39, 56].

Let $\mathscr{H}$ be the set of non-planar rooted forests (including the empty forest 1; we denote $\mathscr{H}_+ := \mathscr{H} \setminus \{1\}$), and $\mathscr{T}$ its subset of non-planar rooted trees. We will denote elements of $\mathscr{H}$ by $\mathscr{f}, \mathscr{g}, \ldots$ and elements of $\mathscr{T}$ by $\mathscr{s}, \mathscr{t}, \ldots$, and for $\mathscr{f} \in \mathscr{H}$ we let $[\mathscr{f}] \in \mathscr{T}$ denote the tree obtained by joining every root in $\mathscr{f}$ to a new vertex, the new root. We denote by $\{\mathscr{f}\}$ the underlying set of vertices of $\mathscr{f} \in \mathscr{H}$. This sets inherits a natural partial ordering from its forest structure. Given a set A, we denote $\mathbb{S}_A$ the permutation group of A, $\mathbb{S}_k := \mathbb{S}_{\{1,\ldots,k\}}$. We define $\mathbb{S}_\mathscr{f}$ to be the group of order-preserving permutations of $\{\mathscr{f}\}$ (the order coming from the forest structure), which we may view as a subgroup of the ordinary permutation group $\mathbb{S}_{\{\mathscr{f}\}}$. The set $\mathbb{S}_\mathscr{f}$ has the following inductive description: $\mathbb{S}_{[\mathscr{f}]} = \mathbb{S}_\mathscr{f}$, and for $\mathscr{f} = \mathscr{t}_1^{k_1} \cdots \mathscr{t}_n^{k_n}$ with $\mathscr{t}_i \neq \mathscr{t}_j$ for $i \neq j$, $\sigma \in \mathbb{S}_{\{\mathscr{f}\}}$ belongs to $\mathbb{S}_\mathscr{f}$ if

$$
\begin{gathered}
\sigma = ((\sigma^1_1 \sqcup \ldots \sqcup \sigma^1_{k_1}) \circ \pi_1) \sqcup \ldots \sqcup ((\sigma^n_1 \sqcup \ldots \sqcup \sigma^n_{k_n}) \circ \pi_n) \\
\sigma^j_i \in \mathbb{S}_{\mathscr{t}_i} \text{ acting on the } j^{\text{th}} \text{ copy of } \mathscr{t}_i, \quad \pi_i \in \mathbb{S}_{k_i}
\end{gathered}
\tag{1.1}
$$

i.e. acts disjointly on the vertices of each individual tree $\mathscr{t}_i$ by an element of $\mathbb{S}_{\mathscr{t}_i}$ and then disjointly permutes the sets of trees that are all equal.

Given a (real) finite-dimensional vector space U, we will now define a way of "labelling" the vertices of a forest without having to choose a basis of U. This and similar constructions have already appeared in the literature: [8] (motivated by the need to consider U infinite-dimensional), and [33, Ch. 5] (using $\mathbb{S}$-modules); we choose to adopt this notation in order to develop the theory in a coordinate-free manner.

Definition 1.1. For $\mathscr{f} = \mathscr{t}_1^{k_1} \cdots \mathscr{t}_n^{k_n}$, we define recursively

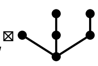
$$
U^{\boxtimes 1} := \mathbb{R}, \qquad U^{\boxtimes \mathscr{f}} := (U^{\boxtimes \mathscr{t}_1})^{\boxdot k_1} \boxtimes \cdots \boxtimes (U^{\boxtimes \mathscr{t}_n})^{\boxdot k_n}, \qquad U^{\boxtimes [\mathscr{f}]} := U^{\boxtimes \mathscr{f}} \boxtimes U\,, \tag{1.2}
$$

where $\boxtimes$ is the usual tensor product among vector spaces and $\boxdot$ is the symmetric tensor product (we prefer these symbols to $\otimes$, $\odot$, since the latter will be used more extensively later for external tensor products). ▲

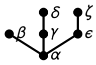
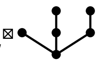
The presence of $\boxdot$ ensures that $U^{\boxtimes \mathscr{f}}$ encodes all the symmetries of the forest $\mathscr{f}$. From (1.1) it follows that there is a canonical isomorphism between this definition that of [8] as a quotient by the action of $\mathbb{S}_\mathscr{f}$

$$
U^{\boxtimes \mathscr{f}} \cong \frac{U^{\boxtimes \{\mathscr{f}\}}}{\mathbb{S}_\mathscr{f}}; \tag{1.3}
$$

both descriptions will be useful. At the two extremes, denoting ℓ_n the totally ordered tree with n vertices (called the n-*ladder*) and $\mathscr{r}_n$ the product of n disjoint roots, we have $U^{\boxtimes \ell_n} = U^{\boxtimes n}$ and $U^{\boxtimes \mathscr{r}_n} = U^{\boxdot n}$. A more generic example would be

$U^{\boxtimes}$  $= U \boxtimes (U \boxtimes U)^{\boxdot 2} \boxtimes U\,.$

Elementary tensors in $U^{\boxtimes \mathcal{f}}$ should be thought of as decorations of $\mathcal{f}$ with elements of U; we denote $\mathscr{H}(U)$ such generators and continue to use cursive letters to denote such decorated forests and trees, while we use Greek letters to denote elements of U. For example

 $= \beta \boxtimes ((\delta \boxtimes \gamma) \boxdot (\zeta \boxtimes \epsilon)) \boxtimes \alpha \ \in U^{\boxtimes}$  .

When considering decorated forests, we will use the notation $[\mathcal{f}]_\alpha$ to refer to the same operation as $[\mathcal{f}]$ in which we decorate the root with α.

Definition 1.2. Define the vector space, graded by *weight* of the forest (i.e. number of vertices, denoted $|\mathcal{f}|$)

$$\mathcal{H}(U) := \mathrm{span}_{\mathbb{R}\mathscr{H}(U) = \bigoplus_{\mathcal{f}\in\mathscr{H}} U^{\boxtimes \mathcal{f}}}, \qquad \mathcal{T}(U) := \mathrm{span}_{\mathbb{R}\mathscr{T}(U) = \bigoplus_{t\in\mathscr{T}} U^{\boxtimes t}} \tag{1.4}$$

and note that the vector space freely generated by unlabelled forests is recovered by taking $U = \mathbb{R}$. We denote $\mathcal{H}_+(U) := \mathrm{span}_{\mathbb{R}\mathscr{H}_+(U)}$. ▲

It is also helpful to introduce *proper forests* $\mathscr{F} := \mathscr{H} \setminus \mathscr{T}$, i.e. forests that are not trees (this includes the empty forest), and their span $\mathcal{F}$, so that $\mathcal{H} = \mathcal{T} \oplus \mathcal{F}$. Let Vec denote the category of finite-dimensional $\mathbb{R}$-vector spaces. Thanks to functoriality of $\boxtimes$ and $\boxdot$, $\mathcal{H}$ can be viewed as a covariant functor Vec $\to$ Vec. Let $\bigtimes(B) := \bigsqcup_{n\in\mathbb{N}} B^n$ denote the (disjoint) union of $\mathbb{N} \ni n$-fold Cartesian products (including the empty one) of a set B, and $\bigotimes(\,\cdot\,) = \bigoplus_{n\in\mathbb{N}}(\,\cdot\,)^{\otimes n}$ the tensor algebra functor Vec $\to$ Vec (note how we are now using the symbol $\otimes$ for external tensor product). The next proposition guarantees that certain operations performed on individual vertices of forests are promoted to natural transformations.

Proposition 1.3. *Let Φ be a collection of real numbers indexed by bijections of vertex sets*

$$\Phi = \big\{\phi_b \in \mathbb{R} \mid b \colon \{\mathcal{f}_1\} \sqcup \cdots \sqcup \{\mathcal{f}_m\} \leftrightarrow \{\mathcal{g}_1\} \sqcup \ldots \sqcup \{\mathcal{g}_n\},\ (\mathcal{f}_1, \ldots, \mathcal{f}_m), (\mathcal{g}_1, \ldots, \mathcal{g}_n) \in \bigtimes(\mathscr{H})\big\}$$

with the property that $\phi_b = \phi_{b\circ\sigma}$ for any $\sigma \in \mathbb{S}_{\mathcal{f}_1} \times \cdots \times \mathbb{S}_{\mathcal{f}_m}$ (viewed as a subgroup of $\mathbb{S}_{\{\mathcal{f}_1\}\sqcup\ldots\sqcup\{\mathcal{f}_m\}}$). For any finite-dimensional vector space U we set

$$\Phi_U(\mathcal{f}_1 \otimes \cdots \otimes \mathcal{f}_m) = \sum_{b \colon \{\mathcal{f}_1\}\sqcup\cdots\sqcup\{\mathcal{f}_m\} \leftrightarrow \{\mathcal{g}_1\}\sqcup\ldots\sqcup\{\mathcal{g}_n\}} \phi_b \mathcal{g}_1 \otimes \cdots \otimes \mathcal{g}_n\,, \tag{1.5}$$

where each $\mathcal{f}_i$ is a decorated forest (an elementary tensors in $U^{\boxtimes \mathcal{f}_i}$) and the undecorated forests $\mathcal{g}_j$ are given the decorations corresponding to the images of their vertices through the bijections b. Then Φ defines a natural endomorphism of the functor $\bigotimes \circ \mathcal{H} \colon$ Vec $\to$ Vec.

We comment on the definition of Φ. The fact that Φ is indexed by bijections is to be interpreted by saying that it corresponds to performing certain weight-preserving operations on an ordered collection of forests, e.g. cutting and grafting of edges, and that ϕ_b is the coefficient of $(\mathcal{g}_1, \ldots, \mathcal{g}_n)$ for the operation applied to $(\mathcal{f}_1, \ldots, \mathcal{f}_m)$, where the vertices of the forests $\mathcal{g}_j$ correspond to specific vertices of the forests $\mathcal{f}_i$ through b. It would not be possible to lift an element of $\mathrm{End}(\bigotimes(\mathcal{H}(\mathbb{R})))$ to one of $\mathrm{End}(\bigotimes(\mathcal{H}(U)))$ without knowing the precise correspondence between vertices: for example, knowing that the former maps ⫶ $\mapsto$ ⫶ does not, on its own, determine whether the induced one should map $= \alpha \boxtimes \beta$ to itself or to $= \beta \boxtimes \alpha$. We are considering endomorphisms of the whole tensor algebra over $\mathcal{H}$ for maximum flexibility, but often we will only need to consider Φ to be $\mathcal{H}^{\otimes 2} \to \mathcal{H}$ or similar, which can be obtained by the proposition by setting many of the coefficients to zero. The map b is required to be a bijection, since injections may leave certain labels undefined and surjections may lead to ill-defined maps (for example assigning the sole non-zero coefficient to the map ⫶ $\to \bullet$ that outputs the root does not lead to a well define map $U \boxtimes U \ni \alpha \boxtimes \beta \mapsto \beta \in U$); note also that allowing for multi-valued functions would violate linearity (e.g. duplicating a vertex would lead to $\Phi_U(\lambda\bullet_\alpha) = \lambda^2 \bullet_\alpha\bullet_\alpha$). The invariance requirement rules out maps that would not be well-defined on non-planar forests, e.g. $\bullet_\alpha\bullet_\beta \mapsto \bullet_\alpha \otimes \bullet_\beta$ (while $\bullet_\alpha\bullet_\beta \mapsto \bullet_\alpha \otimes \bullet_\beta + \bullet_\beta \otimes \bullet_\alpha$ is acceptable).

Proof of Proposition 1.3. First of all, (1.5) is a well defined linear map $\bigotimes(\mathcal{H}(U)) \to \bigotimes(\mathcal{H}(U))$, as a direct sum of its restrictions to $U^{\boxtimes \mathcal{f}_1} \otimes \cdots \otimes U^{\boxtimes \mathcal{f}_m}$. Indeed, viewing each of these as a map from $U^{\boxtimes\{\mathcal{f}_1\}} \otimes \cdots \otimes U^{\boxtimes\{\mathcal{f}_m\}}$, which it is equal to a sum of maps induced by permutations and thus linear, it passes to the quotient (where each $U^{\boxtimes \mathcal{f}_i}$ is given the description of (1.3)) thanks to $\mathbb{S}_{\mathcal{f}_1} \times \cdots \times \mathbb{S}_{\mathcal{f}_m}$-invariance. Next, we must show that for any $\mathbb{R}$-linear map $\theta \colon V \to W$ the square

$$\begin{array}{ccc}
\bigotimes(\mathcal{H}(V)) & \xrightarrow{\Phi_V} & \bigotimes(\mathcal{H}(V)) \\
{\scriptstyle \bigotimes(\mathcal{H}(\theta))}\big\downarrow & & \big\downarrow{\scriptstyle \bigotimes(\mathcal{H}(\theta))} \\
\bigotimes(\mathcal{H}(W)) & \xrightarrow{\Phi_W} & \bigotimes(\mathcal{H}(W))
\end{array}$$

commutes. We have

$$\begin{aligned}&\bigotimes(\mathcal{H}(\theta))\circ\Phi_V(\ell_1\otimes\cdots\otimes\ell_m)\\&=\sum_{b\colon\{\ell_1\}\sqcup\cdots\sqcup\{\ell_m\}\leftrightarrow\{g_1\}\sqcup\ldots\sqcup\{g_n\}}\phi_b\mathcal{H}(\theta)(g_1)\otimes\cdots\otimes\mathcal{H}(\theta)(g_n)\\&=\Phi_W\circ\bigotimes(\mathcal{H}(\theta))(\ell_1\otimes\cdots\otimes\ell_m)\end{aligned}$$

where the second identity follows from the fact that $\mathcal{H}(\theta)$ and b are interchangeable, by functoriality of (symmetric) tensor products. ◻

This result makes it possible to regard $\mathcal{H}$ and spaces defined in terms of it as undecorated, so that operations on it given of the form above, indexed by precise bijections between the vertex sets, automatically induce maps on the decorated spaces, natural in the vector space of decorations. This will be made precise once and for all in Corollary 1.17 for the main structures of interest to us. Recall that the *Connes-Kreimer Hopf algebra* $\mathcal{H}_{\mathrm{CK}} := (\mathcal{H}, \cdot, \Delta_{\mathrm{CK}})$ is the free Abelian algebra over $\mathcal{T}$, with coproduct Δ_{CK} defined in terms of admissible cuts. By applying Proposition 1.3 with $\Phi = \cdot, \Delta$, we obtain a graded connected Hopf algebra $\mathcal{H}_{\mathrm{CK}}(U) := (\mathcal{H}(U), \cdot, \Delta_{\mathrm{CK}})$. Let $\underline{\mathrm{gcHopf}}$ denote the category of graded, connected Hopf algebras, and recall that the forgetful functor from this category to that of graded, connected bialgebras is an equivalence, i.e. each graded connected bialgebra has a unique antipode, and these are automatically preserved by maps of graded, connected bialgebras, see [54]. This is what makes it possible to treat the category $\underline{\mathrm{gcHopf}}$ without ever mentioning antipodes. We now introduce the secondary structure on $\mathcal{H}_{\mathrm{CK}}(U)$ which will prove fundamental to our goals: cofree coalgebras. We refer the [34, Section 1.2.] for a complete reference on the topic. In what follows, we will always consider graded, connected coassociative coalgebras (C, Δ, ε). In particular, as a vector space $C \cong \bigoplus_{n\geq 0} C_n$, with $C_0 = \mathbb{R}1_C$ of dimension one. The reduced coproduct will be denoted by

$$\widetilde{\Delta}h := \Delta h - 1\otimes h - h\otimes 1,$$

and the space of primitive elements

$$\mathrm{Prim}(C) := \ker\widetilde{\Delta} = \{p\in C : \Delta p = 1\otimes p + p\otimes 1\}.$$

Both Δ and $\widetilde{\Delta}$ can be iterated, thanks to coassociativity, to yield linear maps $\Delta^n, \widetilde{\Delta}^n\colon C\to C^{\otimes(n+1)}$ for $n\geq 2$ defined recursively by the condition

$$\Delta^n = (\mathrm{id}\otimes\Delta^{n-1})\circ\Delta,\quad \widetilde{\Delta}^n = (\mathrm{id}\otimes\widetilde{\Delta}^n)\circ\widetilde{\Delta}.$$

We extend these operators by setting $\Delta^0 = \mathrm{id}$ and $\Delta^{-1}(\ell) = \varepsilon(\ell)1$, where ε is the counit of C. We will frequently use (possibly sum-free) Sweedler notation

$$\begin{aligned}\Delta h &=: \sum_{(h)} h_{(1)}\otimes h_{(2)} =: h_{(1)}\otimes h_{(2)}\\ \widetilde{\Delta}h &=: \sum_{(h)} h^{(1)}\otimes h^{(2)} =: h^{(1)}\otimes h^{(2)} =: h'\otimes h''\end{aligned}$$

and similar for higher order (reduced) coproducts.

Definition 1.4. We say that C is cofree over a vector space P with projection $\pi\colon C\to P$, sending 1_C to 0, if for all other graded and connected coalgebras D and linear maps $\phi\colon D\to P$ sending 1_D to 0 there exists a unique graded coalgebra map $\Phi\colon D\to C$ making the diagram

$$\begin{tikzcd} & C \arrow[d, "\pi"] \\ D \arrow[ur, dashed, "\Phi"] \arrow[r, "\phi"'] & P \end{tikzcd}\tag{1.6}$$

commute. We call π a *cofreeness projection*. ▲

The canonical way of constructing a model of cofree coalgebra by considering the tensor algebra $\bigotimes(P)$ with the deconcatenation coproduct

$$\Delta_\otimes(p_1\cdots p_n) = \sum_{k=0}^{n}(p_1\cdots p_k)\otimes(p_{k+1}\cdots p_n)\,.\tag{1.7}$$

and π is the canonical projection $\pi_1\colon\bigotimes(P)\to P$, see [34, Proposition 1.2.7]. It follows from the universal property (1.6) that if C is cofree over P with projection $\pi\colon C\to P$, π always restricts to a linear isomorphism $\mathrm{Prim}(C)\to P$, i.e. the cofreeness map essentially consists of a projection of C onto $\mathrm{Prim}(C)$. We emphasise that the structure of cofreeness depends on this projection: different maps π will yield different coalgebra isomorphisms $C\cong\bigotimes(P)$. Coalgebra morphisms from, and between, cofree coalgebras can be explicitly described in terms of their composition with the cofreeness projection.

Proposition 1.5. *If $C = \bigotimes(P)$ and $\pi = \pi_1$ in are as in Definition 1.4, the unique coalgebra map Φ is given explicitly by*

$$\Phi := \sum_{n=0}^{\infty} \phi^{\otimes n} \circ \widetilde{\Delta}^{n-1}. \tag{1.8}$$

In particular, if $D = (\bigotimes(R), \Delta_\otimes)$ for some vector space R, letting $n^h := n_1 + \ldots + n_h$

$$\Phi(r_1 \otimes \cdots \otimes r_n) = \sum_{\substack{k=1,\ldots,n \\ n_1,\ldots,n_k \geq 1 \\ n_1+\ldots+n_k=n}} \phi(r_1 \otimes \cdots \otimes r_{n_1}) \otimes \cdots \otimes \phi(r_{n^{k-1}+1} \otimes \cdots \otimes r_n)\,. \tag{1.9}$$

and Φ is invertible if and only if $\phi|_P$ is invertible.

Proof. The first identity follows from [34, Proposition 1.2.7]. The second follows from the explicit computation of $\Delta_\otimes^{(k-1)}$, see also [17, Theorem 11.2], and [17, Theorem 11.3] for the claim of bijectivity. □

Remark 1.6. The characterisation of cofreeness here presented works because we consider only connected graded coalgebras, whose coproduct is always conilpotent. Further generalisation to the general category of coalgebra would lead to a much more complicate universal object, see e.g. [53, Theorem 6.4.1]. ♦

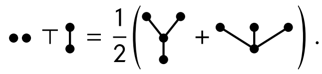
We now proceed to view $\mathcal{H}_{\mathrm{CK}}$ as cofree algebra. This is made possible thanks to the following operation:

Definition 1.7 (Natural growth, [5])**.** Define $\top \colon \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$, called *natural growth* as the unique linear map satisfying the conditions

$$f \top 1 = f = 1 \top f$$

on any forest f and for $f \neq 1 \neq g$, $f \top g$ is the sum over all the forests obtained by grafting f onto every vertex of g and normalizing [1] by $|g|$. In each term, all roots of f must be grafted onto the same vertex of g. ▲

For instance, we have



Note that $\top$ is not an associative operation, and when considering n-fold iterations of it, we will always compose it from the left, i.e.

$$h_1 \top \ldots \top h_n := (h_1 \top \ldots \top h_{n-1}) \top h_n.$$

The key property of this operation is obtained by relating $\top$ with the Connes-Kreimer primitive elements

$$\mathcal{P} := \mathrm{Prim}(\mathcal{H}_{\mathrm{CK}})\,.$$

Namely, the map $\top$ restricted to $\mathcal{H} \otimes \mathcal{P}$ satisfies the property

$$\Delta_{\mathrm{CK}}(h \top p) = (h \top p) \otimes 1 + (\mathrm{id} \otimes (\cdot\ \top p)) \circ \Delta_{\mathrm{CK}}(h)\,. \tag{1.10}$$

which is expressed by saying that for each $p \in \mathcal{P}$ the map $(\cdot\ \top p) \colon \mathcal{H}_{\mathrm{CK}} \to \mathcal{H}_{\mathrm{CK}}$ is a 1-cocycle for the coproduct [11, p.230]; we will give this property an interpretation in terms of integration in Section 2.1 below. The following result uses the 1-cocycle property to establish the fundamental fact that $\mathcal{H}_{\mathrm{CK}}$ is cofree.

Theorem 1.8 ([18, Théorème 48])**.** *The following map (denoted by the same symbol)*

$$\top \colon \bigotimes(\mathcal{P}) \to \mathcal{H}_{\mathrm{CK}}, \quad p_1 \otimes \cdots \otimes p_n \mapsto p_1 \top \ldots \top p_n \tag{1.11}$$

is a coalgebra isomorphism.

The projection associated to this isomorphism is given by the following

Proposition 1.9 ([17, Theorem 9.6])**.** *The projection $\pi \colon \mathcal{H}_{\mathrm{CK}} \to \mathcal{P}$ is given by the recursive formula*

$$\pi = \mathrm{id} - \top \circ (\mathrm{id} \otimes \pi) \circ \widetilde{\Delta}_{\mathrm{CK}} \tag{1.12}$$

From now on, π will refer to the specific projection of the proposition above. We thus have two gradings on $\mathcal{H}_{\mathrm{CK}}$:

$$\mathcal{H}_{\mathrm{CK}} = \bigoplus_{n \in \mathbb{N}} \mathcal{H}^{(n)}, \qquad \mathcal{H}_{\mathrm{CK}} = \bigoplus_{m \in \mathbb{N}} \mathcal{P}^{\top m} \tag{1.13}$$

where $\mathcal{H}^{(n)}$ denotes the space generated by forests of weight n and $\mathcal{P}^{\top m} := \top(\mathcal{P}^{\otimes m})$. Note that although the second grading is not locally finitely-dimensional (since $\mathcal{P}$ is infinite-dimensional), $\mathcal{P}$ can itself be graded into finitely-dimensional

[1] Here we follow [18] by including the normalizing factor, not present in the original definition.

spaces $\mathcal{P} = \bigoplus_n \mathcal{P}^{(n)}$, $\mathcal{P}^{(n)} := \pi(\mathcal{H}^{(n)})$. One may view elements of $\mathcal{P}^{\top m}$ corresponding to elementary tensor as "m-ladders decorated by primitives", since $\top(\bullet^{\otimes m}) = \ell_m$. We will refer to such m as the *degree of primitiveness*. Sometimes we will consider the associated filtrations, which we denote $\mathcal{H}^n := \bigoplus_{k=0}^n \mathcal{H}^{(k)}$ in the first case. On the other hand, we do not have a particular notation for the filtration associated to the second grading, which coincides with the *coradical filtration* of the coalgebra $(\mathcal{H}_{\mathrm{CK}}, \Delta_{\mathrm{CK}})$, see [17, section 11]; note that this does not require cofreeness to be defined. Since $\top$ preserves weight, degree of primitiveness is bounded by weight, these two gradings can be combined as

$$\mathcal{H}_{\mathrm{CK}} = \bigoplus_{0 \le m \le n} \pi_m(\mathcal{H}^{(n)}) \tag{1.14}$$

where we denote π_m the canonical projection $\mathcal{H}_{\mathrm{CK}} \twoheadrightarrow \mathcal{P}^{\top m}$. We can use Proposition 1.5 to express the inverse isomorphism $\top^{-1} \colon \mathcal{H}_{\mathrm{CK}} \to \bigotimes(\mathcal{P})$ in terms of π:

$$\top^{-1} = \sum_{n=0}^{\infty} \pi^{\otimes n} \circ \widetilde{\Delta}_{\mathrm{CK}}^{n-1} \tag{1.15}$$

from which, writing $\top \circ \top^{-1} = \mathrm{id}$, it follows that

$$\pi_m = \top \circ \pi^{\otimes m} \circ \widetilde{\Delta}_{\mathrm{CK}}^{m-1} \tag{1.16}$$

It is also possible to obtain a non-recursive formula for π_m, $m \ge 1$ using the operator $R_n \colon \mathcal{H}^{\otimes n} \to \mathcal{H}$, defined by

$$R_{n+1}(f_0 \otimes \cdots \otimes f_n) := f_0 \top R_n(f_1 \otimes \cdots \otimes f_n)\,;$$

notice that $R_n|_{\mathcal{P}^{\otimes n}} \neq \top|_{\mathcal{P}^{\otimes n}}$, due to lack of associativity of the binary operation $\top$.

Proposition 1.10. *For any finite-dimensional vector space U the following non-recursive formulae for $\pi \colon \mathcal{H}_{\mathrm{CK}} \to \mathcal{P}$ and $\top^{-1} \colon \mathcal{H}_{\mathrm{CK}} \to \bigotimes(\mathcal{P})$ hold*

$$\begin{aligned}
\pi &= \sum_{k \ge 1} (-1)^{k+1} R_k \circ \widetilde{\Delta}_{\mathrm{CK}}^{(k-1)},\\
\top^{-1} &= \sum_{n \ge 0} \sum_{k \ge n} \sum_{k_1 + \cdots + k_n = k} (-1)^{k+n} (R_{k_1} \otimes \cdots \otimes R_{k_n}) \circ \widetilde{\Delta}_{\mathrm{CK}}^{(k-1)}\,.
\end{aligned}$$

Proof. The first identity is obtained by iterating (1.12); the second by substituting it into (1.15). □

Remark 1.11. Identity (1.10) implies that the pair of operations $\top$ and Δ_{CK} turn $\mathcal{H}_{\mathrm{CK}}$ into an infinitesimal bialgebra [32]. The structure theorems and some of the formulae we obtain are a reflection of this fact. ♦

Example 1.12 (Change of coordinates for $\mathcal{H}_{\mathrm{CK}}^3$)**.** All elements of order 1, i.e. single vertices, are primitive. At level 2, we have

$$\pi(\bullet\alpha\bullet\beta) = \bullet\alpha\bullet\beta - \begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} - \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\end{smallmatrix}$$

so that

$$\bullet\alpha\bullet\beta = \pi(\bullet\alpha\bullet\beta) + \begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\end{smallmatrix}, \qquad \begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} = \bullet\alpha \top \bullet\beta$$

At level 3 we have

$$\begin{aligned}
\pi(\bullet\alpha\bullet\beta\bullet\gamma) &= \bullet\alpha\bullet\beta\bullet\gamma - \frac{1}{2}\sum_{\sigma \in \mathbb{S}_3} \bullet\sigma(\alpha)\begin{smallmatrix}\bullet\sigma(\gamma)\\ \vert\\ \bullet\sigma(\beta)\end{smallmatrix} + \frac{1}{2}\sum_{\sigma \in \mathbb{S}_3} \begin{smallmatrix}\bullet\sigma(\gamma)\\ \vert\\ \bullet\sigma(\beta)\\ \vert\\ \bullet\sigma(\alpha)\end{smallmatrix}\\
\pi\left(\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix}\right) &= \frac{1}{2}\left(\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} - \bullet\gamma\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\end{smallmatrix}\right) + \frac{1}{2}\left(\overset{\bullet\gamma\;\bullet\beta}{\underset{\bullet\alpha}{\vee}} - \overset{\bullet\alpha\;\bullet\beta}{\underset{\bullet\gamma}{\vee}}\right) + \frac{1}{2}\left(\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\\ \vert\\ \bullet\alpha\end{smallmatrix} - \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix}\right).
\end{aligned}$$

In particular, the last term is antisymmetric in the lower two indices; notice how this is not visible in the 1-dimensional undecorated case: this is another example of how, using the notation of Proposition 1.3, Φ_U can be zero for a particular choice of U even though Φ itself is not. More complicated relations can be expected at higher orders. From these computations we obtain the change of coordinates at level 3 (ℓ_3 is already expressed in the $\mathcal{P}^{\top\bullet}$ grading)

$$\begin{aligned}
\overset{\bullet\alpha\;\bullet\beta}{\underset{\bullet\gamma}{\vee}} &= \pi(\bullet\alpha\bullet\beta) \top \bullet\gamma + \begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix}\\
\bullet\alpha\bullet\beta\bullet\gamma &= \pi(\bullet\alpha\bullet\beta\bullet\gamma) + \sum_{\sigma \in \mathbb{S}_3}\Big(\bullet\sigma(\alpha) \top \pi(\bullet\sigma(\beta)\bullet\sigma(\gamma)) + \pi(\bullet\sigma(\beta)\bullet\sigma(\gamma)) \top \bullet\sigma(\alpha)\Big) + \sum_{\sigma \in \mathbb{S}_3} \begin{smallmatrix}\bullet\sigma(\gamma)\\ \vert\\ \bullet\sigma(\beta)\\ \vert\\ \bullet\sigma(\alpha)\end{smallmatrix}\\
\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} &= \pi\left(\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix}\right) + \bullet\beta \top \pi(\bullet\alpha\bullet\gamma) + \pi(\bullet\alpha\bullet\beta) \top \bullet\gamma + \begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\\ \vert\\ \bullet\alpha\end{smallmatrix}
\end{aligned}$$

♦

Remark 1.13 (Bases of primitives)**.** Let B be a basis of U, and denote $\mathscr{H}(B)$ to be elementary tensors in $\mathscr{H}(U)$ which are decorated with elements of B, a basis of $\mathcal{H}(U)$. The previous example shows that $\{\pi_1(f) \mid f \in \mathscr{H}(B)\}$, which generates $\mathcal{P}(U)$, is not an independent set. Extracting a basis from this set would involve understanding the precise structure of $\mathcal{P}(U)$, which seems like a complex task (and perhaps similar, in flavour, to how Hall bases are defined [48]). Since in this article we are interested in describing natural operations between algebraic structures, and since such a basis would necessarily "break the symmetry", we will avoid it and rather carry out all computations using forests natural maps. We will sometimes find it convenient to fix the basis B and perform computations in the induced forest basis $\mathscr{H}(B)$. One can exactly compute the dimension of the span of primitive elements with exactly n nodes as a function of $d := \dim U$. Denoting $h_n(d) := \dim \mathcal{P}(U) \cap \mathcal{H}_{\mathrm{CK}}^{(n)}(U)$, following the same argument as in [17, Proposition 7.2] one can show that the generating function

$$H_d(t) := \sum_{n\geq 1} h_n(d) t^n$$

satisfies

$$1 - H_d(t) = \frac{1}{R_d(t)}$$

where R_d is the generating function of the sequence $\dim \mathcal{H}_{\mathrm{CK}}^{(n)}(U)$. Using the coefficients found in [20, Example 1.2] we obtain, for the first few terms:

$$\begin{aligned}
h_1(d) &= d, \\
h_2(d) &= \frac{d(d+1)}{2}, \\
h_3(d) &= \frac{d(2d^2+1)}{3}, \\
h_4(d) &= \frac{d(9d^3+2d^2+3d+2)}{8}, \\
h_5(d) &= \frac{d(64d^4+20d^3-5d^2+6)}{30}.
\end{aligned}$$

♦

Cofree (graded, connected) Hopf algebras are intimately connected to the following notion for which we refer to [19, Ch. 2].

Definition 1.14 (commutative $\mathbf{B}_\infty$-algebra)**.** A vector space P equipped with a map $\langle\!\langle \cdot, \cdot \rangle\!\rangle \colon \bigotimes(P) \odot \bigotimes(P) \to P$, for which we will use the shorthand $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{i,j} := \langle\!\langle \cdot, \cdot \rangle\!\rangle|_{P^{\otimes i} \odot P^{\otimes j}}$, is said to be a commutative $\mathbf{B}_\infty$-algebra if the following conditions are satisfied:

1 for any $k \geq 0$,

$$\langle\!\langle \cdot, \cdot \rangle\!\rangle_{k,0} = \begin{cases} \mathrm{id}_P & k = 1 \\ 0 & \text{otherwise} \end{cases}$$

2 for any tensors $u, v, w \in \bigotimes(P)$ (and using Sweedler notation w.r.t. the unreduced $\Delta_\otimes^k$)

$$\sum_{k\geq 1} \sum_{(u)^k,(v)^k} \langle\!\langle \langle\!\langle u_{(1)}, v_{(1)} \rangle\!\rangle \otimes \cdots \otimes \langle\!\langle u_{(k)}, v_{(k)} \rangle\!\rangle, w \rangle\!\rangle = \sum_{k\geq 1} \sum_{(v)^k,(w)^k} \langle\!\langle u, \langle\!\langle v_{(1)}, w_{(1)} \rangle\!\rangle \otimes \cdots \otimes \langle\!\langle v_{(k)}, w_{(k)} \rangle\!\rangle \rangle\!\rangle,$$

▲

Theorem 1.15 ([19, Theorem 2])**.** *Let* $\mathrm{Bialg}(P)$ *be the set of commutative products on* $\bigotimes(P)$ *compatible with* $\Delta_\otimes$*, and* $\mathbf{B}_\infty(P)$ *be the set of* $\mathbf{B}_\infty$ *structures on* P*. These two sets are in bijection via the maps:*

$$\begin{aligned}
\mathbf{B}_\infty(P) &\longrightarrow \mathrm{Bialg}(P) \\
\langle\!\langle \cdot, \cdot \rangle\!\rangle &\longmapsto u * v := \sum_{k\geq 1} \sum_{(u)^k,(v)^k} \langle\!\langle u_{(1)}, v_{(1)} \rangle\!\rangle \otimes \cdots \otimes \langle\!\langle u_{(k)}, v_{(k)} \rangle\!\rangle \\
\mathrm{Bialg}(P) &\longrightarrow \mathbf{B}_\infty(P) \\
* &\longmapsto \langle\!\langle u, v \rangle\!\rangle := \pi_1(u * v)
\end{aligned}$$

We note that symmetry of $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ corresponds to commutativity of the product; the result holds more generally for non-symmetric brackets and non-commutative bialgebra products. Applied to our setting, we have that the free commutative product on forests, written in $\bigotimes(\mathcal{P}) \cong_{\mathrm{T}} \mathcal{H}_{\mathrm{CK}}$, can be recovered from the projection π by (m^h and n^h defined as in Proposition 1.5).

Corollary 1.16. *Let $p_1, \ldots, p_n, q_1, \ldots, q_m \in \mathcal{P}$. Then*

$$\begin{aligned}(p_1 \top \cdots \top p_m) \cdot (q_1 \top \cdots \top q_n) \\ = \sum_{k \geq 0} \sum_{\substack{m_1 + \ldots + m_k = m \\ n_1 + \ldots + n_k = n}} \pi((p_1 \top \ldots \top p_{m_1}) \cdot (q_1 \top \ldots \top q_{n_1})) \top \cdots \\ \cdots \top \pi((p_{m^{k-1}+1} \top \ldots \top p_m) \cdot (q_{n^{k-1}+1} \top \ldots \top q_n))\end{aligned} \tag{1.17}$$

We regard the category of commutative $\mathbf{B}_\infty$-algebras, $\underline{\text{comm}\mathbf{B}_\infty}$, as the full subcategory of Hopf algebras, consisting of graded, connected, commutative and cofree Hopf algebras with specified cofreeness projection, i.e. isomorphism to $\bigotimes(\mathcal{P})$ endowed with a compatible product. We do not require the cofreeness projection to be preserved: this would amount to only considering morphisms that preserve the primitiveness grading, whereas general Hopf maps only preserve the associated filtration, and would be too restrictive (with a notable exception following shortly). We end this subsection with the promised functoriality statement:

Corollary 1.17 ($\mathbf{B}_\infty$-Connes-Kreimer functor)**.** *$(\mathcal{H}_{\mathrm{CK}}, \cdot, \Delta_{\mathrm{CK}}, \pi)$ defines a functor $\underline{\mathrm{Vec}} \to \underline{\text{comm}\mathbf{B}_\infty}$, naturally isomorphic through $\top$ to $\bigotimes(\mathcal{P})$ (where $\mathcal{P}$ is viewed as a functor $\underline{\mathrm{Vec}} \to \underline{\mathrm{Vec}}$ and $\bigotimes$ as one $\underline{\mathrm{Vec}} \to \underline{\mathrm{Coalg}}$) endowed with the product (1.17). Moreover, for a linear map $\theta \colon V \to W$, the induced map preserves the cofreeness projection: $\pi_W \circ \mathcal{H}(\theta) = \mathcal{H}(\theta) \circ \pi_V$.*

Proof. We must show that induced maps $\mathcal{H}_{\mathrm{CK}}(\theta)$ are bialgebra maps that commute with the projections π. We perform this check for the coproduct, the other two checks are analogous: this amounts to the statement that, for $\theta \colon V \to W$, $\mathcal{H}(\theta)^{\otimes 2} \circ \Delta_{\mathrm{CK},V} = \Delta_{\mathrm{CK},W} \circ \mathcal{H}(\theta)$. Define Φ as in Proposition 1.3 by $\phi_b = 1$ if $m = 1$, $n = 2$ and b is precisely the bijection between the top part g_1 and the bottom part g_2 of an admissible cut of f_1, and $\phi_b = 0$ otherwise. Then for every finite-dimensional vector space U, $\Phi_U = \Delta_{\mathrm{CK},U}$, the Connes-Kreimer coproduct on $\mathcal{H}_{\mathrm{CK}}(U)$, and the required statement follows from the proposition. Note how, in particular, this implies that $\mathcal{H}_{\mathrm{CK}}(\theta)$ restricts to a map $\mathcal{P}(\theta) \colon \mathcal{P}(V) \to \mathcal{P}(W)$, which is required for functoriality of $\mathcal{P}$. ◻

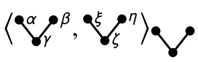
1.2. **Duality with Grossman-Larson.** The *Grossman-Larson Hopf algebra*, which we denote $\mathcal{H}_{\mathrm{GL}}$, was defined independently of (and prior to) the Connes-Kreimer Hopf algebra, in [23]. While $\mathcal{H}_{\mathrm{CK}}$ is useful to encode iterated integration, $\mathcal{H}_{\mathrm{GL}}$ was introduced to represent composition of differential operators. The fact that these two Hopf algebras could be related by duality was noticed only a decade later [28, 46]. The main purpose of this section is to revisit this dual pairing in the setting of functors to (co)free Hopf algebras. While we follow [28] as our main reference on the Grossman-Larson Hopf algebra, we proceed with one technical difference. We take the underlying vector space of $\mathcal{H}_{\mathrm{GL}}$ to be $\mathcal{H} = \mathrm{span}_{\mathbb{R}(\mathscr{H})}$, not $\mathrm{span}_{\mathbb{R}(\mathscr{T})}$ (this was also done, for example, in [26]). In this notation, the Grossman-Larson product of two forests is defined by

$$f \star t_1 \cdots t_n = (f_{(1)} \curvearrowleft t_1) \cdots (f_{(n)} \curvearrowleft t_n) f_{(n+1)} \quad t_k \in \mathcal{T},\ f \in \mathcal{H}, \tag{1.18}$$

where the term $f \curvearrowleft g$ is defined by summing over all possible graftings of roots in f to vertices in g, and the coproduct is the cofree cocommutative one over $\mathcal{T}$ (which partitions the multiset of trees that makes up f in all possible ways), denoted Δ_{GL}; in particular $s \star g = sg + s \curvearrowleft g$. Note that unlike in the definition of $f \top g$, different roots in f can be grafted to different vertices (also, there is no normalising factor in the definition of $\star$). The triple $\mathcal{H}_{\mathrm{GL}} \coloneqq (\mathcal{H}, \star, \Delta_{\mathrm{GL}})$ is shown to be a graded connected bialgebra, which we view again as a functor $\underline{\mathrm{Vec}} \to \underline{\mathrm{gcHopf}}$. We now wish to define a graded dual pairing $\mathcal{H}(V^*) \otimes \mathcal{H}(V) \to \mathbb{R}$, or equivalently (under the identification $\mathcal{H}(V^*)^* = \mathcal{H}(V)$, where the external duality is intended as graded duality) a graded map $\mathcal{H}(V) \to \mathcal{H}(V)$ which is full-rank at each level. In the notation of Proposition 1.3, set $\Phi_b = 1$ when $m = 1 = n$, $f_1 = g_1$ and $b \in \mathbb{S}_{f_1}$ and $\Phi_b = 0$ otherwise: this induces a natural map

$$\langle \cdot, \cdot \rangle \colon \mathcal{H}_{\mathrm{CK}}(V^*) \otimes \mathcal{H}_{\mathrm{GL}}(V) \to \mathbb{R} \tag{1.19}$$

given by a direct sum of pairings $V^{*\boxtimes f} \otimes V^{\boxtimes f} \to \mathbb{R}$ which counts the order automorphisms of f. For example



$$\begin{aligned} &= \langle \alpha \boxdot \beta, \xi \boxdot \eta \rangle \langle \gamma, \zeta \rangle \\ &= \langle \alpha, \xi \rangle \langle \beta, \eta \rangle \langle \gamma, \zeta \rangle + \langle \alpha, \eta \rangle \langle \beta, \xi \rangle \langle \gamma, \zeta \rangle \end{aligned}$$

Define $\mathcal{N}$ to be the orthogonal complement to $\mathcal{P}$ defined by $\top$, i.e. $\mathcal{N} \coloneqq \bigoplus_{n=2}^{\infty} \mathcal{P}^{\top n}$, so that $\mathcal{H}_{\mathrm{CK}} = \mathbb{R} \oplus \mathcal{P} \oplus \mathcal{N}$, and define $\mathcal{Q}$ to be the annihilator of $\mathbb{R} \oplus \mathcal{N}$ in $\mathcal{H}_{\mathrm{GL}}$ w.r.t. the pairing (1.19), i.e. "dual primitives"

$$\mathcal{Q} \coloneqq \{ q \in \mathcal{H}_{\mathrm{GL}} \mid \forall h \in \mathbb{R} \oplus \mathcal{N}\ \langle h, q \rangle = 0 \} \tag{1.20}$$

The isomorphism of Theorem 1.8 dualises to one $\top^* \colon (\mathcal{H}_{\mathrm{GL}}(V), \star) \to (\bigotimes(\mathcal{Q}(V)), \otimes)$ which identifies $\mathcal{H}_{\mathrm{GL}}(V)$ with the free algebra over $\mathcal{Q}(V)$, naturally in V. Putting everything together, we have the following result, original to [28, Proposition 4.4] (see [8, Theorem 4.8] for the decorated case); recall that by a pairing of bialgebras it is intended that the products and coproducts are dual to each other, where the pairing on the tensor product of bialgebras is given in the obvious way.

Theorem 1.18 (Duality between Connes-Kreimer and Grossman-Larson)**.** *The pairing* (1.19) *is a non-degenerate pairing of graded Hopf bialgebras, and the following diagram*

$$\begin{tikzcd} \mathcal{H}_{\mathrm{CK}}(V^*) \otimes \mathcal{H}_{\mathrm{GL}}(V) \arrow[d, "\top^{-1}\otimes\top^*"', "\cong"] \arrow[dr, "\langle\cdot,\cdot\rangle"] & \\ \bigotimes\mathcal{P}(V^*) \otimes \bigotimes\mathcal{Q}(V) \arrow[r, "\langle\cdot,\cdot\rangle_{\mathcal{P},\mathcal{Q}}"'] & \mathbb{R} \end{tikzcd} \tag{1.21}$$

in which $\langle\cdot,\cdot\rangle|_{\mathcal{P},\mathcal{Q}}$ *denotes the canonical extension to the tensor algebras of the restriction of* (1.19) *to* $\mathcal{P}(V^*)\otimes\mathcal{Q}(V)$*, commutes.*

Note that an application of Proposition 1.3 shows how this diagram is functorial in pairs of maps $\psi\colon W \to V$, $\varphi\colon V \to W$ (the maps going in opposite directions because of contravariance of the dualisation functor); from now on we omit V and V^* from the notation whenever possible, with the understanding that $\mathcal{P}$ and $\mathcal{Q}$ are to be evaluated on dual vector spaces, with functoriality intended in the above sense. We now consider computations in coordinates. Let B be a basis of V and B^* be its dual basis, and we identify $f \in \mathscr{H}(B)$ with its dual basis element in $\mathscr{H}(B^*)$. For $x \in \mathcal{H}_{\mathrm{GL}}$, $y \in \mathcal{H}_{\mathrm{CK}}$ and $f \in \mathscr{H}(B)$, $g \in \mathscr{H}(B^*)$ we denote $x^g := \langle g, x\rangle$ and $y_f := \langle y, f\rangle$. Given an undecorated forest $g \in \mathscr{H}$ we denote $\ell_B(f)$ the set of labellings of f with elements of B, i.e. its elements are maps $l\colon \{f\} \to B$ and $f_l \in \mathscr{H}(B)$ is the forest so labelled. In the unlabelled case, it follows directly from the definition of the pairing that

$$\langle f, g\rangle = \delta_{fg}\varsigma(f), \qquad f, g \in \mathscr{H}. \tag{1.22}$$

In the labelled case, if $f \in \mathscr{H}(B)$, $g \in \mathscr{H}(B^*)$, $\langle f, f\rangle$ is the cardinality of the *labelled* symmetry group of f, i.e. the group of order automorphisms of f that preserve the labelling, which may not coincide with $\varsigma(f)$.

Lemma 1.19 (Coordinates and pairing in the forest basis)**.**

1 *For* $f \in \mathscr{H}(B)$*, and* $\widetilde{f} \in \mathscr{H}$ *the same forest stripped of its labels, there are* $\frac{\varsigma(\widetilde{f})}{\langle f,f\rangle}$ *labellings* $l \in \ell(\widetilde{f})$ *s.t.* $\widetilde{f}_l = f$*;*

2 *For* $x \in \mathcal{H}_{\mathrm{GL}}$*,*
$$x = \sum_{f\in\mathscr{H}(B)} \langle f, f\rangle^{-1} x^f f = \sum_{h\in\mathscr{H}} \varsigma(h)^{-1} \sum_{l\in\ell(h)} x^{h_l} h_l\,;$$
and an analogous statement holds for $y \in \mathcal{H}_{\mathrm{CK}}$*;*

3 *For* $x \in \mathcal{H}_{\mathrm{GL}}$ *and* $y \in \mathcal{H}_{\mathrm{CK}}$
$$\langle y, x\rangle = \sum_{f\in\mathscr{H}(B)} \langle f, f\rangle^{-1} x^f y_f = \sum_{h\in\mathscr{H}} \varsigma(h)^{-1} \sum_{l\in\ell(h)} x^{h_l} y_{h_l}\,.$$

Proof. The set of labellings of $\widetilde{f}$ that result in $\widetilde{f}_l = f$ carries a transitive $\mathbb{S}_{\widetilde{f}}$-action whose stabiliser is the subgroup of label-preserving order automorphisms of f. The first assertion then follows from the interpretation of $\langle f, f\rangle$ as the cardinality of this subgroup and the orbit-stabiliser theorem. For all $g \in \mathscr{H}(B^*)$

$$\Big\langle g, \sum_{f\in\mathscr{H}(B)} \langle f, f\rangle^{-1} x^f f \Big\rangle = \sum_{f\in\mathscr{H}(B)} x^f \delta_{fg} = \langle g, x\rangle\,,$$

implying the first identity in (2) by non-degeneracy of the pairing; the second identity follows by decomposing the sum over $f \in \mathscr{H}(B)$ into one over $h \in \mathscr{H}$ and one over $l \in \ell(h)$ and applying (1) to count the number of collisions by which to divide. (3) is now a straightforward application of (2). □

We now describe how to compute the isomorphism $\top^*\colon (\mathcal{H}_{\mathrm{GL}}(\cdot), \star) \to (\bigotimes\mathcal{Q}(\cdot), \otimes)$ and its inverse. This broadly follows the same principle identified by [27, 46], according to which grafting is the dual operation to cutting. Since $\top$ involves a grafting operation on $\mathcal{H}_{\mathrm{CK}}$, its dual will involve a cutting operation on $\mathcal{H}_{\mathrm{GL}}$.

Definition 1.20. We define a *snip* of $h \in \mathscr{H}$ to be a non-empty collection of edges which share their lower endpoint; removing such edges separates h into two forests, call the one that contains the roots g and the one consisting of everything above the snipped edges f: we use the notation $(f, g) \in \mathrm{Snip}(h)$ to denote this correspondence (note that the forests (f, g) themselves do not identify the snip, and the same term in the sum may appear more than once). We then define

$$\widetilde{\Delta}_\top\colon \mathcal{H}_{\mathrm{GL}} \to \mathcal{H}_{\mathrm{GL}} \otimes \mathcal{H}_{\mathrm{GL}}, \qquad h \mapsto \sum_{(f,g)\in\mathrm{Snip}(h)} \frac{1}{|g|} f \otimes g\,.$$ ▲

Note that this operation does not define a coproduct, as it is not coassociative (we have only used the tilde as a reminder that both f and g are never empty). We define its iterations by

$$\widetilde{\Delta}_{\top}^{(n)} := (\widetilde{\Delta}_{\top}^{(n-1)} \otimes \mathrm{id}) \circ \widetilde{\Delta}_{\top} \colon \mathcal{H}_{\mathrm{GL}} \to \mathcal{H}_{\mathrm{GL}}^{\otimes(n+1)} \tag{1.23}$$

Proposition 1.21. *The adjoint of the binary operation $\top$ Definition 1.7 restricted to $\mathcal{H}_+^{\otimes 2}$ is $\widetilde{\Delta}_{\top}$.*

Proof. We follow the same ideas of [28, Proposition 4.4] with the main difference that snips replace admissible cuts (in fact, the proof is simplified since snips are easier to handle than admissible cuts). In the notation of Proposition 1.3, we consider two collections $\Phi = \{\phi_b\}_b, \Psi = \{\psi_b\}_b$, both set to vanish unless $m = 2$ (call the two forests f and g) and $n = 1$ (call it h). Set $\phi_b = |g|^{-1}$ if b carries f and g to two subforests of h, in a way that is an order isomorphism for both, and so that all the roots of the image of f are directly connected to the same vertex in the image of g; set $\phi_b = 0$ otherwise. Set $\psi_b = |g|^{-1}$ for each b obtained by first grafting f onto g, with all roots grafted onto the same vertex, and mapping the resulting forest to h order-isomorphically; set $\psi_b = 0$ otherwise. Then (under the identification $\mathcal{H}(V^*)^* = \mathcal{H}(V)$) the induced maps $\Phi_{V^*}, \Psi_{V^*} \colon \mathcal{H}(V^*)^{\otimes 2} \otimes \mathcal{H}(V) \to \mathbb{R}$

$$\begin{aligned} x \otimes y \otimes z &\mapsto \langle x \otimes y, \widetilde{\Delta}_{\top} z\rangle \\ x \otimes y \otimes z &\mapsto \langle x \top y, z\rangle \end{aligned}$$

To show that, for all V, that these two maps are equal, we show that $\phi_b = \psi_b$ for all f, g, h and b as above. Assume $\phi_b = |g|^{-1}$: then, since all roots of $b(f)$ are grafted onto the same vertex v in $b(g)$, the order isomorphisms $b|_f$ and $b|_g$ glue to an order isomorphism going from the forest obtained by grafting f onto $b^{-1}(v)$. This implies $\phi_b = |g|^{-1}$. Conversely, assume the latter: b must restrict to an order-isomorphism on both f and g, and calling u the vertex of g onto which the roots of f are grafted in the construction of b, $b(f)$ will be grafted onto $b(u)$. Therefore $\phi_b = |g|^{-1}$ and since ϕ_b, ψ_b can only take two values the proof is complete. □

The following explicit description of $\top^*$ immediately follows by dualising Theorem 1.8, (1.12) and (1.16).

Corollary 1.22 (Description of $\mathcal{H}_{\mathrm{GL}} \cong \bigotimes(Q)$). *The map*

$$\bigotimes(Q) \to \mathcal{H}_{\mathrm{GL}}, \qquad q_1 \otimes \cdots \otimes q_n \mapsto q_1 \star \cdots \star q_n$$

is an algebra isomorphism with inverse $\top^*$, and, denoting $\pi_n^* \colon \mathcal{H}_{\mathrm{GL}} \to Q^{\star n}$ the canonical projections, $\pi^* := \pi_1^*$, we have

$$\begin{aligned} \pi^* &= \mathrm{id} - \star \circ (\mathrm{id} \otimes \pi^*) \circ \widetilde{\Delta}_{\top} \\ \pi_n^* &= \star^{(n-1)} \circ \pi^{*\otimes n} \circ \widetilde{\Delta}_{\top}^{(n-1)} \end{aligned}$$

where $\star^{(n-1)}$ denotes the Grossman-Larson product of n arguments.

Note the slight abuse of notation regarding the projections: strictly speaking, π_n^* is not the dual of $\pi_n \colon \mathcal{H}_{\mathrm{CK}} \twoheadrightarrow \mathcal{P}^{\top n}$, rather of $\top^n \circ \pi_n \colon \mathcal{H}_{\mathrm{CK}} \to \mathcal{H}_{\mathrm{CK}}$.

Example 1.23 (Dual change of coordinates up to level 3)**.** The following is a system of generators of Q^3 obtained from $\pi^*(f), f \in \mathcal{H}_{\mathrm{GL}}^3$; to make it a basis it suffices pick a system of representatives for the symmetry and antisymmetry relations and take the labels to range in a basis of V.

$$\left\{\bullet\gamma,\ \bullet\alpha\bullet\beta,\ \bullet\alpha\bullet\beta\bullet\gamma,\ \tfrac{1}{2}\left(\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} - \bullet\beta\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix}\right)\ :\ \alpha, \beta, \gamma \in V\right\}$$

Note how, while ladders are well-adapted to the primitiveness grading of $\mathcal{H}_{\mathrm{CK}}$ ($\ell_n \in \mathcal{P}^{\top n}$), products of nodes $r_n \in Q$, as can be seen from Corollary 1.22 and the fact that $\widetilde{\Delta}_{\top} r_n = 0$ since there are no edges. We proceed to write all forest of weight ≤ 3 (not already contained in Q) in the $Q^{\star\bullet}$ grading:

$$\begin{aligned}
\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} &= \bullet\alpha\bullet\beta - \bullet\alpha \star \bullet\beta \\
\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} &= \frac{1}{2}\left(\bullet\beta\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix} - \bullet\gamma\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} + \bullet\alpha\bullet\beta\bullet\gamma\right) - \bullet\alpha\bullet\beta \star \bullet\gamma - \frac{1}{2}\bullet\alpha \star \bullet\beta\bullet\gamma + \bullet\alpha \star \bullet\beta \star \bullet\gamma \\
\begin{smallmatrix}\bullet\alpha & & \bullet\beta\\ \diagdown & & \diagup\\ & \bullet\gamma & \end{smallmatrix} &= \frac{1}{2}\left(\bullet\gamma\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\beta\end{smallmatrix} - \bullet\beta\begin{smallmatrix}\bullet\alpha\\ \vert\\ \bullet\gamma\end{smallmatrix} + \bullet\gamma\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\end{smallmatrix} - \bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix}\right) - \frac{1}{2}\left(\bullet\alpha \star \bullet\beta\bullet\gamma + \bullet\beta \star \bullet\alpha\bullet\gamma\right) + \bullet\alpha\bullet\beta \star \bullet\gamma \\
\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} &= \frac{1}{2}\left(\bullet\alpha\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\gamma\end{smallmatrix} - \bullet\gamma\begin{smallmatrix}\bullet\beta\\ \vert\\ \bullet\alpha\end{smallmatrix} - \bullet\alpha\bullet\beta\bullet\gamma\right) + \frac{1}{2}\bullet\gamma \star \bullet\alpha\bullet\beta
\end{aligned}$$

Note that the first term in each expansion is the image through π^* of the left hand side. ♦

1.3. **Pre-Lie algebras and higher-order differential operators.** We start by recalling the notion of (left) pre-Lie algebra.

Definition 1.24. A pre-Lie algebra is a vector space L endowed with a bilinear operation $\rhd : L \otimes L \to L$ such that the associator

$$\mathrm{a}_{\rhd}(x, y, z) := x \rhd (y \rhd z) - (x \rhd y) \rhd z$$

is symmetric in x and y. ▲

It follows immediately from the definition that $[-,-]_{\rhd} : L \otimes L \to L$ defines a Lie bracket on L. We are led thus to consider the universal enveloping algebra $\mathcal{U}(L)$, defined as the quotient of $\bigotimes(L)$ by the Hopf ideal generated by the relation $x \otimes y - y \otimes x - [x, y]_{\rhd}$. Describing the product of $\mathcal{U}(L)$ is difficult for a general Lie algebra, but in our present setting, one can characterize it as a modification of the symmetric algebra $\bigodot(L)$. We now briefly recall the construction, due to J.-M. Oudom and D. Guin [45].

The first step is to extend the pre-Lie product to $L \otimes \bigodot(L) \to \bigodot(L)$. This is done by simply requiring that it be a derivation, that is, $x \rhd 1 := 0$

$$x \rhd y_1 \odot \cdots \odot y_n := \sum_{j=1}^{n} y_1 \odot \cdots \odot y_{j-1} \odot (x \rhd y_j) \odot y_{j+1} \odot \cdots \odot y_n. \tag{1.24}$$

The extension to $\bigodot(L) \otimes \bigodot(L) \to \bigodot(L)$ is more involved. Consider $\bigodot(L)$ as a bialgebra with its cofree cocommutative coproduct ($\Delta_{\odot} x = x \otimes 1 + 1 \otimes x$ for $x \in L$). Define recursively, for $x, y \in L$ and $A, B, C \in \bigodot(L)$,

$$1 \rhd x := x \tag{1.25}$$
$$(x \odot A) \rhd y := x \rhd (A \rhd y) - (x \rhd A) \rhd y \tag{1.26}$$
$$A \rhd (B \odot C) := (A_{(1)} \rhd B) \odot (A_{(2)} \rhd C). \tag{1.27}$$

Given this extension, one introduces a non-commutative product on $\bigodot(L)$ by setting

$$A \circledast B := (A_{(1)} \rhd B) \odot A_{(2)}. \tag{1.28}$$

Note that from (1.27) it immediately follows that,

$$A \circledast (y_1 \odot \cdots \odot y_n) = (A_{(1)} \rhd y_1) \odot \cdots \odot (A_{(n)} \rhd y_n) \odot A_{(n+1)}. \tag{1.29}$$

Theorem 1.25 ([45]). *The map $\mathcal{U}(L) \to (\bigodot(L), \circledast, \Delta)$ induced by id_L is a Hopf isomorphism.*

We note the following property of the product.

Theorem 1.26 ([45]). *Let $A, B, C \in \bigodot(L)$. Then*

$$A \rhd (B \rhd C) = (A \circledast B) \rhd C.$$

In particular, restricting to $\bigodot(L) \otimes L \to L$, we see that L becomes a symmetric brace algebra.

The vector space $\bigodot(L)$ is now equipped with two products, related via (1.28).

Definition 1.27. We define the maps $r_{\rhd}^{(n)} : L^{\otimes n} \to L$ recursively by

$$r^{(1)}(x) := x$$
$$r^{(n+1)}(x \otimes x_1 \otimes \cdots \otimes x_n) := x \rhd r_{\rhd}^{(n)}(x_1 \otimes \cdots \otimes x_n).$$ ▲

Let us denote $[n] := \{1, \ldots, n\}$ and let $\mathrm{P}(n)$ be the collection of set partitions of $[n]$. Given $x_1, \ldots, x_n \in L$ and $I = \{i_1 < \cdots < i_k\} \subset [n]$, set

$$r_{\rhd}(x_I) := r_{\rhd}^{(k)}(x_{i_1} \otimes \cdots \otimes x_{i_k}).$$

We have the following formula.

Proposition 1.28. *Let $x_1, \ldots, x_n \in L$, then*

$$x_1 \circledast \cdots \circledast x_n = \sum_{B \in \mathrm{P}(n)} \bigodot_{I \in B} r_{\rhd}(x_I).$$

in $\mathcal{U}(L)$.

Proof. The proof is by induction, the case $n = 2$ being straightforward. Note that, by definition

$$x_1 \circledast \cdots \circledast x_n \circledast x_{n+1} = \Big((x_{1,(1)} \circledast \cdots \circledast x_{n,(1)}) \rhd x_{n+1} \Big) \odot (x_{1,(2)} \circledast \cdots \circledast x_{n,(2)}).$$

By Theorem 1.26 the right-hand side equals

$$r_{\rhd}^{(n+1)}(x_{1,(1)} \otimes \cdots \otimes x_{n,(1)} \otimes x_{n+1}) \odot (x_{1,(2)} \circledast \cdots \circledast x_{n,(2)}).$$

However, since the x_i are primitive, this equals

$$\sum_{I \subset [n]} r_{\rhd}(x_I \otimes x_{n+1}) \odot x^{\circledast}_{[n]\setminus I}$$

where $x^{\circledast}_J := x_{j_1} \circledast \cdots \circledast x_{j_k}$. By the induction hypothesis we obtain

$$\begin{aligned} x^{\circledast}_{[n]} \circledast x_{n+1} &= \sum_{I \subset [n]} \sum_{J \in \mathcal{P}([n]\setminus I)} r_{\rhd}(x_I \otimes x_{n+1}) \odot r_{\rhd}(x_{J_1}) \odot \cdots \odot r_{\rhd}(x_{J_k}) \\ &= \sum_{B \in \mathcal{P}(n+1)} \bigodot_{I \in B} r_{\rhd}(x_I). \end{aligned}$$

□

Concisely, this formula relates the two exponentials in $\bigodot(L)$, namely

$$\exp_{\circledast}\left(\sum_{i=1}^{k} x_i \right) = \exp_{\odot}\left(\sum_{n=1}^{\infty} r_{\rhd}(x_{[n]}) \right).$$

The following two examples of pre-Lie algebras are relevant to this article. Given two (undecorated) trees $\mathscr{s}, \mathscr{t} \in \mathscr{T}$ and a vertex $v \in \{\mathscr{t}\}$, the tree $\mathscr{s} \curvearrowleft_v \mathscr{t}$ is obtained by joining the root of $\mathscr{s}$ onto v by means of a new edge. We then extend this to an operator $\curvearrowleft \colon \mathscr{T} \otimes \mathscr{T} \to \mathscr{T}$ by setting

$$\mathscr{s} \curvearrowleft \mathscr{t} := \sum_{v \in \{\mathscr{t}\}} \mathscr{s} \curvearrowleft_v \mathscr{t}.$$

It is not difficult to show that this turns $\mathscr{T}$ into a pre-Lie algebra. Moreover, we have the following result.

Theorem 1.29 ([10]). *The pair $(\mathscr{T}, \curvearrowleft)$ is the free pre-Lie algebra on one generator.*

In this particular case, one can obtain an explicit form of the Oudom–Guin extension of grafting.

Proposition 1.30 ([45]). *Let $\mathscr{t}_1, \ldots, \mathscr{t}_n \in \mathscr{T}$ and $\mathscr{f} \in \mathscr{H}$. Then*

$$\mathscr{t}_1 \cdots \mathscr{t}_n \curvearrowleft \mathscr{f} = \sum_{(v_1,\ldots,v_n) \in \{\mathscr{f}\}^n} \mathscr{t}_1 \curvearrowleft_{v_1} (\cdots (\mathscr{t}_n \curvearrowleft_{v_n} \mathscr{f})).$$

By the Cartier–Milnor–Moore theorem, the Oudom–Guin product obtained in the previous section coincides with the Grossman-Larson product (1.18).

Now consider the space of smooth vector fields on a finite-dimensional vector space W, denoted here by $\mathfrak{X} = \mathfrak{X}(W) := C^\infty(W, W)$. Recall that for $F \in C^\infty(U, V)$, its n^{th} derivative can be seen as a map $\mathrm{D}^n F \in C^\infty(U, \mathcal{L}(U^{\otimes n}, V))$. Here and for the rest of the article we denote $\mathcal{L}(W, W')$ the space of linear maps between two vector spaces W and W'. We also write $\mathrm{End}(W) := \mathcal{L}(W, W)$.

Definition 1.31. Let $F, G \in \mathfrak{X}$. Define $F \rhd G \in \mathfrak{X}$ by

$$(F \rhd G)(x) := \mathrm{D}G(x)F(x).$$

▲

Proposition 1.32. *The pair $(\mathfrak{X}, \rhd)$ is a left pre-Lie algebra.*

Proof. Follows from an easy computation: from the chain rule we see that

$$\begin{aligned} (F \rhd (G \rhd H))(x) &= \mathrm{D}^2 H(x)(F(x), G(x)) + \mathrm{D}H(x)\mathrm{D}G(x)F(x) \\ ((F \rhd G) \rhd H)(x) &= \mathrm{D}H(x)\mathrm{D}G(x)F(x). \end{aligned}$$

Subtracting both expressions yields

$$\mathrm{a}_{\rhd}(F, G, H)(x) = \mathrm{D}^2 H(x)(F(x), G(x))$$

which is symmetric in F and G. □

Recall that $\mathfrak{X}$ is a Lie algebra with bracket

$$[F,G](x) = \mathrm{D}G(x)F(x) - \mathrm{D}F(x)G(x) = (F \triangleright G)(x) - (G \triangleright F)(x),$$

that is, the Lie bracket is induced by the pre-Lie product.

One obtains a representation ϱ of $\mathfrak{X}(W)$ on $C^\infty(W)$ by setting

$$\varrho(F)\varphi(x) \coloneqq \mathrm{D}\varphi(x)F(x).$$

Indeed, note that if $F, G \in \mathfrak{X}$

$$\begin{aligned}\varrho(F)\varrho(G)\varphi(x) - \varrho(G)\varrho(F)\varphi(x) &= \mathrm{D}\varphi(x)\mathrm{D}G(x)F(x) + \mathrm{D}^2\varphi(x)(F(x),G(x)) \\ &\quad - \mathrm{D}\varphi(x)\mathrm{D}F(x)G(x) - \mathrm{D}^2\varphi(x)(F(x),G(x)) \\ &= \mathrm{D}\varphi(x)[F,G](x) \\ &= \varrho([F,G])\varphi(x).\end{aligned}$$

The universal property of the universal envelope ensures that ϱ lifts uniquely to a representation of $\mathcal{U}(\mathfrak{X})$ on $\mathrm{End}(C^\infty(W))$. We let $\mathfrak{D} = \mathfrak{D}(W) \coloneqq \mathrm{im}(\varrho) \subset \mathrm{End}(C^\infty(W))$ be the algebra of differential operators on W. In particular, for any $F_1, \dots, F_k \in \mathfrak{X}(W)$ we have that

$$\varrho(F_1) \circ \cdots \circ \varrho(F_k)\varphi = \varrho(F_1 \circledast \cdots \circledast F_k)\varphi.$$

The following result makes the computation of ϱ fully explicit.

Lemma 1.33. *Let $F_1, \dots, F_n \in \mathfrak{X}(W)$ and $\varphi \in C^\infty(W)$. Then*

$$\varrho(F_1 \odot \cdots \odot F_n)\varphi(x) = \mathrm{D}^n\varphi(x)(F_1(x), \dots, F_n(x)).$$

Proof. The proof is by induction, the base case being true by definition. Let $F_0, \dots, F_n \in \mathfrak{X}$. From (1.29) we see that

$$F_0 \circledast (F_1 \odot \cdots \odot F_n) = F_0 \odot F_1 \odot \cdots \odot F_n + \sum_{j=1}^{n} F_1 \odot \cdots \odot (F_0 \triangleright F_j) \odot \cdots \odot F_n.$$

By the induction hypothesis and the isomorphism $\mathcal{U}(\mathfrak{X}) \cong \mathfrak{D}(W)$ the action of the left-hand side on φ equals

$$\mathrm{D}^{n+1}\varphi(x)(F_0(x), F_1(x), \dots, F_n(x)) + \sum_{j=1}^{n} \mathrm{D}^n\varphi(x)(F_1(x), \dots, \mathrm{D}F_j(x)F_0(x), \dots, F_n(x)).$$

Likewise, the action of the second term on the right-hand side equals

$$\sum_{j=1} \mathrm{D}^n\varphi(x)(F_1(x), \dots, (F_0 \triangleright F_j)(x), \dots, F_n(x)),$$

and since by definition we have that $(F_0 \triangleright F_j)(x) = \mathrm{D}F_j(x)F_0(x)$ the result follows. □

With this proposition in place we may compute composition of vector fields algebraically. As an example, consider $\varphi \in C^\infty(W,V)$ and $F, G, H \in \mathfrak{X}$. Then

$$\begin{aligned}\varrho(F)\varphi &= \mathrm{D}\varphi F \\ \varrho(F)\varrho(G)\varphi &= \varrho(F \circledast G)\varphi \\ &= \varrho(F \triangleright G + F \odot G)\varphi \\ &= \mathrm{D}\varphi(\mathrm{D}GF) + \mathrm{D}^2\varphi(F,G) \\ \varrho(F)\varrho(G)\varrho(H)\varphi &= (F \circledast G \circledast H)\varphi \\ &= \big(F \triangleright (G \triangleright H) + F \odot (G \triangleright H) + G \odot (F \triangleright H) + H \odot (F \triangleright G) + F \odot G \odot H\big)\varphi \\ &= \mathrm{D}\varphi\mathrm{D}^2H(F,G) + \mathrm{D}\varphi\mathrm{D}H\mathrm{D}GF + \mathrm{D}^2\varphi(F,\mathrm{D}HG) + \mathrm{D}^2\varphi(G,\mathrm{D}HF) \\ &\quad + \mathrm{D}^2\varphi(H,\mathrm{D}GF) + \mathrm{D}^3\varphi(F,G,H)\,.\end{aligned}$$

We consider now a map $F \in C^\infty(W, \mathcal{L}(V,W)) \cong \mathcal{L}(V, \mathfrak{X}(W))$. By Theorem 1.29 there is a unique pre-Lie morphism $\widehat{F}\colon \mathcal{T}(V) \to \mathfrak{X}(W)$ extending $\bullet_\alpha \mapsto F_\alpha$. Moreover, the universal property of the universal envelope ensures that it further extends to a Hopf morphism $\mathbf{F}\colon \mathcal{H}_{\mathrm{GL}} \to \mathfrak{D}(W)$ such that $\mathbf{F}|_{\mathcal{T}(V)} = \widehat{F}$. In particular from Lemma 1.33 we directly obtain the following recursive formula for $\widehat{F}$.

Proposition 1.34. *Let $\mathcal{f} = \mathcal{b}_1 \cdots \mathcal{b}_n \in \mathscr{H}(V)$, $t \in \mathscr{T}(V)$ and $F \in \mathcal{L}(V, \mathfrak{X})$. Then*

$$\widehat{F}_{\mathcal{f} \curvearrowright t}(x) = \mathrm{D}^n \widehat{F}_t(x)(\widehat{F}_{\mathcal{b}_1}(x), \ldots, \widehat{F}_{\mathcal{b}_n}(x)).$$

2. The Itô formula

2.1. **Branched rough paths and the change-of-variables formula.** In this section we relate the algebraic structure developed before with the notion of branched rough paths, original to [25]. We will follow [26], with the caveat that the dual pairing introduced in the previous section is not used therein, which will make our formulae look a little different. For $T \geq 0$ let $\Delta_T := \{(s,t) \in [0,T]^2 \mid s \leq t\}$. A *control* on $[0,T]$ is a continuous function $\omega \colon \Delta_T \to [0, +\infty)$ s.t. $\omega(t,t) = 0$ for $0 \leq t \leq T$ and is super-additive, i.e. $\omega(s,u) + \omega(u,t) \leq \omega(s,t)$ for $0 \leq s \leq u \leq t \leq T$; the main example of a control is $\omega(s,t) := t - s$, but allowing for more general ones make it possible to consider paths of bounded ρ-variation that may only be ρ^{-1}-Hölder up to reparametrisation (e.g. general continuous semimartingales as opposed to just Brownian motion). Throughout this article, ρ will denote a real number $\in [1, +\infty)$. Recall that $\mathcal{H}^n(V) := \bigoplus_{|\mathcal{f}| \leq n} V^{\otimes \mathcal{f}}$ (and similar) where $|\cdot|$ refers to forest weight, i.e. number of vertices. (Co)algebra operations on such a space are understood to be automatically truncated. To distinguish forests of different weight we will also adopt the notation $\mathcal{H}^{(n)}(V) := \bigoplus_{|\mathcal{f}|=n} V^{\otimes \mathcal{f}}$.

Definition 2.1 (Branched rough path)**.** A V-valued *ρ-branched rough path* on $[0,T]$ controlled by ω is a map

$$\mathbf{X} \colon \Delta_T \to \mathcal{H}^{\lfloor \rho \rfloor}_{\mathrm{GL}}(V), \quad (s,t) \mapsto \mathbf{X}_{s,t} \tag{2.1}$$

satisfying the following three properties:

- **Chen property**: $\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \star \mathbf{X}_{u,t}$, i.e. in coordinates
$$\langle \mathcal{f}, \mathbf{X}_{s,t} \rangle = \langle \mathcal{f}_{(1)}, \mathbf{X}_{s,u} \rangle \langle \mathcal{f}_{(2)}, \mathbf{X}_{u,t} \rangle$$
for $\mathcal{f} \in \mathscr{H}^{\lfloor \rho \rfloor}(V^*)$, and $0 \leq s \leq u \leq t \leq T$;
- **Character property**: $\Delta_{\mathrm{GL}} \mathbf{X}_{s,t} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{s,t}$, i.e. in coordinates
$$\langle \mathcal{f}\mathcal{g}, \mathbf{X}_{s,t} \rangle = \langle \mathcal{f}, \mathbf{X}_{s,t} \rangle \langle \mathcal{g}, \mathbf{X}_{s,t} \rangle$$
for $\mathcal{f}, \mathcal{g} \in \mathscr{H}(V^*)$ with $|\mathcal{f}| + |\mathcal{g}| \leq \lfloor \rho \rfloor$, and $0 \leq s \leq t \leq T$;
- **Regularity**: for $\mathcal{f} \in \mathscr{H}^{\lfloor \rho \rfloor}(V^*)$
$$\sup_{0 \leq s < t \leq T} \frac{|\langle \mathcal{f}, \mathbf{X}_{s,t} \rangle|}{\omega(s,t)^{|\mathcal{f}|/\rho}} < \infty\,.$$

We denote the set of these $\mathscr{C}^\rho_\omega([0,T], V)$. ▲

The intuitive meaning of a branched rough path is given by $\langle 1, \mathbf{X}_{s,t} \rangle = 1$, $\langle t_1 \cdots t_n, \mathbf{X}_{s,t} \rangle = \langle t_1 \mathbf{X}_{s,t} \rangle \cdots \langle t_n, \mathbf{X}_{s,t} \rangle$ for $t_k \in \mathscr{T}(V^*)$ and

$$\langle [\mathcal{f}]_\alpha, \mathbf{X}_{s,t} \rangle \text{ “=” } \int_s^t \langle \mathcal{f}, \mathbf{X}_{s,u} \rangle \, \mathrm{d}X^\alpha_u\,. \tag{2.2}$$

The meaning of this last identity is only heuristic: the integral is not well-defined by Stieltjes/Young integration when $\rho \geq 2$. When equipped with an initial value X_0, the components of $\mathbf{X}$ indexed by single labelled vertices are the increments of components X^α of a continuous function $X \colon [0,T] \to V$ called the *trace*; X belongs to $C^\rho_\omega([0,T], V)$, the set of functions $Z \colon [0,T] \to V$ with the property that, assuming V from now on to be finite-dimensional

$$\sup_{0 \leq s < t \leq T} \frac{|Z_{s,t}|}{\omega(s,t)^{1/\rho}} < \infty \tag{2.3}$$

where $Z_{s,t} := Z_t - Z_s$ is the increment and $|\cdot|$ is a generic norm on V.

In what follows we will write $\approx_m$ between two real-valued quantities dependent on $0 \leq s \leq t \leq T$ to mean that their difference is of order $O(\omega(s,t)^{m/\rho})$ on $[0,T]$ and simply $\approx$ (*almost* equal) to mean $\approx_{\lfloor \rho \rfloor + 1}$. Note that, since the vector spaces in which these quantities take values will always be finite-dimensional, the meaning of $\approx_m$ is independent of the choice of a norm of W. When considering expansions, quantities that are ≈ 0 will become negligible, which is why will will always be able to truncate sums at order $\lfloor \rho \rfloor$. We now give the definition of path controlled by a branched rough path, which will be used to define integration. We use sub/superscripts to elements of $\mathcal{H}$ to refer to (forest weight) grading and $\mathcal{L}$ refers to the space of linear maps. The following definition is original to [24, 25], see also [26] for a formulation more similar to the one below.

Definition 2.2. Let $\mathbf{X} \in \mathscr{C}^\rho_\omega([0,T],V)$. A U-valued $\mathbf{X}$-controlled path $\mathbf{H}$ is an element of $C^\rho_\omega([0,T],\mathcal{L}(\mathcal{H}^{\lfloor\rho\rfloor-1}_{\mathrm{GL}}(V),U))$ such that for $n = 0,\dots,\lfloor\rho\rfloor-1$

$$U \ni \langle \mathbf{H}_t, h\rangle \approx_{\lfloor\rho\rfloor-n} \langle \mathbf{H}_s, \mathbf{X}_{s,t} \star h\rangle, \quad h \in \mathcal{H}^{(n)}_{\mathrm{GL}}(V), \tag{2.4}$$

where to apply the pairing we are identifying $\mathcal{L}(\mathcal{H}^{\lfloor\rho\rfloor-1}_{\mathrm{GL}}(V),U) = \mathcal{H}^{\lfloor\rho\rfloor-1}_{\mathrm{CK}}(V^*) \otimes U$. Call the set of these $\mathscr{D}^\rho_{\mathbf{X}}(U)$. ▲

The intuition behind Definition 2.2 is that the *trace* $H \coloneqq \langle \mathbf{H}, 1\rangle \in C^\rho_\omega([0,T],U)$ is a path defined in terms of $\mathbf{X}$ (such as the solution to an $\mathbf{X}$-driven differential equation); higher terms in $\mathbf{H}$, the so-called *Gubinelli derivatives*, represent formal derivatives of H w.r.t. each component $\mathbf{X}$. We denote the portion of $\mathbf{H}$ that constitutes the Gubinelli derivatives [24], i.e. the restriction of $\mathbf{H}$ to forests of positive degree, $\mathbf{H}_\bullet$.

An essential property of the space of $\mathbf{X}$-controlled paths is that it is stable under images through smooth functions.

Lemma 2.3 (Smooth functions of controlled paths are controlled paths, [25, Lemma 8.4])**.** *Let* $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(W)$ *and* $\varphi \in C^\infty(W,Z)$. *Then the definition*

$$\varphi(\mathbf{H})_\bullet \coloneqq \sum_{n=1}^{\lfloor\rho\rfloor-1} \frac{1}{n!}\mathrm{D}^n\varphi(H) \circ \mathbf{H}^{\otimes n} \circ \widetilde{\Delta}^{(n-1)}_{\mathrm{GL}} \tag{2.5}$$

defines an element $\varphi(\mathbf{H}) \in \mathscr{D}_{\mathbf{X}}(Z)$ *with trace* $\varphi(H)$.

Recall that $\mathrm{D}^n\varphi \in C^\infty(W,\mathcal{L}(W^{\otimes n},Z))$; in coordinates this reads

$$\langle \varphi(\mathbf{H}), f\rangle = \sum_{n=1}^{\lfloor\rho\rfloor-1} \frac{1}{n!}\mathrm{D}^n\varphi(H)(\mathbf{H}_{f^{(1)}} \otimes \cdots \otimes \mathbf{H}_{f^{(n)}}),$$

where recall that the Sweedler notation indicates summing over all ways of partitioning up the forest f into n non-empty subforests.

Integration against controlled paths is usually introduced by considering $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(V,W))$ and defining $\int_0^T \mathbf{H}\mathrm{d}\mathbf{X} \in W$ as the limit

$$\int_0^T \mathbf{H}\mathrm{d}\mathbf{X} = \lim_{|\pi_n|\to 0} \sum_{[s,t]\in\Pi_n} \langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle \tag{2.6}$$

where $|\pi_n|$ is the mesh-size of the partition π_n of the interval $[0,T]$. The intuition is the same as in the better-known case of $\rho \in [2,3)$ [21]: compensating the ordinary Riemann sums with the higher terms of $\mathbf{X}$ contracted against the Gubinelli derivatives of $\mathbf{H}$ result in almost additivity of the two parameter function $\langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle$. Here we are interested in a generalisation of this operation of integration, which will be sufficiently expressive to write general controlled paths (with an extra degree of controlledness — see Remark 2.13 below) as integrals.

We begin by observing that primitive elements have a special significance in the context of rough paths. Indeed, $X^p_{s,t} \coloneqq \langle p, \mathbf{X}_{s,t}\rangle$ for $p \in \mathcal{P}$ are increments of paths: for $s \le u \le t$

$$X^p_{s,u} + X^p_{u,t} = \langle p \otimes 1 + 1 \otimes p, \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}\rangle = \langle \Delta_{\mathrm{CK}} p, \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}\rangle = \langle p, \mathbf{X}_{s,t}\rangle = X^p_{s,t}$$

(note the use of normal font for X to emphasise evaluation on primitives, i.e. this is viewed as an extension of the trace of $\mathbf{X}$). It is therefore natural to wish for integration against $X^{\alpha\in V}$ to extend to $X^{p\in\mathcal{P}}$. The algebraic counterpart to this operation is provided by $\top$: the following heuristic identity

$$\langle h \top p, \mathbf{X}_{s,t}\rangle \text{“="} \int_s^t \langle h, \mathbf{X}_{s,u}\rangle \,\mathrm{d}X^p_u \tag{2.7}$$

is supported by the fact that it satisfies the correct rule for breaking up the integral on $[s,t]$ into one on $[s,u]$ and one on $[u,t]$ (the so-called Chasles relation):

$$\begin{aligned}\langle h \top p, \mathbf{X}_{s,t}\rangle \text{“="} \int_s^t \langle h, \mathbf{X}_{sr}\rangle \,\mathrm{d}X^p_r &= \int_s^u \langle h, \mathbf{X}_{sr}\rangle \,\mathrm{d}\mathbf{X}^p_r + \int_u^t \langle h, \mathbf{X}_{sr}\rangle \,\mathrm{d}X^p_r \\ &= \int_s^u \langle h, \mathbf{X}_{sr}\rangle \,\mathrm{d}X^p_r + \langle h_{(1)}, \mathbf{X}_{s,u}\rangle \int_u^t \langle h_{(2)}, \mathbf{X}_{ur}\rangle \,\mathrm{d}\mathbf{X}^p_r \\ &\text{“="} \langle h \top p, \mathbf{X}_{s,u}\rangle + \langle h_{(1)} \otimes (h_{(2)} \top p), \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}\rangle\,.\end{aligned}$$

Even though the first and last identities above are heuristic, the first and last items are actually equal by the Chen identity and the 1-cocycle property of $\top$, that is, (1.10). Note that (2.7) extends (2.2) thanks to the fact that $h \top \bullet_\alpha = \mathcal{B}^\alpha_+(h)$ where $\mathcal{B}_+$ is defined by

$$\mathcal{B}_+ \colon \mathcal{H}_{\mathrm{CK}} \otimes V^* \to \mathcal{H}_{\mathrm{CK}}, \qquad \mathcal{B}^\alpha_+(f) \coloneqq [f]_\alpha$$

still satisfies (1.10) but is defined on a restricted space. The presence of this operator is what made it possible, in the literature published so far, to define integration against $X^{\alpha\in V}$. Below we will see examples of Hopf algebras that do not have such an operator, and that are thus not good candidates to describe an integration theory.

In fact, something more can be said in the way of making (2.7) rigorous. Recall that for $\mathbf{X} \in \mathscr{C}^\rho_\omega([0,T],V)$ we can extend it to arbitrarily high degree by defining its *branched signature* as the convergent limit of Grossman-Larson products

$$\mathbb{S}(\mathbf{X})_{s,t} := \lim_{n\to\infty} \bigstar_{[u,v]\in\Pi_n} \mathbf{X}_{u,v} \quad \in \mathcal{H}_{\mathrm{GL}}((V)) \tag{2.8}$$

independently of the partition Π_n with vanishing mesh size as $n \to \infty$ [25, Theorem 7.3]. The signature extends the rough path, in the sense that the projection of $\mathbb{S}(\mathbf{X})$ onto $\mathcal{H}_{\mathrm{GL}}^{\lfloor\rho\rfloor}$ equals $\mathbf{X}$, but $\mathbb{S}(\mathbf{X})$ belongs to the algebraic (ungraded) dual of $\mathcal{H}_{\mathrm{CK}}(V^*)$, defined by taking a direct product in (1.4) instead of the direct sum. From this it follows that $\top^*(\mathbb{S}(\mathbf{X})_{s,t}) \in \bigotimes((Q((V))))$ satisfies a similar expression to the above, with a tensor product over intervals in Π_n in place of the Grossman-Larson product, the same expression valid for the signature of a geometric rough path [38, Theorem 2.2.1]. In particular, if $p \in \mathcal{P}$ and $h \in \mathcal{H}_{\mathrm{CK}} \top (\bigoplus_{n>\lfloor\rho\rfloor-|p|} \mathcal{P}^{(n)})$, i.e. h is a linear combination of elements of the form $k \top q$ with $|p|+|q| \geq \lfloor\rho\rfloor+1$, the path $u \mapsto \langle h, \mathbb{S}(\mathbf{X})_{s,u}\rangle$ is Young integrable against the path $u \mapsto X^p_u := \langle p, \mathbb{S}(X)_{s,u}\rangle$, and (2.7) (with $\langle h, \mathbf{X}\rangle$ replaced with $\langle h, \mathbb{S}(\mathbf{X})_{s,u}\rangle$) holds literally, with the integral taken in the sense of Young.

The next lemma extends this notion of integration to general $\mathbf{X}$-controlled paths.

Proposition 2.4 (Rough integration of controlled paths against $X^{p\in\mathcal{P}}$)**.** *Let* $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(Q,W))$. *Then*

$$W \ni \langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle - \langle \mathbf{H}_s, \mathbf{X}_{s,u}\rangle - \langle \mathbf{H}_u, \mathbf{X}_{u,t}\rangle \approx 0$$

where we are identifying

$$\mathcal{L}(\mathcal{H}_{\mathrm{GL}}, \mathcal{L}(Q,W)) = \mathcal{L}(\mathcal{H}_{\mathrm{GL}} \otimes Q, W) = (\mathcal{H}_{\mathrm{CK}} \otimes \mathcal{P}) \otimes W \xrightarrow{\top\otimes \mathrm{id}} \mathcal{H}_{\mathrm{CK}} \otimes W \tag{2.9}$$

to apply the pairing. This implies that the limit

$$W \ni \int_s^t \mathbf{H}\,\mathrm{d}\mathbf{X} := \lim_{n\to\infty} \sum_{[u,v]\in\Pi_n} \langle \mathbf{H}_u, \mathbf{X}_{uv}\rangle$$

exists independently of the sequence of partitions $(\Pi_n)_n$ on $[s,t]$ with vanishing mesh size as $n \to \infty$; it is the unique function on the simplex $\Delta[0,T]$, additive in the sense that $\int_s^u \mathbf{H}\,\mathrm{d}\mathbf{X} + \int_u^t \mathbf{H}\,\mathrm{d}\mathbf{X}$ for $s \leq u \leq t$, s.t. $\int_s^t \mathbf{H}\,\mathrm{d}\mathbf{X} \approx \langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle$. Moreover, setting $\langle \int \mathbf{H}\,\mathrm{d}\mathbf{X}, h \star q\rangle := \langle \mathbf{H}_q, h\rangle \in W$ for $h \in \mathcal{H}_{\mathrm{GL}}$ and $q \in Q$ (where now $\mathbf{H}_q \in \mathcal{L}(\mathcal{H}_{\mathrm{GL}}, W) = \mathcal{H}_{\mathrm{CK}} \otimes W$) defines an element of $\mathscr{D}_{\mathbf{X}}(W)$.

Remark 2.5. Strictly speaking the contractions in the statement of Proposition 2.4 contain summations over forests up to a certain degree depending on the regularity of $\mathbf{X}$. We chose not to include this in the notation at the expense of possibly including terms of high regularity. In practice this is not an issue since condition (2.4) and its version in Proposition 2.4 only hold up to a remainder, so when doing computations one can just discard these extra terms. ♦

Proof of Proposition 2.4. The proof is standard and easily adapted from the case of geometric rough paths. Denote $\mathbf{X}^{m_1\ldots m_k}$ the restriction of $\mathbf{X}$ to $\mathcal{P}^{m_1\top\ldots\top m_k} := \top(\mathcal{P}^{(m_1)} \otimes \cdots \otimes \mathcal{P}^{(m_k)})$ (recall that $\mathcal{P}^{(m)}$ are Connes-Kreimer primitives of homogeneous weight equal to m), and similarly for $\mathbf{H}_{m_1\ldots m_k}$ the projection of $\mathbf{H}$ onto $\mathcal{P}^{m_1\top\ldots\top m_k} \otimes W$. Then, implying below a sum on k and on $m_i \geq 1$ subject to $m_1 + \ldots + m_k \leq \lfloor\rho\rfloor$

$$\begin{aligned}
&\langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle - \langle \mathbf{H}_s, \mathbf{X}_{s,u}\rangle - \langle \mathbf{H}_u, \mathbf{X}_{u,t}\rangle \\
&= \langle \mathbf{H}_{m_1\ldots m_k;s}, \mathbf{X}^{m_1\ldots m_k}_{s,t}\rangle - \langle \mathbf{H}_{m_1\ldots m_k;s}, \mathbf{X}^{m_1\ldots m_k}_{s,u}\rangle - \langle \mathbf{H}_{m_1\ldots m_k;u}, \mathbf{X}^{m_1\ldots m_k}_{u,t}\rangle \\
&= \langle \mathbf{H}_{m_1\ldots m_k;s}, \mathbf{X}^{m_1\ldots m_k}_{s,t} - \mathbf{X}^{m_1\ldots m_k}_{s,u} - \mathbf{X}^{m_1\ldots m_k}_{u,t}\rangle - \langle \mathbf{H}_{m_1\ldots m_k;su}, \mathbf{X}^{m_1\ldots m_k}_{u,t}\rangle \\
&= \sum_{i=1}^{k-1} \langle \mathbf{H}_{m_1\ldots m_k;s}, \mathbf{X}^{m_1\ldots m_i}_{s,u} \otimes \mathbf{X}^{m_{i+1}\ldots m_k}_{u,t}\rangle - \langle \mathbf{H}_{m_1\ldots m_k;su}, \mathbf{X}^{m_1\ldots m_k}_{u,t}\rangle \\
&\approx \sum_{i=1}^{k-1} \langle \mathbf{H}_{m_1\ldots m_k;s}, \mathbf{X}^{m_1\ldots m_i}_{s,u} \otimes \mathbf{X}^{m_{i+1}\ldots m_k}_{u,t}\rangle - \langle \mathbf{H}_s, \mathbf{X}_{s,u} \otimes \mathbf{X}^{m_1\ldots m_k}_{u,t}\rangle \\
&= 0
\end{aligned}$$

where we have used (2.4) combined with the fact that $y = \mathbf{X}^{m_1\ldots m_k}_{u,t}$ contributes an additional $O(\omega(s,t)^{m_1+\ldots+m_k})$. The classical statement on almost-additivity [38, Theorem 3.3.1] implies existence of the limit, uniqueness of the additive functional, and controlledness at level 0. Since the statement extends linearly, we may assume y to be of the form $y = h \star q$, and we have

$$\langle \textstyle\int_0^t \mathbf{H}\,\mathrm{d}\mathbf{X}, y\rangle = \langle \mathbf{H}_{q;t}, h\rangle \approx_{\lfloor p\rfloor - |h|} \langle \mathbf{H}_{q;s}, \mathbf{X}_{s,t} \star h\rangle = \langle \textstyle\int_0^s \mathbf{H}\,\mathrm{d}\mathbf{X}, \mathbf{X}_{s,t} \star y\rangle$$

and notice that integration has introduced one extra (inhomogeneous) degree of controlledness (corresponding to $|q|$ above); stated otherwise, some terms in this controlled path could be discarded, but this is of little importance. This concludes the proof. ◻

We now consider computations in coordinates. When considering, say, the integral against a V-valued bounded variation path $\int H\,\mathrm{d}X$ with H is a $\mathcal{L}(V, W)$-valued path, given a basis of V it is often helpful to decompose the integral as a sum $\int H_\gamma\,\mathrm{d}X^\gamma$ over the chosen coordinates. This carries over straightforwardly to the case of rough integrals in the case in which $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(V, W))$, but if $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(Q, W))$ a basis of Q would be needed. As stated in Remark 1.13 such bases are difficult to construct, and we will not do so here; the question then becomes whether it is possible to express rough integrals as linear combinations of ones against one-dimensional components of the rough path (in which the integrand however is still controlled by the whole of $\mathbf{X}$), using only a basis of V and the projection $\pi\colon \mathcal{H}_{\mathrm{CK}} \twoheadrightarrow \mathcal{P}$ to non-injectively obtain the components over which to sum. To do this, having picked a basis B of V (with dual basis B^* of V^*), we introduce the family of linear maps indexed by $f \in \mathscr{H}(B)$

$$\Pi_f\colon \mathcal{H}_{\mathrm{CK},+} \to \mathcal{H}_{\mathrm{CK},+}, \qquad h \top p \mapsto \frac{\langle p, f\rangle}{\langle \pi(f), f\rangle} h \top \pi(f) \tag{2.10}$$

where the argument is uniquely decomposed in the form considered with $p \in \mathcal{P}$ and $h \in \mathcal{H}_{\mathrm{CK}}$. Note that if p is in the linear span of $\pi(f)$, say for $p = \lambda\pi(f)$

$$\Pi_f(h \top p) = \frac{\langle \lambda\pi(f), f\rangle}{\langle \pi(f), f\rangle} h \top \pi(f) = h \top p$$

implying that Π_f is a projection onto $\{g \top \pi(f) \mid g \in \mathscr{H}(B^*)\}$. In the next lemma we use the same notation as in Lemma 1.19.

Lemma 2.6 (Rough integral in coordinates)**.**

1 *$y \in \mathcal{H}_{\mathrm{CK},+}$ can be written as*

$$y = \sum_{f \in \mathscr{H}(B)} \frac{\langle \pi(f), f\rangle}{\langle f, f\rangle} \Pi_f(y)\,;$$

2 *Let $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(Q, W))$. Then for $f \in \mathscr{H}(B)$, identifying $\mathbf{H}$ with an element of $\mathcal{H}_{\mathrm{CK}} \otimes W$ thanks to (2.9), $\Pi_f(\mathbf{H}) \in \mathscr{D}_{\mathbf{X}}(\mathcal{L}(Q, W))$ and*

$$\langle \pi(f), f\rangle^{-1} \int_s^t \mathbf{H}_{\pi^*(f)}\,\mathrm{d}\mathbf{X}^{\pi(f)} \coloneqq \int_s^t \Pi_f(\mathbf{H})\,\mathrm{d}\mathbf{X} \approx \langle \pi(f), f\rangle^{-1} \sum_{g \in \mathscr{H}} \varsigma(g)^{-1} \sum_{l \in \ell(g)} \mathbf{H}_{g_l \star \pi^*(f);s} \mathbf{X}_{s,t}^{g_l \top \pi(f)}$$

where again we use (2.9) to identify $\mathbf{H}$ with an element of $\mathcal{H}_{\mathrm{CK}} \otimes W$ which we evaluate on $g_l \star f$;

3 *Finally,*

$$\int_s^t \mathbf{H}\,\mathrm{d}\mathbf{X} = \sum_{f \in \mathscr{H}} \varsigma(f)^{-1} \sum_{l \in \ell(f)} \int_s^t \mathbf{H}_{\pi^*(f_l)}\,\mathrm{d}\mathbf{X}^{\pi(f_l)} \approx \sum_{f, g \in \mathscr{H}} \varsigma(f)^{-1} \varsigma(g)^{-1} \sum_{\substack{l \in \ell(f)\\ m \in \ell(g)}} \mathbf{H}_{g_m \star \pi^*(f_l);s} \mathbf{X}_{s,t}^{g_m \top \pi(f_l)}\,.$$

Proof. For $h \in \mathcal{H}_{\mathrm{CK}}$ and $p \in \mathcal{P}$ we have

$$p = \sum_{f \in \mathscr{H}(B)} \frac{\langle p, f\rangle}{\langle f, f\rangle} \pi(f)$$

so that

$$h \top p = \sum_{f \in \mathscr{H}(B)} \frac{\langle p, f\rangle}{\langle f, f\rangle} h \top \pi(f) = \sum_{f \in \mathscr{H}(B)} \frac{\langle \pi(f), f\rangle}{\langle f, f\rangle} \Pi_f(h \top p),$$

and (1) follows from (1.19). Now, for $k \in \mathcal{H}_{\mathrm{GL}}$ and $q \in Q$

$$\langle \Pi_f(h \top p), k \star q\rangle = \frac{\langle p, f\rangle}{\langle \pi(f), f\rangle} \langle h, k\rangle \langle \pi(f), q\rangle = \langle h \top p, \frac{\langle f, q\rangle}{\langle f, \pi^*(f)\rangle} k \star \pi^*(f)\rangle,$$

whence

$$\Pi_f(k \star q) = \frac{\langle f, q\rangle}{\langle f, \pi^*(f)\rangle} k \star \pi^*(f).$$

This allows us to check the controlledness condition (making the necessary adjustments to take into account that we are operating under the identification of (2.9)):

$$\begin{aligned}\langle \Pi_{f}(\mathbf{H}_t), k \star q\rangle &= \langle \mathbf{H}_t, \Pi_{f}^*(k \star q)\rangle \\ &= \frac{\langle f, q\rangle}{\langle f, \pi^*(f)\rangle}\langle \mathbf{H}_t, k \star \pi^*(f)\rangle \\ &\approx_{\lfloor \rho \rfloor - |k|} \frac{\langle f, q\rangle}{\langle f, \pi^*(f)\rangle}\langle \mathbf{H}_s, \mathbf{X}_{s,t} \star k \star \pi^*(f)\rangle \\ &= \langle \mathbf{H}_t, \Pi_{f}^*(\mathbf{X}_{s,t} \star k \star q)\rangle \\ &= \langle \Pi_{f}(\mathbf{H}_t), \mathbf{X}_{s,t} \star k \star q\rangle\end{aligned}$$

Using Lemma 1.19, we compute

$$\begin{aligned}\langle \Pi_{f}(h \top p), k \star q\rangle &= \frac{\langle p, f\rangle}{\langle \pi(f), f\rangle}\langle h, k\rangle\langle \pi(f), q\rangle \\ &= \frac{\langle p, \pi^*(f)\rangle}{\langle \pi(f), f\rangle}\sum_{g \in \mathscr{H}} \varsigma(g)^{-1} \sum_{l \in \ell(g)} \langle h, g_l\rangle\langle g_l, k\rangle\langle \pi(f), q\rangle \\ &= \langle \pi(f), f\rangle^{-1} \sum_{g \in \mathscr{H}} \varsigma(g)^{-1} \sum_{l \in \ell(g)} \langle h \top p, g_l \star \pi^*(f)\rangle\langle g_l \top \pi(f), k \star q\rangle\end{aligned}$$

and (2) now follows $\int_s^t \Pi_{f}(\mathbf{H})\, \mathrm{d}\mathbf{X} \approx \langle \mathbf{H}_s, \mathbf{X}_{s,t}\rangle$ and bilinearity of the pairing. Assertion (3) is a straightforward consequence of (1) and (2). □

Note that the coefficients $\langle \pi(f), f\rangle$ (which depend in a complicated way on the labelling) disappear in the final version of the integral in coordinates. The definition

$$\langle \pi(f), f\rangle^{-1} \int_0^\cdot \mathbf{H}_{\pi^*(f)}\, \mathrm{d}\mathbf{X}^{\pi(f)} := \int_0^\cdot \Pi_{f}(\mathbf{H})\, \mathrm{d}\mathbf{X}$$

is what makes this happen, and is justified by the presence of this coefficient as a factor to the Davie expansion that follows, while the use of integral notation is justified that this is an integral of the controlled path $\Pi_{f}(\mathbf{H})$. If we were provided with a basis $\mathscr{P}(B)$ of $\mathcal{P}$ (say, extracted from the generating set $\{\pi(f) \mid f \in \mathscr{H}(B)\}$) with dual basis $\mathbb{Q}(B)$ (in the sense that $\langle p_i, q_j\rangle = 0$ if $i \neq j$), in which case we could rewrite (3) as

$$\int_s^t \mathbf{H}\, \mathrm{d}\mathbf{X} = \sum_{\langle p,q\rangle \neq 0} \langle p, q\rangle^{-1} \int_s^t \mathbf{H}_p\, \mathrm{d}\mathbf{X}^q \approx \sum_{\substack{n=1,\ldots,\lfloor \rho \rfloor \\ \langle q_k, p_k\rangle \neq 0}} \left(\prod_{k=1}^n \langle p_k, q_k\rangle\right)^{-1} \mathbf{H}_{q_1 \star \cdots \star q_n; s} \mathbf{X}_{s,t}^{p_1 \top \ldots \top p_n} \tag{2.11}$$

We prefer, however, to provide formulae in the canonical forest coordinates, since this does not mean stating results in terms of a basis which is complicated to extract. Finally, we mention there is a third way of writing the expansion, which combines the $(\star, \top)$-powers of (2.11) and the basis-free formulation of Lemma 2.6:

$$\int_s^t \mathbf{H}\, \mathrm{d}\mathbf{X} \approx \sum_{f_1, \ldots, f_n} \left(\prod_{k=1}^n \varsigma(f_n)\right)^{-1} \mathbf{H}_{\pi^*(f_1) \star \cdots \star \pi^*(f_n); s} \mathbf{X}_{s,t}^{\pi(f_1) \top \cdots \top \pi(f_n)}$$

The following definition extends the one original to [25], for which F is only allowed to take values in $\mathcal{L}(V, W)$, thanks to the natural inclusion $V \hookrightarrow Q(V)$.

Definition 2.7 (RDEs driven by $X^{p \in \mathcal{P}}$)**.** Let $F \in C^\infty(W, \mathcal{L}(Q, W))$ and $y_0 \in W$. We say that $Y : [0, T] \to W$ solves the *rough differential equation* (RDE)

$$\mathrm{d}Y_t = F(Y_t)\, \mathrm{d}\mathbf{X}_t, \quad Y_0 = y_0 \tag{2.12}$$

if there exists $\mathbf{Y} \in \mathscr{D}_{\mathbf{X}}^{\lfloor \rho \rfloor}(W)$ with trace Y such that

$$\mathbf{Y}_t = y_0 + \int_0^t F(\mathbf{Y})\, \mathrm{d}\mathbf{X},$$

where $F(\mathbf{Y})$ is defined according to (2.5) and the integral is defined thanks to Proposition 2.4. ▲

Unlike the case where the equation is driven by the paths X^α, the Davie expansion of the solution to (2.12) requires a bit more care since the map $h \mapsto \langle \mathbf{Y}, h\rangle$ will not be induced by a pre-Lie morphism (see Proposition 1.34). Nonetheless, we can make use of the pre-Lie structure of vector fields to obtain a recursive formula for the coefficients.

In order to keep the formulas compact, we introduce the notation

$$\widetilde{\Delta}_{\top} f = f^1 \otimes f^0,$$

and we refer to Definition 1.20 for the definition of the map $\widetilde{\Delta}_{\top}$. The reason for the 0 index in $\widetilde{\Delta}_{\top}$ is to reflect the fact that this operator is not coassociative and so can only be iterated on the left.

Definition 2.8. Let $F \in \mathcal{L}(\mathcal{Q}(V), \mathfrak{X}(W))$. We define recursively a map $\widehat{F}\colon \mathcal{H}_{\mathrm{GL}}(V) \to \mathfrak{X}(W)$ by

$$\widehat{F}_h := F_{\pi^*(h)} + \sum_{n\geq 1} \frac{1}{n!} \widehat{F}_{(h^1)^{(1)}} \odot \cdots \odot \widehat{F}_{(h^1)^{(n)}} \triangleright F_{\pi^*(h^0)}. \tag{2.13}$$

We introduce the map $\mathbf{F}\colon \mathcal{H}_{\mathrm{GL}}(V) \to \mathfrak{D}(W)$ by

$$\mathbf{F}_h := \sum_{n\geq 1} \frac{1}{n!} \widehat{F}_{h^{(1)}} \odot \cdots \odot \widehat{F}_{h^{(n)}}. \tag{2.14}$$

▲

We note that the map $\widehat{F}$ is *not* a pre-Lie morphism, as e.g.

$$\widehat{F}_{\bullet_\alpha \!\!\!\!\! \raisebox{0.5ex}{$\bullet_\beta$}}(y) = \mathrm{D}F_{\bullet\alpha}(y) F_{\bullet\beta}(y) - F_{\bullet\alpha\bullet\beta}(y) \neq (\widehat{F}_{\bullet\alpha} \triangleright \widehat{F}_{\bullet\beta})(y).$$

However, in the case $F \in \mathcal{L}(V, \mathfrak{X}(W))$, $\widehat{F}$ is the unique pre-Lie morphism extending $\bullet\alpha \mapsto F_\alpha$, hence $\mathbf{F}$ is induced by a pre-Lie morphism and the simpler formula in Proposition 1.34 may be applied. Nonetheless, $\mathbf{F}$ is a morphism for the Grossman–Larson product in both cases.

Lemma 2.9. *The map* $\mathbf{F}$ *defined in* (2.14) *is a Hopf morphism.*

Proof. We begin by noting that by Theorem 1.25, the coalgebra $(\mathfrak{D}(W), \Delta)$ is cofree cocommutative over $\mathfrak{X}(W)$. In particular, given any other cocommutative coalgebra (C, Δ_C) every map $\phi\colon C \to \mathfrak{X}(W)$ extends uniquely to a coalgebra morphism $\Phi\colon C \to \mathfrak{D}(W)$ such that $\pi_{\mathbf{X}(W)}\Phi = \phi$. Moreover,

$$\Phi(c) = \sum_{n\geq 1} \frac{1}{n!} \phi(c_{(1)}) \odot \cdots \odot \phi(c_{(n)}).$$

Applying this to the map $\widehat{F}\colon \mathcal{H}_{\mathrm{GL}} \to \mathfrak{X}(W)$ we immediately see that $\mathbf{F}$ is a coalgebra morphism. We note that in particular $\pi_{\mathfrak{X}}\mathbf{F} = \widehat{F}$. A cofreeness argument, together with the fact that in $\mathfrak{D}(W)$ we have $\pi_{\mathfrak{X}(W)}(G \circledast H) = G \triangleright \pi_{\mathfrak{X}(W)} H$ for any $F, G \in \mathfrak{D}(W)$, yields that in order for $\mathbf{F}$ to be an algebra morphism it is enough that

$$\widehat{F}_{h\star h'} = \mathbf{F}_h \triangleright \widehat{F}_{h'}. \tag{2.15}$$

The proof proceeds in two steps. First, we show that (2.15) holds for $q \in \mathcal{Q}$ and $h \in \mathcal{H}_{\mathrm{GL}}^+$. Indeed, noting that for any $f \in \mathcal{H}_{\mathrm{CK}}^+$ and $p \in \mathcal{P}$ the formula

$$\widetilde{\Delta}_{\mathrm{CK}}(f \top p) = f' \otimes f'' \top p + f \otimes p$$

holds, we see immediately that

$$\begin{aligned}
\langle \widetilde{\Delta}_{\top}(h \star q), f \otimes p \rangle &= \langle h \star q, f \top p \rangle \\
&= \langle h \otimes q, f' \otimes f'' \top p + f \otimes p \rangle \\
&= \langle h, f \rangle \langle q, p \rangle,
\end{aligned}$$

that is, $\widetilde{\Delta}_{\top}(h \star q) = h \otimes q$. It follows immediately that

$$\widehat{F}_{h\star q} = \mathbf{F}_h \triangleright \widehat{F}_q.$$

Second, any $h \in \mathcal{H}_{\mathrm{GL}}^+$ may be written as $h = \hat{h} \star q$ for some $\hat{h} \in \mathcal{H}_{\mathrm{GL}}^+$ with $|\hat{h}| < |h|$ and $q \in \mathcal{Q}$. Taking $h, h' \in \mathcal{H}_{\mathrm{GL}}^+$ and letting $h' = \hat{h}' \star q$ inductively we see that

$$\begin{aligned}
\widehat{F}_{h\star h'} &= \mathbf{F}_{h\star \hat{h}'} \triangleright \widehat{F}_q \\
&= (\mathbf{F}_h \circledast \mathbf{F}_{\hat{h}'}) \triangleright \widehat{F}_q \\
&= \mathbf{F}_h \triangleright (\mathbf{F}_{\hat{h}'} \triangleright \widehat{F}_q) \\
&= \mathbf{F}_h \triangleright \widehat{F}_{h'}.
\end{aligned}$$

□

The following theorem extends Davie expansion to RDEs driven by the collection $X^{p\in\mathcal{P}}$.

Theorem 2.10 (Davie expansion with drifts). *Let $F \in \mathcal{L}(\mathcal{Q}, \mathfrak{X}(W))$. A path $Y : [0,T] \to W$ solves the RDE $\mathrm{d}Y = F(Y)\,\mathrm{d}\mathbf{X}$, $Y_0 = y_0 \in V$ if and only if*

$$Y_{s,t} \approx \sum_{f \in \mathscr{H}} \varsigma(f)^{-1} \widehat{F}_{f_l}(Y_s)\mathbf{X}^{f_l}_{s,t} \tag{2.16}$$

and the map defined by $\langle \mathbf{Y}_t, h\rangle := \widehat{F}_h(Y_t)$ belongs to $\mathscr{D}^{\lfloor\rho\rfloor}_{\mathbf{X}}(W)$.

Proof. Let us first suppose that Y solves (2.12). By evaluating both sides on the empty forest we immediately obtain from Proposition 2.4 that

$$Y_{s,t} \approx \langle F(\mathbf{Y}_s), \mathbf{X}_{s,t}\rangle = \sum_{f \in \mathscr{F}_+} \varsigma(f)^{-1}\langle F(\mathbf{Y}_s), f\rangle \mathbf{X}^{f}_{s,t}.$$

We now show the expression for $\langle \mathbf{Y}_t, f\rangle$ by induction. In the case of $q \in \mathcal{Q}$ we immediately see that

$$\langle \mathbf{Y}_t, q\rangle = \langle F_q(\mathbf{Y}_t), 1\rangle = F_q(Y_t).$$

If $h \in \ker \pi^*$, assume the identity is true for all $h' \in \mathcal{H}_{\mathrm{GL}}$ with $|h'| < |h|$. Noting that $h = \pi^*(h) + h^1 \star \pi^*(h^0)$ we obtain

$$\begin{aligned}
\langle \mathbf{Y}_t, h\rangle &= F_{\pi^*(h)}(Y_t) + \langle \mathbf{Y}_t, h^1 \star \pi^*(h^0)\rangle \\
&= F_{\pi^*(h)}(Y_t) + \left\langle \int F(\mathbf{Y}_t)\,\mathrm{d}\mathbf{X}, h^1 \star \pi^*(h^0)\right\rangle \\
&= F_{\pi^*(h)}(Y_t) + \langle F_{\pi^*(h^0)}(\mathbf{Y}_t), h^1\rangle \\
&= F_{\pi^*(h)}(Y) + \sum_{n\geq 1}^{\lfloor\rho\rfloor-1} \frac{1}{n!}\mathrm{D}^n F_{\pi^*(h^0)}(Y_t)\Big(\mathbf{Y}_{(h^1)^{(1)};t}, \dots, \mathbf{Y}_{(h^1)^{(n)};t}\Big) \\
&= F_{\pi^*(h)}(Y_t) + \sum_{n\geq 1}^{\lfloor\rho\rfloor-1} \frac{1}{n!}\mathrm{D}^n F_{\pi^*(h^0)}(Y_t)\Big(\widehat{F}_{(h^1)^{(1)}}(Y_t), \dots, \widehat{F}_{(h^1)^{(n)}}(Y_t)\Big) \\
&= F_{\pi^*(h)}(Y_t) + \sum_{n\geq 1}^{\lfloor\rho\rfloor-1} \frac{1}{n!}\Big(\widehat{F}_{(h^1)^{(1)}} \odot \cdots \odot \widehat{F}_{(h^1)^{(n)}} \triangleright F_{\pi^*(h^0)}\Big)(Y_t).
\end{aligned}$$

To show the converse, assume that Y has the local expansion in (2.16) and that $\mathbf{Y}$ is a controlled path above it with components given by $\langle \mathbf{Y}_t, h\rangle = \widehat{F}_h(Y_t)$. Fix $n \geq 1$ and $h \in \mathcal{H}^{(n)}_{\mathrm{GL}}$. By Taylor expansion and (2.16) we can easily see that

$$\begin{aligned}
\widehat{F}_h(Y_t) - \widehat{F}_h(Y_s) &\approx_{\lfloor\rho\rfloor-n} \sum_{k=1}^{\lfloor\rho\rfloor-n-1} \frac{1}{k!}\mathrm{D}^k \widehat{F}_h(Y_s) Y^{\otimes k}_{s,t} \\
&\approx_{\lfloor\rho\rfloor-n} \sum_{k=1}^{\lfloor\rho\rfloor-n-1} \frac{1}{k!} \sum_{f_1,\dots,f_k \in \mathcal{F}_+} \left(\prod_{j=1}^{k} \varsigma(f_k)\right)^{-1} \mathrm{D}^k \widehat{F}_h(Y_s)(\widehat{F}_{f_1}(Y_s) \otimes \cdots \otimes \widehat{F}_{f_k}(Y_s))\mathbf{X}^{f_1\cdots f_k}_{s,t} \\
&\approx_{\lfloor\rho\rfloor-n} \sum_{f \in \mathcal{F}_+} \varsigma(f)^{-1} (\mathbf{F}_f \triangleright \widehat{F}_h)(Y_s)\mathbf{X}^{f}_{s,t} \\
&\approx_{\lfloor\rho\rfloor-n} \sum_{f \in \mathcal{F}_+} \varsigma(f)^{-1} \widehat{F}_{f \star h}(Y_s)\mathbf{X}^{f}_{s,t},
\end{aligned}$$

that is, $\mathbf{Y}$ satisfies Definition 2.2. □

Example 2.11 (Davie expansion for $\lfloor\rho\rfloor = 3$). Let us consider the RDE with drifts

$$\mathrm{d}Y_t = F_{\bullet\alpha}(Y)\,\mathrm{d}\mathbf{X}^{\bullet\alpha}_t + \frac{1}{2}F_{\bullet\alpha\bullet\beta}(Y)\,\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\bullet\beta)}_t + \frac{1}{6}F_{\bullet\alpha\bullet\beta\bullet\gamma}(Y)\,\mathrm{d}\mathbf{X}^{\bullet\alpha\bullet\beta\bullet\gamma}_t + F_{\frac{1}{2}\left(\bullet\alpha \overset{\bullet\beta}{\underset{\bullet\gamma}{|}} - \bullet\gamma \overset{\bullet\beta}{\underset{\bullet\alpha}{|}}\right)}(Y)\,\mathrm{d}\mathbf{X}_t^{\pi\left(\bullet\alpha \overset{\bullet\beta}{\underset{\bullet\gamma}{|}}\right)}.$$

Its solution has the following Davie expansion (suppressing Y_s on the right-hand side in the notation):

$$
\begin{aligned}
Y_{s,t} \approx F_{\bullet\alpha}\mathbf{X}^{\bullet\alpha}_{s,t} + (F_{\bullet\alpha} \rhd F_{\bullet\beta} - F_{\bullet\alpha\bullet\beta})\mathbf{X}^{\substack{\bullet\alpha\\ \bullet\beta}}_{s,t} + \frac{1}{2}F_{\bullet\alpha\bullet\beta}\mathbf{X}^{\bullet\alpha\bullet\beta}_{s,t} \\
+ \big((F_{\bullet\gamma} \rhd F_{\bullet\beta} - F_{\bullet\gamma\bullet\beta}) \rhd F_{\bullet\alpha} - F_{\bullet\gamma} \rhd F_{\bullet\alpha\bullet\beta}\big)\mathbf{X}^{\substack{\bullet\gamma\\ \bullet\beta\\ \bullet\alpha}}_{s,t} \\
+ \frac{1}{2}\big((F_{\bullet\beta\bullet\gamma} + F_{\bullet\beta}F_{\bullet\gamma}) \rhd F_{\bullet\alpha} - F_{\bullet\beta} \rhd F_{\bullet\alpha\bullet\gamma} - F_{\bullet\gamma} \rhd F_{\bullet\beta\bullet\alpha}\big)\mathbf{X}^{\substack{\bullet\beta\ \bullet\gamma\\ \bullet\alpha}}_{s,t} \\
+ \left(F_{\frac{1}{2}\left(\bullet\alpha \substack{\bullet\beta\\ \bullet\gamma} - \bullet\gamma \substack{\bullet\beta\\ \bullet\alpha}\right)} + F_{\bullet\gamma} \rhd F_{\bullet\alpha\bullet\beta}\right)\mathbf{X}^{\bullet\alpha \substack{\bullet\gamma\\ \bullet\beta}}_{s,t} + \frac{1}{6}F_{\bullet\alpha\bullet\beta\bullet\gamma}\mathbf{X}^{\bullet\alpha\bullet\beta\bullet\gamma}_{s,t}. \quad \blacklozenge
\end{aligned}
$$

We come to one of our main results, which is intimately linked to the so-called Davie expansion [12] of our extended type of RDE; such expansions in the case of $F \in C^\infty(W, \mathcal{L}(V, W))$ were already considered in [26]. The purpose of an Itô formula is to express functionals of a path, in our case smooth functions of RDE solutions, as integrals; see Example 2.23 below for the link with Itô's celebrated formula for continuous semimartingales. A type of Itô formula for branched rough paths that depends on additional structure is to be found in [31, Ch. 5], see Remark 2.15 below. We continue to implicitly use the identification (2.9).

Theorem 2.12 (Itô formula for RDEs with drifts). *Let $\varphi \colon W \to U$ be a function of class $C^{\lfloor\rho\rfloor}$. Then the controlled path $\varphi(Y)$ satisfies*

$$\varphi(Y_t) = \varphi(y_0) + \int_0^t \mathbf{F}\varphi(Y_u)\,\mathrm{d}\mathbf{X}_u.$$

Proof. It is enough to check that

$$\varphi(Y_t) - \varphi(Y_s) \approx \langle \mathbf{F}\varphi(Y_s), \mathbf{X}_{s,t}\rangle.$$

By Taylor's formula, we have

$$\varphi(Y_t) - \varphi(Y_s) \approx \sum_{n=1}^{\lfloor\rho\rfloor} \frac{1}{n!}\mathrm{D}^n\varphi(Y_s)Y_{s,t}^{\otimes n}.$$

Since Y solves (2.12), by Theorem 2.10 we may substitute the expansion in (2.16) for $Y_{s,t}$ to obtain

$$
\begin{aligned}
\varphi(Y_t) - \varphi(Y_s) &\approx \sum_{n=1}^{\lfloor\rho\rfloor}\frac{1}{n!}\sum_{f_1,\dots f_n \in \mathcal{F}_+} \mathrm{D}^n\varphi(Y_s)\Big(\widehat{F}_{f_1}(Y_s) \otimes \cdots \otimes \widehat{F}_{f_n}(Y_s)\Big)\mathbf{X}^{f_1\cdots f_n}_{s,t} \\
&\approx \sum_{\substack{f \in \mathcal{F}_+\\ |f| \le \lfloor\rho\rfloor}} \left(\sum_{n=1}^{|f|}\frac{1}{n!}\mathrm{D}^n\varphi(Y_s)\Big(\widehat{F}_{f^{(1)}} \otimes \cdots \otimes \widehat{F}_{f^{(n)}}\Big)\right)\mathbf{X}^{f}_{s,t} \\
&= \sum_{\substack{f \in \mathcal{F}_+\\ |f| \le \lfloor\rho\rfloor}} \mathbf{F}_f\varphi(Y_s)\mathbf{X}^f_{s,t} \\
&= \big\langle \mathbf{F}\varphi(Y_s), \mathbf{X}_{s,t}\big\rangle.
\end{aligned}
$$

□

Remark 2.13 (Itô formula for controlled paths). In the extended setting of (Q, W)-valued controlled integrands, it could be more generally stated that, given any $\mathbf{H} \in \mathscr{D}_{\mathbf{X}}(W)$,

$$\varphi(H)_{s,t} = \sum_{f \in \mathcal{H}} \varsigma(f)^{-1} \sum_{l \in \ell(f)} \int_s^t \varphi(\mathbf{H})_{\pi^*(f_l)}\mathrm{d}\mathbf{X}^{\pi(f_l)}$$

at the cost of requiring $\mathbf{H}$ to have "an extra degree of controlledness by primitiveness". This is not an issue with RDE solutions, which have enough Gubinelli derivatives. ♦

Example 2.14 (Itô formula for $\lfloor\rho\rfloor = 3$). Continuing with Example 2.11, we see that for any smooth function $\varphi \in C^\infty(W)$,

$$
\begin{aligned}
\varphi(Y_t) - \varphi(Y_s) = &\int_s^t \mathbf{F}_{\bullet\gamma}\varphi(Y_u)\mathrm{d}\mathbf{X}^{\bullet\gamma}_u \\
&+ \frac{1}{2}\int_s^t \mathbf{F}_{\bullet\alpha\bullet\beta}\varphi(Y_u)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\bullet\beta)}_u + \frac{1}{6}\int_s^t \mathbf{F}_{\bullet\alpha\bullet\beta\bullet\gamma}\varphi(Y_u)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\bullet\beta\bullet\gamma)}_u \\
&+ \int_s^t \mathbf{F}_{\frac{1}{2}\left(\bullet\alpha \substack{\bullet\beta\\ \bullet\gamma} - \bullet\gamma \substack{\bullet\beta\\ \bullet\alpha}\right)}\varphi(Y_u)\mathrm{d}\mathbf{X}_u^{\pi\left(\bullet\alpha \substack{\bullet\beta\\ \bullet\gamma}\right)}
\end{aligned}
$$

where in each integral, the sum is intended over labels, not labelled forests. Here, the differential operators are given explicitly by

$$
\begin{aligned}
\mathbf{F}_{\bullet\gamma} &= F_{\bullet\gamma}\\
\mathbf{F}_{\bullet\alpha\bullet\beta} &= F_{\bullet\alpha\bullet\beta} + F_{\bullet\alpha} \odot F_{\bullet\beta}\\
\mathbf{F}_{\bullet\alpha\bullet\beta\bullet\gamma} &= F_{\bullet\alpha\bullet\beta\bullet\gamma} + F_{\bullet\alpha} \odot F_{\bullet\beta\bullet\gamma} + F_{\bullet\beta} \odot F_{\bullet\alpha\bullet\gamma} + F_{\bullet\gamma} \odot F_{\bullet\alpha\bullet\beta} + F_{\bullet\alpha} \odot F_{\bullet\beta} \odot F_{\bullet\gamma}\\
\mathbf{F}_{\frac{1}{2}\left(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}}-\bullet\gamma\overset{\bullet\beta}{\underset{\bullet\alpha}{\vert}}\right)} &= F_{\frac{1}{2}\left(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}}-\bullet\gamma\overset{\bullet\beta}{\underset{\bullet\alpha}{\vert}}\right)} + F_{\bullet\alpha} \odot F_{\bullet\beta} \rhd F_{\bullet\gamma} - F_{\bullet\gamma} \odot F_{\bullet\beta} \rhd F_{\bullet\alpha} + F_{\bullet\gamma} \odot F_{\bullet\alpha\bullet\beta} - F_{\bullet\alpha} \odot F_{\bullet\gamma\bullet\beta},
\end{aligned}
$$

and we recall Lemma 1.33 for the evaluation of the symmetric products. We further Davie-expand each integral:

$$
\begin{aligned}
\int_s^t \mathbf{F}_\gamma \varphi(Y)\mathrm{d}\mathbf{X}^\gamma &\approx \mathbf{F}_{\bullet\gamma}\varphi(Y_s)\mathbf{X}^{\bullet\gamma}_{s,t} + \mathbf{F}_{\bullet\alpha\star\bullet\beta}\varphi(Y_s)\mathbf{X}^{\bullet\alpha\top\bullet\beta}_{s,t} + \frac{1}{2}\mathbf{F}_{\bullet\alpha\bullet\beta\star\bullet\gamma}\varphi(Y_s)\mathbf{X}^{\bullet\alpha\bullet\beta\top\bullet\gamma}_{s,t}\\
\int_s^t \mathbf{F}_{\bullet\alpha\bullet\beta}\varphi(Y)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\bullet\beta)} &\approx \mathbf{F}_{\bullet\alpha\bullet\beta}\varphi(Y_s)\mathbf{X}^{\pi(\bullet\alpha\bullet\beta)}_{s,t} + \mathbf{F}_{\bullet\gamma\star\bullet\alpha\bullet\beta}\varphi(Y_s)\mathbf{X}^{\bullet\gamma\top\pi(\bullet\alpha\bullet\beta)}_{s,t}\\
\int_s^t \mathbf{F}_{\bullet\alpha\bullet\beta\bullet\gamma}\varphi(Y)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\bullet\beta\bullet\gamma)} &\approx \mathbf{F}_{\bullet\alpha\bullet\beta\bullet\gamma}\varphi(Y_s)\mathbf{X}^{\pi(\bullet\alpha\bullet\beta\bullet\gamma)}_{s,t}\\
\int_s^t \mathbf{F}_{\frac{1}{2}(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}}-\bullet\gamma\overset{\bullet\beta}{\underset{\bullet\alpha}{\vert}})}\varphi(Y)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}})} &= \int_s^t \mathbf{F}_{\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}}}\varphi(Y)\mathrm{d}\mathbf{X}^{\pi(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}})} \approx \mathbf{F}_{\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}}}\varphi(Y_s)\mathbf{X}^{\pi(\bullet\alpha\overset{\bullet\beta}{\underset{\bullet\gamma}{\vert}})}_{s,t}
\end{aligned}
$$

where we were able to simplify the antisymmetrisation in the last integral due to the fact that there are no higher derivatives (if $\lfloor\rho\rfloor$ were 4 or higher, this might not be possible, as is revealed by an inspection of the proof of Lemma 2.6). The reader may check that extracting a dual pair of bases of $\mathcal{P}$ and $\mathcal{Q}$ (which is straightforward up to degree 3 by working with symmetric/antisymmetric tensors, see Example 1.12, Example 1.23) and writing the alternative expansion (2.11) one obtains the same result. ♦

Remark 2.15 (Kelly's approach via bracket extensions)**.** As already mentioned, a version of Theorem 2.12 was already obtain by Kelly [31, Theorem 5.3.11] in the case of $F \in C^\infty(W, \mathcal{L}(V, W))$. The main difference between the two approaches is that we leverage an algebraic property of $\mathcal{H}_{\mathrm{GL}}$ — its freeness over $\mathcal{Q}$ — in order to write a change-of-variable formula that does not depend on anything that is not already contained in the original branched rough path $\mathbf{X}$. On the other hand, Kelly introduces the notion of *bracket extension*, a branched rough path whose trace takes its values in an enlarged space, which in our notation would be $\mathcal{H}(V)$: that is, while the original $\mathbf{X}$ is indexed by forests labelled with $\alpha \in V$, the extended $\widehat{\mathbf{X}}$ is indexed by forests *indexed with by forests labelled with* $\alpha \in V$. These new labels $\mathscr{f}$ actually represent the primitives $\pi(\mathscr{f})$, so that $\widehat{\mathbf{X}}$ is meant to encode branched integrals indexed by primitives (which in our approach is handled by (2.7)). Beyond extending $\mathbf{X}$, $\widehat{\mathbf{X}}$ is required to satisfy certain "bracket polynomial" relations from which it follows that a result analogous to the theorem above with integration against $\mathrm{d}\widehat{\mathbf{X}}^{(\mathscr{f}_1\cdots\mathscr{f}_n)}$. The bracket extension $\widehat{\mathbf{X}}$, however, is highly non unique, and only shown to exist using non-constructive methods, which have been avoided here. Another point worth noting is that, in the bracket extension approach, $\mathbf{X}$-controlled paths can be expanded as integrals against $\widehat{\mathbf{X}}$; however, $\widehat{\mathbf{X}}$-controlled paths would need a further iteration of the bracket extension to be represented as integrals; this, too, is avoided in our approach, since we do not need to extend $\mathbf{X}$, only re-organise its internal structure. We remark that Kelly's change of variable formula takes a slightly simplified expression in coordinates than ours; in particular the $\pi(\mathscr{f})$ that indexes the integrator for us is a separate, newly added label for him, and $\mathscr{f}$ replaces $\pi^*(\mathscr{f})$ as subscript of the integrand: this substitution, which appears not to be possible in our case (see a special case where it is, in Example 2.14), jointly with the fact that F is of simplified form, makes it possible for him to sum over trees.

The two approaches can be reconciled by observing that setting, recursively, for $\mathscr{f} \in \mathscr{H}(\mathscr{H}(V))$ and $\mathscr{g} \in \mathscr{H}(V^*)$

$$
\langle [\mathscr{f}]_{(\mathscr{g})}, \widehat{\mathbf{X}}\rangle = \langle \mathscr{f} \top \pi(\mathscr{g}), \widehat{\mathbf{X}}\rangle \tag{2.17}
$$

defines a *canonical* bracket extension: the bracket relations to be checked read

$$
\langle \pi(\mathscr{g}), \widehat{\mathbf{X}}\rangle = \langle \mathscr{g} - \mathscr{g}' \top \pi(\mathscr{g}''), \widehat{\mathbf{X}}\rangle, \qquad \mathscr{g} \in \mathscr{H}(V)
$$

which is the content of Proposition 1.10. One must also check that (2.17) defines a branched rough path, which is implied by the fact that the algebra morphism $\top \colon \mathcal{H}_{\mathrm{CK}}(\mathcal{H}(V^*)) \to \mathcal{H}_{\mathrm{CK}}(V^*)$ defined recursively by

$$
[\mathscr{f}]_{(\mathscr{g})} \mapsto \top(\mathscr{f}) \top \pi(\mathscr{g}), \quad \mathscr{f} \in \mathscr{H}(\mathscr{H}(V^*)),\ \mathscr{g} \in \mathscr{H}(V^*)
$$

is a Hopf morphism. This is again checked thanks to the cocycle property (1.10):

$$\begin{aligned}\widetilde{\Delta}_{\mathrm{CK}}\top([\mathcal{f}]_{(g)}) &= \top(\mathcal{f})\otimes\pi(g) + \top(\mathcal{f})'\otimes(\top(\mathcal{f})''\top\pi(g))\\ &= \top(\mathcal{f})\otimes\pi(g) + \top(\mathcal{f}')\otimes(\top(\mathcal{f}'')\top\pi(g))\\ &= \top(\mathcal{f})\otimes\top(\bullet^{(\pi(g))}) + \top(\mathcal{f}')\otimes\top([\mathcal{f}'']_{(\pi(g))})\\ &= (\top\otimes\top)\widetilde{\Delta}_{\mathrm{CK}}([\mathcal{f}]_{(g)})\end{aligned}$$

where in the second step we have argued inductively since $\mathcal{f}$ is of lower degree in the coradical filtration, jointly with the observation that Δ_{CK} is an algebra morphism, so that (taking $\mathcal{f} = t_1\cdots t_n$)

$$\widetilde{\Delta}_{\mathrm{CK}}\top(t_1\cdots t_n) = \widetilde{\Delta}_{\mathrm{CK}}(\top(t_1)\cdots\top(t_n)) = \top^{\otimes 2}\widetilde{\Delta}_{\mathrm{CK}}(t_1)\cdots\widetilde{\Delta}_{\mathrm{CK}}(t_n)$$

The fact that our approach and Kelly's can be reconciled in this way can be helpful when bootstrapping statements in the extended setting that are already known in the classical branched setting: for example, local existence and uniqueness of the RDEs driven by $X^{p\in\mathcal{P}}$ can be inferred from the fact that these are RDEs (in the ordinary branched sense) driven by this canonical bracket extension. ♦

2.2. **A criterion for quasi-geometricity and the simple change-of-variables formula.** We begin by motivating this section by considering the restriction of Theorem 2.12 to the "trivial RDE", from which we obtain the Itô formula for functions of X. The proof is a straightforward application of the recursive definition of $\mathbf{F}$ in Theorem 2.12. We recall the notation $\mathcal{V}(\gamma_1,\ldots,\gamma_n) := \bullet_{\gamma_1}\cdots\bullet_{\gamma_n}$.

Corollary 2.16 (Simple change of variable formula)**.** *Let* $\mathbf{X}\in\mathscr{C}^{\rho}([0,T],V)$*,* $W=Q$ *and* $F(y)\equiv\mathrm{id}_Q$ *in* (2.12)*. Then, for* $\varphi\in C^{\lfloor\rho\rfloor+1}(Q)$ *we have*

$$\mathbf{F}_h\varphi(Y) = \sum_{n\geq 1}\frac{1}{n!}\partial_{\pi^*(h^{(1)})\ldots\pi^*(h^{(n)})}\varphi(X)$$

for any $h\in\mathcal{H}^{\lfloor\rho\rfloor}_{\mathrm{GL}}$*. If, furthermore,* $W=V$*,* $\mathbf{F}_{\mathcal{V}(\gamma_1,\ldots,\gamma_n)}\varphi(Y) = \partial_{\gamma_1\ldots\gamma_n}\varphi(X)$ *and* $\overline{F}_h\varphi=0$ *otherwise, so that*

$$\varphi(X_t)-\varphi(X_s) = \sum_{n=1}^{\lfloor\rho\rfloor}\frac{1}{n!}\int_s^t\partial_{\gamma_1\ldots\gamma_n}\varphi(X_u)\,\mathrm{d}\mathbf{X}_u^{\pi(\mathcal{V}(\gamma_1,\ldots,\gamma_n))} \tag{2.18}$$

The use of the term *simple* is due to Kelly [31], who first pointed out that the terms of a bracket extension needed to make sense of (2.18) were fewer than those needed for the general RDE case; the same formula is available in the quasi-geometric case [3, Theorem 4.20]. There is a setting in which the general change-of-variables formula of Theorem 2.12 takes on the simpler form (2.18). Hoffman's quasi-shuffle algebra, introduced in [27] to study renormalisation of multiple zeta values, has been connected with rough path theory [3] and stochastic differential equations [13]. We refer to [48] for an introduction to the shuffle algebra, and to [27,29] for an introduction to the quasi-shuffle algebra, and assume familiarity with these notions. In the spirit of this paper we view these as covariant functors (see [16, 43] for a similar perspective on quasi-shuffle algebras and their deformations)

$$\sqcup\!\sqcup(U) = (\bigotimes(U),\sqcup\!\sqcup,\Delta_\otimes),\qquad \widetilde{\sqcup\!\sqcup}(U) = (\bigotimes(\bigodot(U)),\widetilde{\sqcup\!\sqcup},\Delta_\otimes) \tag{2.19}$$

We are considering a free commutative bracket for our quasi-shuffle algebra, i.e. $[x,y]$ in the notation of [27] is given by $x\odot y$ here ($x,y\in\bigodot(U)$).

Both structures have indeed a very interpretation even in the branched rough path setting, as it is noted in the literature. A first direct link is provided by the arborifications maps, see [6, Section 2.3] and [26, Section 4.1].

Proposition 2.17 ((Contracting-)arborification maps)**.** *There exist unique surjective Hopf algebra morphisms* $\mathfrak{a}\colon\mathcal{H}_{\mathrm{CK}}(U)\to\sqcup\!\sqcup(U)$*,* $\widetilde{\mathfrak{a}}_\odot\colon\mathcal{H}_{\mathrm{CK}}(\bigodot U)\to\widetilde{\sqcup\!\sqcup}(U)$ *defined recursively by the conditions*

$$\mathfrak{a}([\mathcal{f}]_\gamma) := \mathfrak{a}(\mathcal{f})\otimes\gamma,\qquad \widetilde{\mathfrak{a}}_\odot([\mathcal{f}]_{\gamma_1\odot\cdots\odot\gamma_n}) := \widetilde{\mathfrak{a}}_\odot(\mathcal{f})\otimes(\gamma_1\odot\cdots\odot\gamma_n)\,.$$

Following the literature, we call them the arborification *and* contracting-arborification *maps respectively. These maps are natural and are left inverses to the natural injective coalgebra maps* $\iota\colon\sqcup\!\sqcup(U)\to\mathcal{H}_{\mathrm{CK}}(U)$*,* $\widetilde{\iota}_\odot\colon\widetilde{\sqcup\!\sqcup}(U)\hookrightarrow\mathcal{H}_{\mathrm{CK}}(\bigodot U)$ *defined respectively by the conditions*

$$\iota(\gamma) = \bullet_\gamma\qquad \widetilde{\iota}_\odot(\gamma_1\odot\cdots\odot\gamma_n) = \bullet_{\gamma_1\odot\cdots\odot\gamma_n}\,.$$

and extended to ladder trees according by putting the decoration of the root as last letter.

This result shows how $\sqcup\!\sqcup(U)$ and $\widetilde{\sqcup\!\sqcup}(U)$ can be obtained as natural quotients of the Connes-Kreimer Hopf algebra over U and $\bigodot(U)$ respectively. However, in order to relate quasi-geometric and branched rough paths, we need to represent $\widetilde{\sqcup\!\sqcup}(U)$ as a quotient of $\mathcal{H}_{\mathrm{CK}}(U)$, not $\mathcal{H}_{\mathrm{CK}}(\bigodot U)$: for this reason, from now on and unless otherwise mentioned, we revert to working over U.

We denote by $\mathcal{Q}_{\mathscr{r}}$ the subspace of $\mathcal{F}$ spanned by elements of the form $\mathscr{r}(\gamma_1,\ldots,\gamma_n)$, and as remarked in Example 1.23, $\mathcal{Q}_r \subset \mathcal{Q}$; similarly, denote $\mathcal{P}_{\mathscr{r}} := \pi(\mathcal{Q}_{\mathscr{r}})$. Note that there is an obvious, natural isomorphism $\mathcal{Q}_r(U) \cong \bigodot(U)$. Using the definition of π, we have the following recursive formula:

$$\pi(\mathscr{r}(\gamma_1,\ldots,\gamma_n)) = \mathscr{r}(\gamma_1,\ldots,\gamma_n) - \sum_{I\sqcup J=\{\gamma_1,\ldots,\gamma_n\}} \mathscr{r}(\gamma_I) \top \pi(\mathscr{r}(\gamma_J))\,, \tag{2.20}$$

where I and J are two non-empty subsets partitioning $\{\gamma_1,\ldots,\gamma_n\}$ and $\mathscr{r}(\gamma_I)$, $\mathscr{r}(\gamma_J)$ are the associated forests obtained decorating the roots with elements indexed by I and J respectively.

Lemma 2.18. *$\pi\colon \mathcal{Q}_{\mathscr{r}} \to \mathcal{P}_{\mathscr{r}}$ is an isomorphism.*

Proof. (2.20) implies that, given a basis B of U, $\{\pi(\mathscr{r}(\gamma_1,\ldots,\gamma_n)) \mid \gamma_1,\ldots,\gamma_n \in B\}$ is a basis of $\mathcal{P}_{\mathscr{r}}$. Indeed, $\{\mathscr{r}(\gamma_1,\ldots,\gamma_n) \mid \gamma_1,\ldots,\gamma_n \in B\}$ is a basis of $\mathcal{Q}_{\mathscr{r}}$, and no non-trivial linear combination of the terms of the form $\mathscr{r}(\gamma_I) \top \pi(\mathscr{r}(\gamma_J))$ can belong to $\mathcal{Q}_{\mathscr{r}}$ since all forests involved have at least one edge. □

We are now in a position to define a version of the quasi-arborification on $\mathcal{H}_{\mathrm{CK}}(U)$ instead of on $\mathcal{H}_{\mathrm{CK}}(\bigodot(U))$ and to construct a natural right inverse to it.

Theorem 2.19 (Intrinsic quasi-arborification map)**.** *There exists a surjective Hopf algebra morphism, the* intrinsic quasi-arborification *map* $\widetilde{\mathfrak{a}}\colon \mathcal{H}_{\mathrm{CK}}(U) \to \widetilde{\sqcup\!\sqcup}(U)$*, defined recursively by the condition*

$$\widetilde{\mathfrak{a}}([f]_\gamma) := \widetilde{\mathfrak{a}}(f) \otimes \gamma\,.$$

$\widetilde{\mathfrak{a}}$ is natural and is left inverse to the natural coalgebra map $\widetilde{\iota}\colon \widetilde{\sqcup\!\sqcup}(U) \hookrightarrow \mathcal{H}_{\mathrm{CK}}(U)$ *defined by* $\gamma_1 \odot \cdots \odot \gamma_n \mapsto \pi(\mathscr{r}(\gamma_1,\ldots,\gamma_n))$ *and extended by* $s_1 \otimes \cdots \otimes s_n \mapsto s_1 \top \ldots \top s_n$ *for* $s_1,\ldots,s_n \in \bigodot(U)$*.*

Proof. Existence and uniqueness of an algebra morphism is clear from the defining property and the fact that algebra morphism property forces $\widetilde{\mathfrak{a}}(t_1\cdots t_n) = \widetilde{\mathfrak{a}}(t_1)\,\widetilde{\sqcup\!\sqcup}\ldots\widetilde{\sqcup\!\sqcup}\,\widetilde{\mathfrak{a}}(t_n)$. $\widetilde{\mathfrak{a}}$ is a coalgebra morphism by the universal property of the Connes-Kreimer Hopf algebra (see for example [15, Theorem 3]), since $\widetilde{\mathfrak{a}} \circ \mathcal{B}_+^\gamma = (-\otimes\gamma)\circ\widetilde{\mathfrak{a}}$. We now make use of the following property of quasi-shuffles, which can be verified directly from the definition of $\widetilde{\sqcup\!\sqcup}$: if $w_1,\ldots,w_n \in \widetilde{\sqcup\!\sqcup}(U)$ are words (i.e. elementary tensors) at least one of which has length strictly greater than 1 (where length refers to length in the tensor algebra, disregarding weight of the letters in $\bigodot(U)$), then $w_1\,\widetilde{\sqcup\!\sqcup}\ldots\widetilde{\sqcup\!\sqcup}\,w_n$ is a linear combination of words of length strictly greater than 1. This implies that

$$\pi_1 \circ \widetilde{\mathfrak{a}}(f) = 0 \text{ if } f \text{ is a forest containing at least one edge} \tag{2.21}$$

We then have, by Proposition 1.5 (and using notation therein), for $p_1,\ldots,p_n \in \mathcal{P}$

$$\begin{aligned}\widetilde{\mathfrak{a}}(p_1\top\ldots\top p_n) &= \sum_{\substack{n_1+\ldots+n_k=n\\ n_1,\ldots,n_k\geq 1}} \pi_1\circ\widetilde{\mathfrak{a}}(p_1\top\ldots\top p_{n_1})\otimes\cdots\otimes\pi_1\circ\widetilde{\mathfrak{a}}(p_{n^{k-1}+1}\top\ldots\top p_n)\\ &= \widetilde{\mathfrak{a}}(p_1)\otimes\cdots\otimes\widetilde{\mathfrak{a}}(p_n)\end{aligned}$$

once again because $\top$ always introduces edges, and thus length strictly greater than 1 in the corresponding images through $\widetilde{\mathfrak{a}}$. Finally, using (2.20)

$$\begin{aligned}\widetilde{\mathfrak{a}}\circ\widetilde{\iota}(\gamma_1\odot\cdots\odot\gamma_n) &= \widetilde{\mathfrak{a}}\circ\pi(\mathscr{r}(\gamma_1,\ldots,\gamma_n))\\ &= \pi_1\circ\widetilde{\mathfrak{a}}\circ\pi(\mathscr{r}(\gamma_1,\ldots,\gamma_n))\\ &= \pi_1\circ\widetilde{\mathfrak{a}}\circ\mathscr{r}(\gamma_1,\ldots,\gamma_n)\\ &= \pi_1(\gamma_1\,\widetilde{\sqcup\!\sqcup}\cdots\widetilde{\sqcup\!\sqcup}\,\gamma_n)\\ &= \gamma_1\odot\cdots\odot\gamma_n\end{aligned}$$

and since $\widetilde{\mathfrak{a}}\circ\widetilde{\iota}(s_1\otimes\cdots\otimes s_n) = \widetilde{\mathfrak{a}}(\widetilde{\iota}(s_1)\top\ldots\top\widetilde{\iota}(s_n)) = s_1\otimes\cdots\otimes s_n$ (as $\widetilde{\iota}$ maps $\bigodot(U)$ to $\mathcal{P}_{\mathscr{r}}$), we conclude that $\widetilde{\mathfrak{a}}\circ\widetilde{\iota} = \mathrm{id}$. □

We note that the intrinsic arborification map, despite admitting a right inverse that is defined in terms of $\top$, is not itself defined in terms of the $\mathbf{B}_\infty$-structure of $\mathcal{H}_{\mathrm{CK}}$. The fact that $\sqcup\!\sqcup$ and $\widetilde{\sqcup\!\sqcup}$ are quotients of $\mathcal{H}_{\mathrm{CK}}$ immediately implies that geometric and quasi-geometric rough paths — rough paths defined on these Hopf algebras in total analogy to Definition 2.1 — can be viewed as branched rough paths.

Conversely, one may ask whether a branched rough path comes from a (quasi-)geometric one in this manner. Before we answer this question, we provide some intuition as to what the answer entails from the point of view of integration. Just in (2.7) and the discussion below, we showed that the $\top$ operation could be thought of as the algebraic counterpart to integration, we now show how π can be thought of the algebraic counterpart to a generalised integration-by-parts formula. From Proposition 1.10 it follows that for $f = [g_1]_{\gamma_1} \cdots [g_n]_{\gamma_n}$, the product of the "integrals" (in the interpretation (2.2)) $\langle [g]_{\gamma_k}, \mathbf{X}^{[g]_{\gamma_k}} \rangle$ can be written as a sum of "integrals" as

$$\begin{aligned}\langle f, \mathbf{X}_{s,t} \rangle &= \sum_{(f)} \langle f_{(1)} \top \pi(f_{(2)}), \mathbf{X}_{s,t} \rangle \\ &= \sum_{k=1}^{n} \Big\langle \big[[g_1]_{\gamma_1} \cdots [g_{k-1}]_{\gamma_{k-1}} \cdot g_k \cdot [g_{k+1}]_{\gamma_{k+1}} \cdots [g_n]_{\gamma_n} \big]_{\gamma_k}, \mathbf{X}_{s,t} \Big\rangle \\ &\quad + \sum_{\substack{I \sqcup J = [n] \\ |I| \geq 2}} \Big\langle \prod_{i \in I} g_i \cdot \prod_{j \in J} [g_j]_{\gamma_j} \Big) \top \pi\Big(\prod_{i \in I} \bullet_{\gamma_i} \Big) \Big\rangle \\ &\quad + \text{terms involving cuts above the roots of the forests } g_k\end{aligned} \tag{2.22}$$

Truncating the last expression after its first line sign is the classical integration-by-parts formula that holds for smooth paths, which continues to hold for geometric rough paths (which, despite their roughness, obey the same laws as ordinary calculus). Stochastic examples of geometric rough paths are given by Stratonovich integration of a semimartingale. For quasi geometric rough paths, instead, one must stop after the second line: this means that "interactions" involving n of the integrators ("n-variations") are non-negligible. All $2 < \rho$-branched rough paths are automatically quasi-geometric (this will be obvious from the next result), the Itô integral being the prime stochastic example. The correct integration-by-parts for general branched rough path, however, involves further terms (third line), "interactions between integrals" not captured by those mentioned previously. We note that a characterisation of $\bigodot(V)$-valued branched rough paths that are quasi-geometric was provided in [14, Theorem 3.1]. This is different from a characterisation of V-valued branched rough paths that are quasi geometric which does not require the extension of the rough path to $\bigodot(V)$ to be checked, and which is only possible to state thanks to our intrinsic arborification map above.

Corollary 2.20 (Characterisation of (quasi-)geometric rough paths)**.**

- *A V-valued branched rough path $\mathbf{X}$ is geometric (i.e. lies in the image of $\mathfrak{a}^*$) if and only if, for $p_1, \ldots, p_n \in \mathcal{P}$, $\langle p_1 \top \ldots \top p_n, \mathbf{X} \rangle = 0$ if $\exists k$ $p_k \in \mathcal{P}/\iota(V^*)$ (where the quotient is identified with the direct complement of $\iota(V^*)$ in $\mathcal{P}$ generated by forests that have more than one vertex);*
- *A V-valued branched rough path $\mathbf{X}$ is quasi-geometric (i.e. lies in the image of $\widetilde{\mathfrak{a}}^*$) if and only if, for $p_1, \ldots, p_n \in \mathcal{P}$, $\langle p_1 \top \ldots \top p_n, \mathbf{X} \rangle = 0$ if $\exists k$ $p_k \in \mathcal{P}/\mathcal{P}_{\mathscr{r}}$ (where the quotient is identified with the direct complement of $\mathcal{P}_{\mathscr{r}}$ in $\mathcal{P}$ generated by forests that have an edge).*

Proof. We only prove the second assertion; the first is proved similarly (using Proposition 2.17). Note that if $\mathbf{X} = \widetilde{\mathfrak{a}}^*(\mathbf{Z}) = \widetilde{\mathfrak{a}}^* \circ \widetilde{\iota}^* \circ \widetilde{\mathfrak{a}}^*(\mathbf{Z})$, then $\mathbf{X} = \widetilde{\mathfrak{a}}^* \circ \widetilde{\iota}^*(\mathbf{X})$. By Theorem 1.8, the elements $p_1 \top \ldots \top p_n$ with $p_k \in \mathcal{P}$ span $\mathcal{H}_{\mathrm{CK}}$, so by non-degeneracy of the pairing Theorem 1.18, $\mathbf{X} = \widetilde{\mathfrak{a}}^* \circ \widetilde{\iota}^*(\mathbf{X})$ if and only if

$$\langle p_1 \top \ldots \top p_n, \mathbf{X} \rangle = \langle \widetilde{\iota} \circ \widetilde{\mathfrak{a}}(p_1 \top \ldots \top p_n), \mathbf{X} \rangle = \langle \widetilde{\iota} \circ \pi_1 \circ \widetilde{\mathfrak{a}}(p_1) \top \ldots \top \widetilde{\iota} \circ \pi_1 \circ \widetilde{\mathfrak{a}}(p_n), \mathbf{X} \rangle$$

and the assertion now follows from (2.21). □

We now state the restriction of Theorem 2.12 to the case of quasi-geometric rough paths (in the sense of Corollary 2.20). We use bold letters to denote tuples of elements in a basis B of V.

Corollary 2.21 (Quasi geometric change-of-variable formula for RDEs)**.** *In the notation of Theorem 2.12, if $\mathbf{X}$ is quasi-geometric the formula reduces to*

$$\varphi(Y)_{s,t} = \sum_{\substack{|\boldsymbol{\gamma}^1|,\ldots,|\boldsymbol{\gamma}^n| \geq 1 \\ |\boldsymbol{\gamma}^1|,\ldots,|\boldsymbol{\gamma}^n| \leq \lfloor \rho \rfloor}} \frac{1}{n! |\boldsymbol{\gamma}^1|! \cdots |\boldsymbol{\gamma}^n|!} \int_s^t \partial_{k_1 \ldots k_n} \varphi(\mathbf{Y}) F^{k_1}_{\mathscr{r}(\boldsymbol{\gamma}^1)}(\mathbf{Y}) \cdots F^{k_n}_{\mathscr{r}(\boldsymbol{\gamma}^n)}(\mathbf{Y})\, \mathrm{d}\mathbf{X}^{\pi(\mathscr{r}(\boldsymbol{\gamma}^1 \ldots \boldsymbol{\gamma}^n))}.$$

We consider two special cases of this: first, if $F \in C^\infty(W, \mathcal{L}(V, W))$,

$$\varphi(Y)_{s,t} = \sum_{\substack{\gamma_1, \ldots, \gamma_n \\ 1 \leq n \leq \lfloor \rho \rfloor}} \frac{1}{n!} \int_s^t \partial_{k_1 \ldots k_n} \varphi(\mathbf{Y}) F^{k_1}_{\gamma_1}(\mathbf{Y}) \cdots F^{k_1}_{\gamma_1}(\mathbf{Y})\, \mathrm{d}\mathbf{X}^{\pi(\mathscr{r}(\gamma_1 \ldots \gamma_n))}.$$

Secondly, if W, F are as in Corollary 2.16,

$$\varphi(X)_{s,t} = \sum_{\substack{|\gamma^1|,\ldots,|\gamma^n|\geq 1\\ |\gamma^1|,\ldots,|\gamma^n|\leq\lfloor\rho\rfloor}} \frac{1}{n!|\gamma^1|!\cdots|\gamma^n|!}\int_s^t \partial_{\mathscr{r}(\gamma^1)\ldots\mathscr{r}(\gamma^n)}\varphi(\mathbf{X})\,\mathrm{d}\mathbf{X}^{\pi(\mathscr{r}(\gamma^1\ldots\gamma^n))}$$

The result follows straightforwardly from Theorem 2.12, using the fact that $Q_{\mathscr{r}}$ is a sub-coalgebra of $\mathcal{H}_{\mathrm{GL}}$ to write the integrals explicitly in terms of the vector fields F, without having to resort to the more complicated general expression of $\mathbf{F}$.

Remark 2.22 (Simple change of variable formula and quasi-geometric rough paths)**.** We note that the intersection of the two special cases corresponds to Corollary 2.16 (restricted to the quasi-geometric setting, already shown in [3, Theorem 4.20]). Since $\widetilde{\iota}^*\colon \mathcal{H}_{\mathrm{GL}} \twoheadrightarrow \sqcup\!\sqcup$ is not a Hopf algebra morphism, $\widetilde{\iota}^*(\mathbf{X})$ cannot be expected to define a rough path. Nevertheless, the integration of (2.18) only depends on this projection: this is because, in the Davie expansion of that integral, $\mathbf{X}$ is only evaluated on linear combinations of terms of the form $(\bullet_{\alpha_1}\cdots\bullet_{\alpha_m}) \top \pi(\bullet_{\beta_1}\cdots\bullet_{\beta_n})$; using Proposition 1.10 it is then seen inductively that such elements belong to $\top(\bigotimes\bigodot(V^*))$, i.e. the range of $\widetilde{\iota}$. ⧫

Example 2.23 (Obstructions to quasi-geometricity and Itô/Stratonovich calculus)**.** The degree to which a branched rough path is free to be non (quasi-)geometric is constrained, to a certain extent, by its regularity. When $\rho < 2$ there is only one rough path, and it is geometric. When $2 \leq \rho < 3$, a rough path $\mathbf{X}$ is always quasi-geometric and geometric if and only if $\langle\pi(\bullet_\alpha\bullet_\beta),\mathbf{X}\rangle = 0$: in the notation of [21], this corresponds to $[\mathbf{X}] = 0$. Recall that if X is a continuous semimartingale, this condition is satisfied when $\mathbf{X}^{\ell(\alpha,\beta)}_{s,t} := \int_s^t X^\alpha_{s,u}\circ \mathrm{d}X^\beta_u$ is defined by Stratonovich integration, but when it is defined by Itô integration $\mathbf{X}^{\ell(\alpha,\beta)}_{s,t} := \int_s^t X^\alpha_{s,u}\,\mathrm{d}X^\beta_u$ it holds that $\langle\pi(\bullet_\alpha\bullet_\beta),\mathbf{X}\rangle = [X]$, the quadratic variation of X. We take the opportunity to remark that the framework laid out in this and the previous section makes it possible to recover the classical Itô SDE with drift Equation (0.2) from the equation $\mathrm{d}Y = F(Y)\,\mathrm{d}\mathbf{W}$, with $F_\gamma = \sigma_\gamma$ and $F_{\mathscr{r}(\alpha,\beta)} = \frac{\delta_{\alpha\beta}}{2}\mu$. This is because $\mathbf{W}_t^{\pi(\mathscr{r}(\alpha,\beta))} = [W]^{\alpha\beta}_t = \delta^{\alpha\beta}t$, and the dependence of the coefficients on t is easily handled by coupling the equation with the trivial equation $\mathrm{d}t = \mathrm{d}\mathbf{W}_t^{\pi(\mathscr{r}(1,1))}$. Moving on to higher-order cases, the unique obstruction for quasi-geometricity in the case $3 \leq \rho < 4$ is

$$\langle\pi(\bullet_\alpha \mathrm{I}^{\beta}_{\gamma}),\mathbf{X}\rangle = 0 \tag{2.23}$$

by Example 1.12. Expressions get increasingly complicated as ρ increases, but the quasi-geometricity condition is still tractable for $4 \leq \rho < 5$ and $\dim(V) = 1$, in which case (see Proposition 4.7) it reads

$$\langle\pi(\mathrm{I}\,\mathrm{I}),\mathbf{X}\rangle = 0 \tag{2.24}$$

since the previous order-3 condition is automatically cleared by dimensionality. ⧫

3. The Itô-Stratonovich correction formula

3.1. Commutative B$_\infty$-algebras are shuffle algebras.

The main purpose of this section is to define an explicit, natural Hopf isomorphism between the Connes-Kreimer Hopf algebra and the shuffle algebra over its primitive elements, and to use this isomorphism to write an RDE driven by a branched rough path $\mathbf{X}$ as an equivalent RDE driven by a geometric rough path, defined algebraically in terms of $\mathbf{X}$, and with the new vector fields defined algebraically in terms of the old.

Recall that on any Hopf algebra $(H,\times,\Delta)$ the *convolution product* on linear endomorphisms is defined as

$$f * g := (\,\cdot\times\cdot\,)\circ(f\otimes g)\circ\Delta, \qquad f,g\in\mathcal{L}(H,H)$$

making $(\mathcal{L}(H,H),*)$ an associative algebra. For the following definition and lemma, we follow [47] (see also [1]).

Definition 3.1 (Eulerian idempotent)**.** Given a graded connected Hopf algebra $(H,\times,\Delta)$, which we assume to be commutative or cocommutative, we define its Eulerian idempotent $\mathrm{e}_H\in\mathcal{L}(H,H)$

$$\begin{aligned}
\mathrm{e}_H &:= \log_*(\mathrm{id})\\
&:= \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}(\mathrm{id}-\eta_H\circ\varepsilon_H)^{*n}\\
&= \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}\times^{(n-1)}\circ(\mathrm{id}-\eta_H\circ\varepsilon_H)^{\otimes n}\circ\Delta^{(n-1)}\\
&= \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}\times^{(n-1)}\circ\widetilde{\Delta}^{(n-1)}
\end{aligned}$$

where ε_H denotes the counit $H\twoheadrightarrow\mathbb{R}$ and η_H denotes the unit $\mathbb{R}\hookrightarrow H$, $\times^{(n-1)}$ is the product taking n arguments, and $\widetilde{\Delta}$ the reduced coproduct. ▲

In Sweedler notation the map e_H reads

$$\mathrm{e}_H(h) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} h^{(1)} \times \cdots \times h^{(n)}. \tag{3.1}$$

We summarise a few properties of e_H, which distinguish between the case of H commutative or cocommutative below [47].

Proposition 3.2 (Properties of the Eulerian idempotent)**.** *Let H be as above.*

1 *Setting*

$$\mathrm{e}_n := \frac{1}{n!}\mathrm{e}_H^{*n}, \qquad \text{it holds that} \quad H = \bigoplus_{n=0}^{\infty} \mathrm{e}_n(H) \quad \text{and} \quad \mathrm{e}_m \circ \mathrm{e}_n = \delta_{mn}\mathrm{e}_n.$$

In particular, e_n is idempotent for every $n \geq 1$.

2 *Let H° be the graded dual Hopf algebra of H. Then $\mathrm{e}_H^\circ = \mathrm{e}_{H^\circ}$.*

3 $\mathrm{Prim}(H) \subseteq \mathrm{e}(H)$ *and equality holds if H is cocommutative.*

4 *If H is cocommutative, the inclusion $\mathrm{e}(H) \hookrightarrow H$ induces an algebra isomorphism $\mathcal{U}(\mathrm{e}(H)) \cong H$, where $\mathcal{U}$ denotes the universal enveloping algebra functor.*

5 *If H is commutative, setting $H_+ :=$ ker ε_H and $H_+^{\times 2}$ the space generated by products of elements in H_+, $\ker(\mathrm{e}_H) = H_+^{\times 2}$, and the resulting quotient map splits the exact sequence*

$$\begin{tikzcd} 0 \arrow[r] & H_+^{\times 2} \arrow[r] & H_+ \arrow[r] & H/H_+^{\times 2} \arrow[r] \arrow[l, bend left, "\mathrm{e}_H"] & 0\,. \end{tikzcd}$$

6 *If H is commutative, $\mathrm{e}_H(H) \hookrightarrow H$ induces an algebra isomorphism $\bigodot(\mathrm{e}_H(H)) \cong H$, where $\bigodot$ denotes the symmetric tensor algebra functor.*

We briefly comment on those aspects that are not difficult to show, and the consequences in the context considered here. In statement (1), showing that the maps e_n are idempotents is non-trivial and we just refer to [47]. The decomposition of H as a direct sum involves writing

$$\mathrm{id} = \exp_* \circ \log_*(\mathrm{id}) = \sum_{n=0}^{\infty} \mathrm{e}_n$$

Assertion (2) is an easy consequence of the definition of e_H and duality of the bialgebra operations; we will be using it for the dual pair $(\mathcal{H}_{\mathrm{CK}}, \mathcal{H}_{\mathrm{GL}})$, i.e. $\mathrm{e}_{\mathrm{CK}}^* = \mathrm{e}_{\mathrm{GL}}$. As for (3), the only non-zero term in (3.1) with $h \in \mathrm{Prim}(H)$ corresponds to $n = 0$, i.e. $\mathrm{e}_H\big|_{\mathrm{Prim}(H)} = \mathrm{id}_{\mathrm{Prim}(H)}$, and the inclusion follows by idempotence. We stress that equality does not hold for $\mathcal{H}_{\mathrm{CK}}$, as illustrated by the following simple example:

$$\mathrm{e}_{\mathrm{CK}}\Big(\genfrac{}{}{0pt}{}{\bullet^{\alpha}}{\bullet_{\beta}}\Big) = \genfrac{}{}{0pt}{}{\bullet^{\alpha}}{\bullet_{\beta}} - \tfrac{1}{2}\bullet_{\alpha}\bullet_{\beta} \notin \mathcal{P},$$

of which $\pi(\bullet_\alpha\bullet_\beta)$ is the symmetrisation. Statement (4) is the celebrated Milnor-Moore theorem [40]. To see (5), recall the notion of *infinitesimal character* [39, Proposition 22], in our case taking values in H itself, i.e. elements $c \in \mathcal{L}(H, H)$ s.t.

$$c(xy) = 1_H(x)c(y) + c(x)1_H(y).$$

When H is commutative, $\exp_*$ and $\log_*$ define bijections, inverse to each other, between the group of H-valued characters, i.e. algebra morphisms $H \to H$, and the Lie algebra of H-valued infinitesimal characters. Then if $x, y \in H_+$ it is immediate that $\mathrm{e}_H(xy) = \log_*(\mathrm{id})(xy) = 0$. The splitting of the short exact sequence amounts to saying that, for $h \in H_+^{\times 2}$, $\mathrm{e}_H(h) = h + k$ with $k \in H_+^{\times 2}$, which is again evident from (3.1). In the case of $H = \mathcal{H}_{\mathrm{CK}}$, the exact sequence is

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{F} \arrow[r] & \mathcal{H}_+ \arrow[r] & \mathcal{T} \arrow[r] \arrow[l, bend left, "\mathrm{e}_{\mathrm{CK}}"] & 0\,. \end{tikzcd}$$

Assertion (5) is known as Leray's theorem, see again [40, Theorem 7.5], and can be viewed as of a piece with the statement in the cocommutative case, in the sense that the universal enveloping algebra of the trivial Lie algebra is the symmetric algebra.

We now proceed with $H = \mathcal{H}_{\mathrm{CK}}$, and continue to denote e the corresponding Eulerian idempotent. We denote $\mathcal{E} := \mathrm{e}(\mathcal{H}_{\mathrm{CK}})$ and, as usual, view it as a functor $\underline{\mathrm{Vec}} \to \underline{\mathrm{Vec}}$. We also view e as a natural transformation $\mathcal{H} \Rightarrow \mathcal{E}$, which is possible since e is defined via the natural operations of $\mathcal{H}_{\mathrm{CK}}$, and Proposition 1.3. Proposition 3.2 implies that the inclusion $\mathcal{E} \hookrightarrow \mathcal{H}_{\mathrm{CK}}$ induces a natural isomorphism $\mathcal{H}_{\mathrm{CK}} \cong \bigodot(\mathcal{E})$.

In this section we will show how to define a natural isomorphism from $\mathcal{H}_{\mathrm{CK}} \cong \bigsqcup(\mathcal{P})$. This will be achieved by combining both maps π and e. Before doing this, however, it is instructive to see that considering the two descriptions of $\mathcal{H}_{\mathrm{CK}}$ separately — one as a tensor algebra and one as a symmetric algebra — does not yield the desired result.

Proposition 3.3 (Two failed attempts at a Hopf isomorphism)**.**

1 $\top^{-1}\colon \mathcal{H}_{\mathrm{CK}} \to \shuffle(\mathcal{P})$ *is a coalgebra isomorphism but not an algebra morphism.*
2 *The unique algebra morphism* $\varphi\colon \mathcal{H}_{\mathrm{CK}} \to \shuffle(\mathcal{P})$ *with* $\varphi|_{\mathcal{E}} = \top^{-1}|_{\mathcal{E}}$ *is an algebra isomorphism but not a coalgebra morphism.*

Proof. That $\top^{-1}$ is a coalgebra isomorphism was stated in Theorem 1.8. To check that it is not an algebra isomorphism, the following suffices:

$$\top^{-1}(\bullet_\alpha\cdot\bullet_\beta) = \Big(\bullet_\alpha\bullet_\beta - \begin{smallmatrix}\bullet_\alpha\\ \vert\\ \bullet_\beta\end{smallmatrix} - \begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\Big) + \bullet_\alpha\otimes\bullet_\beta + \bullet_\beta\otimes\bullet_\alpha \neq \bullet_\alpha\otimes\bullet_\beta + \bullet_\beta\otimes\bullet_\alpha = \top^{-1}(\bullet_\alpha)\shuffle\top^{-1}(\bullet_\beta)$$

We now consider the map φ. It is clear from free commutativity of $\mathcal{H}$ over $\mathcal{E}$ that such a map exists and is an algebra morphism. For the purposes of this proof, it is convenient to consider the shuffle operation defined directly on $\mathcal{H}_{\mathrm{CK}}$, call it

$$\shuffle^\top := \top\circ\shuffle\circ\top^{-1}$$

which by Corollary 1.16 is related to the forest product by

$$\begin{aligned}(p_1\top\ldots\top p_m)\shuffle^\top(p_{m+1}\top\ldots\top p_{m+n}) &= \sum_{\sigma\in\mathrm{Sh}(m,n)} p_{\sigma^{-1}(1)}\top\ldots\top p_{\sigma^{-1}(m+n)}\\ &= \pi_{m+n}\big((p_1\top\ldots\top p_m)\cdot(p_{m+1}\top\ldots\top p_{m+n})\big),\end{aligned}$$

although we note that this will not be used in this proof. We are now interested in showing that the algebra morphism $(\mathcal{H}_{\mathrm{CK}},\cdot)\to(\mathcal{H}_{\mathrm{CK}},\shuffle^\top)$, which we continue to call φ with slight abuse of notation, is bijective. To do this, we observe that it admits the expression $\varphi = \exp_*^{\shuffle^\top}\circ\log_*(\mathrm{id})$, where the superscript indicates that the convolution product is taken w.r.t. the product $\shuffle^\top$ instead of the forest product whilst in both cases the coproduct Δ_{CK} is the same. This is because φ maps

$$\begin{aligned}\mathcal{H}_{\mathrm{CK}}\ni h &= \exp_*\circ\log_*(\mathrm{id})(h)\\ &= \sum_{n=0}^\infty\frac{1}{n!}\mathrm{e}(h^{(1)})\cdots\mathrm{e}(h^{(n)})\\ &\mapsto \sum_{n=0}^\infty\frac{1}{n!}\mathrm{e}(h^{(1)})\shuffle^\top\cdots\shuffle^\top\mathrm{e}(h^{(n)})\\ &= \exp_*^{\shuffle^\top}\circ\log_*(\mathrm{id})(h).\end{aligned}$$

From this it follows that for $h\in\mathcal{E}$, denoting $\mathrm{e}_{\shuffle^\top} = \log_*^{\shuffle^\top}(\mathrm{id})$ the Eulerian idempotent of the Hopf algebra $(\mathcal{H},\shuffle^\top,\Delta_{\mathrm{CK}})$

$$\begin{aligned}\mathrm{e}_{\shuffle^\top}(h) &= \sum_{n=1}^\infty\frac{(-1)^{n-1}}{n}h^{(1)}\shuffle^\top\cdots\shuffle^\top h^{(n)}\\ &= \varphi\left(\sum_{n=1}^\infty\frac{(-1)^{n-1}}{n}h^{(1)}\cdots h^{(n)}\right)\\ &= \varphi\circ\mathrm{e}_{\mathrm{CK}}(h)\\ &= \varphi(h)\\ &= h\end{aligned}$$

which by idempotence implies $\mathcal{E}\subseteq\mathrm{e}_{\shuffle^\top}(\mathcal{H})$, and running through the same argument with products reversed yields the other inclusion, i.e. $\mathcal{H}$ is free commutative over the same $\mathcal{E}$ w.r.t. either product. From this it follows immediately that φ is bijective with inverse $\exp_*\circ\log_*^{\shuffle^\top}(\mathrm{id})$. In order to find a counterexample to the coalgebra morphism property, we must consider degree 3 with decorations. From the definition of the operations one has

$$\varphi\Big(\begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\Big) = \begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix} - \frac12\bullet_\alpha\bullet_\beta + \frac12\bullet_\alpha\shuffle^\top\bullet_\beta = \begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix} - \frac12\pi(\bullet_\alpha\bullet_\beta),$$

Thanks to this identity and the trivial one $\varphi(\bullet_\alpha) = \bullet_\alpha$, we compute

$$\begin{aligned}\varphi^{\otimes2}\Delta_{\mathrm{CK}}\left(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\right) &= \varphi^{\otimes2}\left(1\otimes\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\end{smallmatrix}\otimes\bullet_\alpha + \bullet_\gamma\otimes\begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\otimes 1\right)\\ &= 1\otimes\varphi\left(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\right) + \varphi\left(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\right)\otimes 1 + \bullet_\gamma\otimes\varphi\Big(\begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\Big) + \varphi\Big(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\end{smallmatrix}\Big)\otimes\bullet_\alpha\\ &= 1\otimes\varphi\left(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\right) + \varphi\left(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix}\right)\otimes 1 + \bullet_\gamma\otimes\Big(\begin{smallmatrix}\bullet_\beta\\ \vert\\ \bullet_\alpha\end{smallmatrix} - \frac12\pi(\bullet_\alpha\bullet_\beta)\Big) + \Big(\begin{smallmatrix}\bullet_\gamma\\ \vert\\ \bullet_\beta\end{smallmatrix} - \frac12\pi(\bullet_\gamma\bullet_\beta)\Big)\otimes\bullet_\alpha\,.\end{aligned}$$

On the other hand, we pass to compute φ on the same tree via the explicit definition of $\pi(\bullet_\alpha\bullet_\beta\bullet_\gamma)$

$$
\begin{aligned}
\varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) &= \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix} - \frac{1}{2}\Big(\bullet_\alpha \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \bullet_\gamma \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix}\Big) + \frac{1}{3}\bullet_\alpha\bullet_\beta\bullet_\gamma + \frac{1}{2}\bullet_\alpha \sqcup\!\sqcup^{\top} \Big(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} - \frac{1}{2}\bullet_\beta\bullet_\gamma\Big) + \frac{1}{2}\Big(\begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix} - \frac{1}{2}\bullet_\alpha\bullet_\beta\Big) \sqcup\!\sqcup^{\top} \bullet_\gamma + \frac{1}{6}\bullet_\alpha \sqcup\!\sqcup^{\top} \bullet_\beta \sqcup\!\sqcup^{\top} \bullet_\gamma \\
&= \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix} - \frac{1}{2}\Big(\bullet_\alpha \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \bullet_\gamma \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix}\Big) + \frac{1}{3}\bullet_\alpha\bullet_\beta\bullet_\gamma + \frac{1}{6}\sum_{\sigma\in\mathbb{S}_3} \begin{smallmatrix}\bullet\sigma(\gamma)\\ \bullet\sigma(\beta)\\ \bullet\sigma(\alpha)\end{smallmatrix} - \frac{1}{4}\Big(\bullet_\alpha \sqcup\!\sqcup^{\top} \bullet_\beta\bullet_\gamma + \bullet_\alpha\bullet_\beta \sqcup\!\sqcup^{\top} \bullet_\gamma\Big) + \frac{1}{2}\Big(\bullet_\alpha \sqcup\!\sqcup^{\top} \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix} \sqcup\!\sqcup^{\top} \bullet_\gamma\Big) \\
&= 2\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix} + \frac{1}{2}\left(\begin{smallmatrix}\bullet\alpha\\ \bullet\gamma\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\gamma\\ \bullet\alpha\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \bullet\gamma\\ \bullet\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\\ \bullet\gamma\end{smallmatrix}\right) - \frac{1}{2}\Big(\bullet_\alpha \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \bullet_\gamma \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix}\Big) + \frac{1}{3}\pi(\bullet_\alpha\bullet_\beta\bullet_\gamma) \\
&+ \frac{1}{12}\big(\bullet_\alpha \top \pi(\bullet_\beta\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\beta)\top\bullet_\gamma + \pi(\bullet_\beta\bullet_\gamma)\top\bullet_\alpha + \bullet_\gamma\top\pi(\bullet_\alpha\bullet_\beta)\big) + \frac{1}{3}\big(\bullet_\beta\top\pi(\bullet_\alpha\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\gamma)\top\bullet_\beta\big)\,.
\end{aligned}
$$

Applying the coproduct we derive

$$
\begin{aligned}
\Delta_{\mathrm{CK}}\varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) &= 1\otimes\varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) + \varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right)\otimes 1 + \frac{1}{2}\widetilde{\Delta}_{\mathrm{CK}}\left(4\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet\alpha\\ \bullet\gamma\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\gamma\\ \bullet\alpha\\ \bullet\beta\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \bullet\gamma\\ \bullet\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\\ \bullet\gamma\end{smallmatrix}\right) \\
&- \frac{1}{2}\widetilde{\Delta}_{\mathrm{CK}}\left(\bullet_\alpha \begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \bullet_\gamma \begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) + \frac{1}{3}\widetilde{\Delta}_{\mathrm{CK}}\big(\bullet_\beta\top\pi(\bullet_\alpha\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\gamma)\top\bullet_\beta\big) \\
&+ \frac{1}{12}\widetilde{\Delta}_{\mathrm{CK}}\big(\bullet_\alpha\top\pi(\bullet_\beta\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\beta)\top\bullet_\gamma + \pi(\bullet_\beta\bullet_\gamma)\top\bullet_\alpha + \bullet_\gamma\top\pi(\bullet_\alpha\bullet_\beta)\big) \\
&= 1\otimes\varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) + \varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right)\otimes 1 - \frac{1}{6}\big(\bullet_\beta\otimes\pi(\bullet_\alpha\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\gamma)\otimes\bullet_\beta\big) + \frac{1}{12}\big(\bullet_\alpha\otimes\pi(\bullet_\beta\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\beta)\otimes\bullet_\gamma\big) \\
&+ \left(\frac{3}{2}\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\end{smallmatrix} + \frac{1}{2}\begin{smallmatrix}\bullet\beta\\ \bullet\gamma\end{smallmatrix} - \frac{1}{2}\bullet_\beta\bullet_\gamma + \frac{1}{12}\pi(\bullet_\beta\bullet_\gamma)\right)\otimes\bullet_\alpha + \bullet_\gamma\otimes\left(\frac{3}{2}\begin{smallmatrix}\bullet\beta\\ \bullet\alpha\end{smallmatrix} + \frac{1}{2}\begin{smallmatrix}\bullet\alpha\\ \bullet\beta\end{smallmatrix} - \frac{1}{2}\bullet_\alpha\bullet_\beta + \frac{1}{12}\pi(\bullet_\beta\bullet_\alpha)\right).
\end{aligned}
$$

Subtracting the two expressions we obtain

$$
\begin{aligned}
\Delta_{\mathrm{CK}}\varphi\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) - \varphi^{\otimes 2}\Delta_{\mathrm{CK}}\left(\begin{smallmatrix}\bullet\gamma\\ \bullet\beta\\ \bullet\alpha\end{smallmatrix}\right) &= -\frac{1}{6}\big(\bullet_\beta\otimes(\pi(\bullet_\alpha\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\gamma)\otimes\bullet_\beta\big) \\
&+ \frac{1}{12}\big(\bullet_\alpha\otimes\pi(\bullet_\beta\bullet_\gamma) + \pi(\bullet_\alpha\bullet_\beta)\otimes\bullet_\gamma + \pi(\bullet_\beta\bullet_\gamma)\otimes\bullet_\alpha + \bullet_\gamma\otimes\pi(\bullet_\alpha\bullet_\beta)\big)
\end{aligned}
$$

which is not zero when α, β and γ are different.

$\square$

The first map in Proposition 3.3 did not use Eulerian idempotents, and indeed did not require commutativity, while the second failed to leverage cofreeness. In the following theorem we put both structures to use. We note that the use of "Log" and "Exp" is motivated by Hoffman's isomorphism, see Section 3.2 below. Given maps $\varphi_i\colon \mathcal{P}^{\top j_i} \to \mathcal{P}$ we denote $\varphi_1 \top\cdots\top \varphi_m\colon \mathcal{P}^{\top(j_1+\ldots+j_m)} \to \mathcal{P}^{\top m}$ their concatenation using the isomorphism $\mathcal{H}_{\mathrm{CK}} \cong \bigotimes(\mathcal{P})$.

Theorem 3.4 (Natural isomorphism $\mathcal{H}_{\mathrm{CK}} \cong \sqcup\!\sqcup(\mathcal{P})$). *There exists a unique isomorphism* $\mathrm{Log}\colon \mathcal{H}_{\mathrm{CK}} \to \sqcup\!\sqcup(\mathcal{P})$ *with the property that*

$$\pi\circ\mathrm{Log} = \pi\circ\mathrm{e}$$

and it is given by

$$\mathrm{Log} = \sum_{n\geq 1}(\pi\circ\mathrm{e})^{\otimes n}\circ\widetilde{\Delta}_{\mathrm{CK}}^{(n-1)}$$

Its inverse $\mathrm{Exp}\colon \sqcup\!\sqcup(\mathcal{P}) \to \mathcal{H}_{\mathrm{CK}}$ *is given by the following combinatorial expression involving summing over increasing sequences of integers and their compositions*

$$\sum_{\substack{k\geq 0\\ 1\leq n^0<\ldots<n^k\\ n^1_1+\ldots+n^1_{n^0}=n^1\\ \ldots\\ n^k_1+\ldots+n^k_{n^{k-1}}=n^k}} (-1)^k\big[(\pi\circ\mathrm{e}\circ\pi_{n^1_1})\top\cdots\top(\pi\circ\mathrm{e}\circ\pi_{n^1_{n^0}})\big]\circ\cdots\circ\big[(\pi\circ\mathrm{e}\circ\pi_{n^k_1})\top\cdots\top(\pi\circ\mathrm{e}\circ\pi_{n^k_{n^{k-1}}})\big]$$

or equivalently by the recursion

$$\mathrm{Exp}^m_n = -\sum_{m\leq k<n}\mathrm{Exp}^m_k\circ\mathrm{Log}^k_n,$$

where $\mathrm{Exp}^m_n \coloneqq \pi_m\circ\mathrm{Exp}\circ\pi_n\colon \mathcal{P}^{\otimes n}\to\mathcal{P}^{\top m}$*, and similarly for* Log.

Log *and* Exp *are natural isomorphisms between the functors* $\mathcal{H}_{\mathrm{CK}}, \sqcup\!\sqcup\circ\mathcal{P}\colon \underline{\mathrm{Vec}} \to \underline{\mathrm{comm}\widehat{\mathbf{B}}_\infty}$ *restricting to the identity on* $\mathcal{P}$.

Proof. By cofreeness (1.6), there is a unique coalgebra homomorphism Log with a specified projection $\pi \circ \mathrm{Log}$ (in this case equals to $\pi \circ \mathrm{e}$). By Proposition 1.5 it follows that such a map will have the following expression: for $p_1, \ldots, p_n \in \mathcal{P}$

$$\mathrm{Log}(p_1 \top \ldots \top p_n) = \sum_{\substack{n_1+\ldots+n_k=n \\ n_1,\ldots,n_k \geq 1}} (\pi \circ \mathrm{e})(p_1 \top \ldots \top p_{n_1}) \otimes \cdots \otimes (\pi \circ \mathrm{e})(p_{n_1+\ldots+n_{k-1}+1} \top \ldots \top p_n)$$

which by Theorem 1.8 equals the second expression in the statement; the conclusion then follows by linear extension. We must show that this coalgebra morphism is an algebra morphism: denoting $\times = \cdot$ the forest product for better clarity, this is the case if and only if

$$\mathrm{Log} \circ (\,\cdot \times \cdot\,) = (\,\cdot \sqcup\!\sqcup \cdot\,) \circ \mathrm{Log}^{\otimes 2} \colon \mathcal{H}_{\mathrm{CK}}^{\otimes 2} \to \coprod(\mathcal{P})$$

Since $\mathcal{H}_{\mathrm{CK}}$ and $\coprod(\mathcal{P})$ are bialgebras and by the above, both these maps are coalgebra morphisms, and thus again by cofreeness they coincide if and only if

$$\pi \circ \mathrm{e} \circ (\,\cdot \times \cdot\,) = \pi \circ \mathrm{Log} \circ (\,\cdot \times \cdot\,) = \pi_1 \circ (\,\cdot \sqcup\!\sqcup \cdot\,) \circ \mathrm{Log}^{\otimes 2} = 0,$$

where π_1 is the canonical projection $\coprod(\mathcal{P}) \twoheadrightarrow \mathcal{P}$, since π_1 vanishes on $\coprod(\mathcal{P})_+^{\sqcup\!\sqcup 2}$ and on the ground field. But by point 5. of Proposition 3.2, e_{CK} maps $\mathcal{H}_+^{\times 2}$ to zero, concluding the proof that Log is a bialgebra, and hence Hopf algebra morphism. Proposition 1.5 already implies that Log is invertible, since $\mathrm{Log}|_{\mathcal{P}} = (\pi \circ \mathrm{e})|_{\mathcal{P}} = \mathrm{id}_{\mathcal{P}}$, a consequence of point 3. of Proposition 3.2; this also implies that its inverse Exp restricts to the identity on $\mathcal{P}$.

In order to prove the formula for Exp, it is convenient to think of a coalgebra map between cofree coalgebras as a triangular matrix: again by Proposition 1.5, setting $\mathrm{Log}_n^m \coloneqq \pi_m \circ \mathrm{Log} \circ \pi_n \colon \mathcal{H}_{\mathrm{CK}} \to \coprod(\mathcal{P})$, we have

$$\mathrm{Log} = \underbrace{\sum_{1 \leq m < n} \mathrm{Log}_n^m}_{=:L} + \underbrace{\sum_{n \geq 1} \mathrm{Log}_n^n}_{=:D}$$

$$\text{with} \quad \mathrm{Log}_n^m = \sum_{\substack{n_1+\ldots+n_m=n \\ n_1,\ldots,n_m \geq 1}} (\pi_1 \circ \mathrm{e} \circ \pi_{n_1}) \otimes \cdots \otimes (\pi_1 \circ \mathrm{e} \circ \pi_{n_m})$$

Now, $D = \sum_{n \geq 1} \top^{-1} \circ \pi_n$ is invertible with inverse $\sum_{n \geq 1} \top \circ \pi_n$, and

$$D^{-1} \circ \mathrm{Log} = D^{-1} \circ L + \mathrm{id}, \qquad (D^{-1} \circ L + \mathrm{id}) \circ \sum_{n \geq 0} (-1)^n (D^{-1} \circ L)^{\circ n} = \mathrm{id}$$

with the sum locally finite, since $D^{-1} \circ L$ is nilpotent of degree N on $\bigoplus_{n=0}^N \mathcal{P}^{\top n}$. This implies

$$\begin{aligned}
\mathrm{Exp} &= (D^{-1} \circ \mathrm{Log})^{-1} \circ D^{-1} \\
&= \sum_{k \geq 0} (-1)^k (D^{-1} \circ L)^{\circ k} \circ D^{-1} \\
&= \sum_{k \geq 0} \sum_{\substack{1 \leq m^1 < n^1 \\ \cdots \\ 1 \leq m^k < n^k}} (-1)^k (\top \circ \mathrm{Log}_{n^1}^{m^1} \circ \top) \circ \cdots \circ (\top \circ \mathrm{Log}_{n^k}^{m^k} \circ \top) \\
&= \sum_{k \geq 0} (-1)^k \sum_{1 \leq n^0 < \ldots < n^k} (\top \circ \mathrm{Log}_{n^1}^{n^0} \circ \top) \circ \cdots \circ (\top \circ \mathrm{Log}_{n^k}^{n^{k-1}} \circ \top) \\
&= \sum_{k \geq 0} (-1)^k \sum_{1 \leq n^0 < \ldots < n^k} \sum_{\substack{n_1^1+\ldots+n_{n^0}^1=n^1 \\ \cdots \\ n_1^k+\ldots+n_{n^{k-1}}^k=n^k}} [(\pi \circ \mathrm{e} \circ \pi_{n_1^1}) \top \cdots \top (\pi \circ \mathrm{e} \circ \pi_{n_{n^0}^1})] \circ \cdots \\
&\qquad \cdots \circ [(\pi \circ \mathrm{e} \circ \pi_{n_1^k}) \top \cdots \top (\pi \circ \mathrm{e} \circ \pi_{n_{n^{k-1}}^k})]
\end{aligned}$$

The recursive expression is now obvious. Naturality follows again from Proposition 1.3 and the fact that all operations involved in the expressions of Log and Exp are natural. □

Remark 3.5. To keep the presentation concise and concrete we have stated the above theorem in the case which is relevant to the Itô-Stratonovich transformation of branched rough paths, Corollary 3.7 below. However, both the statement and the proof carry over word for word to the case of arbitrary functors $\underline{\mathrm{Vec}} \to \underline{\mathrm{comm}\mathbf{B}_\infty}$, simply by replacing π with the cofreeness projection and e with the Eulerian idempotent for the commutative $\mathbf{B}_\infty$-algebra in question. ♦

Example 3.6 (Explicit computations of Log and Exp at order 3 (labelled) and order 4 (unlabelled))**.** We compute Exp up to level 3 in the primitiveness grading (which yields more compact formulae than the forest basis). The expression in the forest basis can be obtained via Example 1.12.

$$\operatorname{Exp}(p) = p, \qquad \operatorname{Exp}(p \otimes q) = p \top q + \frac{1}{2}\pi(pq)$$
$$\begin{aligned}\operatorname{Exp}(p \otimes q \otimes r) &= p \top q \top r + \frac{1}{2}(\pi(pq) \top r + p \top \pi(qr)) + \frac{1}{6}\pi(pqr) \\ &\quad + \frac{1}{4}(\pi((p \top q)r) + \pi(p(q \top r)) - \pi((q \top p)r) - \pi(p(r \top q)))\end{aligned}$$

We note that everything on the first line is already present in Hoffmann's isomorphism for the quasi-shuffle case (see below Section 3.2), while terms in the second line are specific to the case of Connes-Kreimer. The following is an example of Exp at level 4 in the undecorated case.

$$\begin{aligned}\operatorname{Exp}(\bullet \otimes \bullet \otimes \bullet \otimes \bullet) = \substack{\bullet \\ \vert \\ \bullet \\ \vert \\ \bullet \\ \vert \\ \bullet} + \frac{1}{6}\bullet \top \pi(\bullet\bullet\bullet) + \frac{1}{6}\pi(\bullet\bullet\bullet) \top \bullet + \frac{1}{2}\bullet \top \pi(\bullet\bullet) \top \bullet + \frac{1}{2}\pi(\bullet\bullet) \top \bullet \top \bullet \\ + \frac{1}{2}\bullet \top \bullet \top \pi(\bullet\bullet) + \frac{1}{24}\pi(\bullet\bullet\bullet\bullet) - \frac{1}{2}\pi\Big(\substack{\bullet \\ \vert \\ \bullet}\ \substack{\bullet \\ \vert \\ \bullet}\Big). \quad \blacklozenge\end{aligned}$$

We now discuss the consequences that this has for rough paths. Recall that a $\mathcal{Q}$-valued geometric ρ-rough path (of inhomogeneous regularity) is defined in complete analogy with Definition 2.1, with the only difference that the Hopf algebra $\mathcal{H}_{\mathrm{CK}}$ is replaced with $\sqcup\!\sqcup(\mathcal{P})$. An important caveat is that the grading on $\sqcup\!\sqcup(\mathcal{P})$ used in the regularity requirement is the inhomogeneous one, which takes into account the grading on $\mathcal{P}$ inherited from $\mathcal{H}_{\mathrm{CK}}$

$$|p_1 \otimes \cdots \otimes p_n| = |p_1| + \ldots + |p_n| \tag{3.2}$$

and ρ refers to terms of the worst regularity, $X|_V$. This also affects how the tensor algebra (along with Davie expansions, etc.) is truncated. In particular from the above examples we clearly see that Exp (and Log) are not graded maps when using the standard grading in $\sqcup\!\sqcup(\mathcal{P})$ since it decreases primitiveness, but it is under (3.2).

Recall that the graded dual bialgebra to $\sqcup\!\sqcup(\mathcal{P})$ is $\bigotimes(\mathcal{Q}) \coloneqq (\bigotimes(\mathcal{Q}), \otimes, \Delta_{\sqcup\!\sqcup})$, where $\Delta_{\sqcup\!\sqcup}$ denotes "unshuffling" (see, for instance, [48, §1.5]), and the dual pairing is the tensor product of the dual pairings $\mathcal{P} \otimes \mathcal{Q} \to \mathbb{R}$ inherited from $\mathcal{H}_{\mathrm{CK}} \otimes \mathcal{H}_{\mathrm{GL}} \to \mathbb{R}$ of Theorem 1.18. We can now state the main application of Theorem 3.4.

Corollary 3.7 (Itô-Stratonovich correction formula for RDEs driven by branched rough paths)**.** *Let* $\mathbf{X}$ *be a* ρ*-branched rough path controlled by* ω*. Then* $\overline{\mathbf{X}} \coloneqq \operatorname{Exp}^*(\mathbf{X})$ *defines a* $\mathcal{Q}$*-valued geometric* ρ*-rough path controlled by* ω *and the* $\mathbf{X}$*-driven RDE* (2.12) *is equivalent to the* $\overline{\mathbf{X}}$*-driven one*

$$\mathrm{d}Y = \sum_{f \in \mathscr{H}} \varsigma(f)^{-1} \sum_{l \in \ell(f)} \langle \widehat{F}(Y), \mathrm{e}_{\mathrm{GL}} \circ \pi^*(f_l) \rangle \, \mathrm{d}\overline{\mathbf{X}}^{\pi(f_l)}$$

where $\widehat{F}$ *is taken as in Theorem 2.10.*

Proof. The first statement is an immediate consequence of Theorem 3.4 and the two definitions of rough path, together with the fact that Exp preserves the grading in (3.2). The $\overline{\mathbf{X}}$-driven equation in the statement has Davie expansion

$$Y_{s,t} \approx \langle \operatorname{Log}(\widehat{F}(Y_s)), \operatorname{Exp}^*(\mathbf{X}_{s,t}) \rangle_{\sqcup\!\sqcup(\mathcal{P}), \bigotimes(\mathcal{Q})} \approx \langle \operatorname{Exp} \circ \operatorname{Log} \widehat{F}(Y_s), \mathbf{X}_{s,t} \rangle_{\mathrm{CK,GL}} \approx \langle \widehat{F}(Y_s), \mathbf{X}_{s,t} \rangle_{\mathrm{CK,GL}}$$

which is the Davie expansion for the solution to (2.12). $\square$

Example 3.8 (The Itô-Stratonovich formula for $\lfloor \rho \rfloor = 3$)**.** Continuing with Example 2.11, we rewrite the RDE in geometric terms using Corollary 3.7. One directly obtains:

$$\begin{aligned}\mathrm{d}Y_t &= F_{\bullet\alpha}\, \mathrm{d}\overline{\mathbf{X}}_t^{\bullet\alpha} - \frac{1}{2}\widehat{F}_{\substack{\bullet\alpha \\ \vert \\ \bullet\beta}}\, \mathrm{d}\overline{\mathbf{X}}_t^{\pi(\bullet\alpha\bullet\beta)} \\ &\quad + \frac{1}{6}\left(F_{\bullet\alpha\bullet\beta\bullet\gamma} + \frac{1}{3}\widehat{F}_{\substack{\bullet\gamma \\ \vert \\ \bullet\beta \\ \vert \\ \bullet\alpha}} + \frac{1}{6}\widehat{F}_{\substack{\bullet\beta\ \bullet\gamma \\ \vee \\ \bullet\alpha}}\right) \mathrm{d}\overline{\mathbf{X}}_t^{\pi(\bullet\alpha\bullet\beta\bullet\gamma)} \\ &\quad + \left(F_{\frac{1}{2}\left(\bullet\alpha \substack{\bullet\beta \\ \vert \\ \bullet\gamma} - \bullet\gamma \substack{\bullet\beta \\ \vert \\ \bullet\alpha}\right)} - \frac{1}{2}\left(\widehat{F}_{\substack{\bullet\alpha \\ \vert \\ \bullet\beta \\ \vert \\ \bullet\gamma}} + \widehat{F}_{\substack{\bullet\beta \\ \vert \\ \bullet\gamma \\ \vert \\ \bullet\alpha}} + \widehat{F}_{\substack{\bullet\alpha\ \bullet\beta \\ \vee \\ \bullet\gamma}}\right)\right) \mathrm{d}\overline{\mathbf{X}}_t^{\pi\left(\bullet\alpha \substack{\bullet\beta \\ \vert \\ \bullet\gamma}\right)}\end{aligned}$$

Note that at degree 2, since

$$\widehat{F}_{\substack{\bullet\alpha \\ \vert \\ \bullet\beta}} = F_{\bullet\alpha} \triangleright F_{\bullet\beta} - F_{\bullet\alpha\bullet\beta},$$

we recover the well-known Itô–Stratonovich correction for semimartingales, but we also get the modified vector fields at higher orders. We refer again to Example 2.11 for the concrete expression of $\widehat{F}$ in terms of the original vector fields F. ♦

3.2. **The quasi-geometric case: Hoffman's exponential.** In [27], Hoffman defined a Hopf isomorphism between the quasi shuffle and shuffle bialgebras. In this subsection, we show how this isomorphism can be viewed as a special case of ours. In the context of this paper, and of the notation introduced in Section 2.2, we call it $\widetilde{\mathrm{Exp}}\colon \shuffle(\bigodot(U)) \to \widetilde{\shuffle}(U)$ and its inverse $\widetilde{\mathrm{Log}}\colon \widetilde{\shuffle}(U) \to \shuffle(\bigodot(U))$. We recall their definition: for $s_1, \ldots, s_n \in \bigodot(U)$

$$
\begin{aligned}
\widetilde{\mathrm{Exp}}(s_1 \otimes \cdots \otimes s_n) &:= \sum_{\substack{n_1+\ldots+n_k=n \\ n_1,\ldots,n_k \geq 1}} \frac{1}{n_1! \cdots n_k!} (s_1 \odot \cdots \odot s_{n_1}) \otimes \cdots \otimes (s_{n_1+\ldots+n_{k-1}+1} \odot \cdots \odot s_n) \\
\widetilde{\mathrm{Log}}(s_1 \otimes \cdots \otimes s_n) &:= \sum_{\substack{n_1+\ldots+n_k=n \\ n_1,\ldots,n_k \geq 1}} \frac{(-1)^{n-k}}{n_1 \cdots n_k} (s_1 \odot \cdots \odot s_{n_1}) \otimes \cdots \otimes (s_{n_1+\ldots+n_{k-1}+1} \odot \cdots \odot s_n)
\end{aligned}
\tag{3.3}
$$

The following theorem establishes these isomorphisms as particular cases of Theorem 3.4 (defined w.r.t. the functor $\widetilde{\shuffle}\colon \underline{\mathrm{Vec}} \to \underline{\mathrm{comm}\widehat{\mathbf{B}}_\infty}$, see Remark 3.5); furthermore it shows how the our and Hoffman's isomorphisms are related via the arborification quotient maps of Proposition 2.17, and are furthermore related to the recently introduced arborified exponential introduced in [6], denoted here by $\widetilde{\mathfrak{a}}_{\odot}$. It is defined in complete analogy with $\widetilde{\mathfrak{a}}$ Proposition 2.17, the only difference being that forests in the source are decorated with $\bigodot(U)$, not U, and that decorations are multiplied associatively when taking the shuffle product; it has the alternative right inverse given by $\gamma_1 \odot \cdots \odot \gamma_n \mapsto \bullet_{\gamma_1 \odot \cdots \odot \gamma_n}$ defined without invoking π. The existence of a unique and explicit Hopf algebra morphism $\mathfrak{a}\widetilde{\mathrm{Exp}}$ making the lower parallelogram commute is proved in [6, Theorem 2].

Theorem 3.9. *Hoffman's isomorphisms are given by the same formulae of Theorem 3.4, as specified by Remark 3.5, and the following diagram of Hopf morphisms*

$$
\begin{tikzcd}
 & \shuffle(\mathcal{P}(U)) \arrow[rr, "\mathrm{Exp}"] & & \mathcal{H}_{\mathrm{CK}}(U) \arrow[dl, "\mathcal{H}_{\mathrm{CK}}(j)"'] \arrow[dd, two heads, "\widetilde{\mathfrak{a}}"] \\
\mathcal{H}_{\mathrm{CK}}(\bigodot(U)) \arrow[ur] \arrow[dr, two heads, "\mathfrak{a}"'] \arrow[rr, "\mathfrak{a}\widetilde{\mathrm{Exp}}"] & & \mathcal{H}_{\mathrm{CK}}(\bigodot(U)) \arrow[dr, two heads, "\widetilde{\mathfrak{a}}_{\odot}"] & \\
 & \shuffle(\bigodot(U)) \arrow[uu, hook, "\shuffle(i)"'] \arrow[rr, "\widetilde{\mathrm{Exp}}"] & & \widetilde{\shuffle}(U)
\end{tikzcd}
$$

where $i := \bigodot(U) \hookrightarrow \mathcal{P}(U)$ and $j := U \hookrightarrow \bigodot(U)$ are the natural inclusions (and the unlabelled map is the composition $\mathfrak{a} \circ \shuffle(i)$), commutes.

Proof. We begin with the first statement, which we verify in the case of the logarithm; the analogoues statement for the exponential follows by uniqueness of inverses. By cofreeness, it suffices to check that $\pi_1 \circ \widetilde{\mathrm{Log}} = \pi_1 \circ \mathrm{e}_{\widetilde{\shuffle}}$, where $\pi_1\colon \widetilde{\shuffle}(U) \twoheadrightarrow \bigodot(U)$ and $\mathrm{e}_{\widetilde{\shuffle}}$ is the Eulerian idempotent for the quasi shuffle algebra. We compute, using Definition 3.1 and (3.3)

$$
\begin{aligned}
\pi_1 \circ \mathrm{e}_{\widetilde{\shuffle}}(s_1 \otimes \cdots \otimes s_n) &= \sum_{m \geq 1} \frac{(-1)^{m-1}}{m} \pi_1 \circ \widetilde{\shuffle}^{(m-1)} \circ \widetilde{\Delta}_{\otimes}^{(m-1)}(s_1 \otimes \cdots \otimes s_n) \\
&= \frac{(-1)^{n-1}}{n}(s_1 \odot \cdots \odot s_n) \\
&= \pi_1 \circ \widetilde{\mathrm{Log}}(s_1 \otimes \cdots \otimes s_n)
\end{aligned}
$$

since the only term to survive application of π_1 appears once as a summand in the $(n-1)$-fold deconcatenation. Although redundant, it is informative to see why Hoffman's exponential has a much simpler closed-form expression than Exp for general $\mathbf{B}_\infty$-algebras. We check that $\pi \circ \widetilde{\mathrm{Exp}}$ satisfies the same recursion as Exp (which is already implied by Remark 3.5):

$$\begin{aligned}
&- \sum_{1\le k<n} \widetilde{\mathrm{Exp}}^1_k \circ \widetilde{\mathrm{Log}}^k_n(s_1 \otimes \cdots \otimes s_n) \\
&= - \sum_{1\le k<n} \widetilde{\mathrm{Exp}}^1_k \circ \sum_{\substack{n_1+\ldots+n_k=n \\ n_1,\ldots,n_k\ge 1}} \frac{(-1)^{n-k}}{n_1\cdots n_k}(s_1 \odot \cdots \odot s_{n_1}) \otimes \cdots \otimes (s_{n_1+\ldots+n_{k-1}+1} \odot \cdots \odot s_{n_k}) \\
&= \Bigg(\sum_{\substack{1\le k<n \\ n_1+\ldots+n_k=n \\ n_1,\ldots,n_k\ge 1}} \frac{(-1)^{n-k}}{n_1\cdots n_k} \Bigg) s_1 \odot \cdots \odot s_n
\end{aligned}$$

and the assertion follows from the fact that the coefficient is equal to $1/n!$, as can be verified by writing out $t = \exp \circ \log(t+1) - 1$ in terms of power series. From this we see that the simplifications for the closed-form expression for the exponential in the quasi-geometric case are due to the fact that projection onto primitives behaves associatively, i.e. $\pi_1(x \mathbin{\widetilde{\sqcup\!\sqcup}} y) = \pi_1(\pi_1(x) \mathbin{\widetilde{\sqcup\!\sqcup}} \pi_1(y))$, which is not true on $\mathcal{H}_{\mathrm{CK}}$.

We must show commutativity of two rectangles, two parallelograms, and a rectangle. We begin by considering the alternative rectangle

$$\begin{tikzcd}
\bigsqcup(\mathcal{P}(U)) \arrow[d, two heads, "\bigsqcup(q)"'] & \mathcal{H}_{\mathrm{CK}}(U) \arrow[l, "\mathrm{Log}"'] \\
\bigsqcup(\bigodot(U)) & \widetilde{\sqcup\!\sqcup}(U) \arrow[l, "\widetilde{\mathrm{Log}}"'] \arrow[u, hook, "\widetilde{\iota}"']
\end{tikzcd} \tag{3.4}$$

where $q := \widetilde{\mathfrak{a}}|_{\mathcal{P}(U)}$ and $\widetilde{\iota}$ was defined in Theorem 2.19 (recall that this is not a Hopf morphism). $\bigsqcup(q)$ is a left inverse to $\bigsqcup(i)$ and $\widetilde{\iota}$ is a right inverse to $\widetilde{\mathfrak{a}}$ (as shown in Proposition 2.17), the two vertical maps appearing in the statement of the theorem. We show commutativity, using the first statement and again cofreeness

$$\begin{aligned}
\pi_1 \circ \widetilde{\mathrm{Log}} &= \pi_1 \circ \mathrm{e}_{\widetilde{\sqcup\!\sqcup}} \\
&= \bigsqcup(q) \circ \pi \circ \mathrm{e}_{\mathrm{CK}} \circ \widetilde{\iota} \\
&= \bigsqcup(q) \circ \pi \circ \mathrm{Log} \circ \widetilde{\iota} \\
&= \pi_1 \circ \bigsqcup(q) \circ \mathrm{Log} \circ \widetilde{\iota}
\end{aligned}$$

where we have used that

$$\begin{aligned}
\bigsqcup(q) \circ \pi \circ \mathrm{e}_{\mathrm{CK}} \circ \widetilde{\iota}(s_1 \otimes \cdots \otimes s_n) &= \frac{(-1)^n}{n} s_1 \odot \cdots \odot s_n \\
&= \pi_1 \circ \mathrm{e}_{\widetilde{\sqcup\!\sqcup}}(s_1 \otimes \cdots \otimes s_n)
\end{aligned}$$

by the same argument used in the first part of this proof. Incidentally, note how this proof will show that $\widetilde{\mathrm{Exp}}$ factors as a Hopf isomorphism in terms of Exp, but that $\widetilde{\mathrm{Log}}$ only factors as a coalgebra morphism in terms of Log. Now, returning to the rectangle in the statement, we have

$$\begin{aligned}
\widetilde{\mathrm{Exp}} &= \widetilde{\mathrm{Log}}^{-1} \\
&= (\bigsqcup(q) \circ \mathrm{Log} \circ \widetilde{\iota})^{-1} \\
&= \widetilde{\mathfrak{a}} \circ \mathrm{Exp} \circ \bigsqcup(i)
\end{aligned}$$

using partial inverses.

Returning to the other faces, the triangle on the left commutes by definition. The triangle on the right commutes since $\widetilde{\mathfrak{a}}_\odot$ restricted to U-decorated forests is equal to $\widetilde{\mathfrak{a}}$, since they are defined in the same way. Commutativity of the lower parallelogram is the defining property of $\mathfrak{a}\widetilde{\mathrm{Exp}}$. Finally, commutativity of the upper parallelogram follows from a diagram chase:

$$\begin{aligned}
\mathcal{H}_{\mathrm{CK}}(j) \circ \mathrm{Exp} \circ (\bigsqcup(i) \circ \widetilde{\mathfrak{a}}) &= \widetilde{\iota}_\odot \circ \widetilde{\mathrm{Exp}} \circ \mathfrak{a} \\
&= \mathcal{H}_{\mathrm{CK}}(j) \circ \widetilde{\iota} \circ \widetilde{\mathrm{Exp}} \circ \mathfrak{a} \\
&= \mathfrak{a}\widetilde{\mathrm{Exp}}
\end{aligned}$$

where $\widetilde{\iota}_\odot$ is the canonical right inverse to $\widetilde{\mathfrak{a}}_\odot$ defined in analogy with $\widetilde{\iota}$. □

Remark 3.10. Something slightly unsatisfying about the above proof is that it does not show how the expression for It is also possible to check commutativity of the rectangle in the main diagram more directly, by showing that $\widetilde{\mathrm{Exp}}$ satisfies the same

$$\mathrm{Exp}^1_n = -\sum_{1\le k<n} \mathrm{Exp}^1_k \circ \mathrm{Log}$$

♦

3.3. **Relationship with previous approaches.** As mentioned in the introduction, the problem of transforming a branched rough path into a geometric one over a larger space is not new. It was first tackled by Hairer and Kelly in [26] and later by Boedihardjo and Chevyrev [4]. In this subsection we review these approaches and compare them with ours. To summarise, the picture that will emerge is that our solution is quite distinct from that of [26], which is not canonical and relies on abstract existence results. [4] provides an algebraic solution, but still one that is not canonical, as it depends on a choice of a basis. The dual of our Exp can be identified as a specific case of their isomorphism; however, there is no guarantee that a generic basis considered by them will produce a *natural* isomorphism, as simple examples demonstrate. The approach of [26] is, in fact, partly algebraic. They consider a natural map

$$\psi\colon \mathcal{H}_{\mathsf{CK}} \to \sqcup\!\sqcup(\mathcal{T}), \qquad \psi(t) = t + \psi(t')\otimes t'',\ t\in\mathcal{T}$$

which is fixed by the further requirement that it be an algebra morphism (this is because $\mathcal{H}_{\mathsf{CK}}$ is free commutative over trees $\mathcal{T}$). The map ψ plays the role of our Log, since it maps $\mathcal{H}_{\mathsf{CK}}$ to a shuffle algebra over a larger space, and is injective, since it is injective on $\mathcal{T}$ as seen from the leading term. But, unlike Log, it is not surjective as immediately seen from dimensionality. Then $\mathcal{T}$-valued geometric rough paths $\bar{\mathbf{X}}$ satisfying $\psi^*(\bar{\mathbf{X}}) = \mathbf{X}$ are considered, and the authors prove a theorem analogous to Corollary 3.7 for such rough paths. Since ψ is not invertible, a recursive procedure for proving the existence of such a rough path is employed, and this is achieved via the non-constructive existence theorem of [37]. The lack of uniqueness of this rough path is expressed dually by the fact that there is redundancy in the choice of vector fields in the geometric RDE [26, Remark 5.9]. One may try to avert these difficulties by asking whether ψ admits a Hopf left inverse, which one could then use to define $\bar{\mathbf{X}}$ directly, but it is very unclear whether such a map should exist, especially with the extra requirement of naturality. In [4], the authors use the result, proved independently (in different contexts) in [17, Theorem 8.4] and [9, Corollary 6.3], that $\mathcal{T}(U)$ is free as a Lie algebra over some space $\mathcal{B}(U)$. Specifically, the latter can be summarised by stating that the upper triangle in the following diagram

$$\begin{tikzcd}
 & \underline{\mathsf{pre-Lie}} \arrow[d] \arrow[dd, bend left=60, "(\bigodot(\cdot),\circledast)"] \\
\underline{\mathsf{Vec}} \arrow[ur, "(\mathcal{T},\curvearrowleft)"] \arrow[r, "(\mathcal{T},[\cdot,\cdot]_\star)"] \arrow[d, "\mathcal{B}"] & \underline{\mathsf{Lie}} \arrow[d, "\mathcal{U}"'] \\
\underline{\mathsf{Vec}} \arrow[ur, "\mathcal{L}"] \arrow[r, "\bigotimes"'] & \underline{\mathsf{Alg}}
\end{tikzcd} \tag{3.5}$$

commutes, where $\bigotimes$ is the free algebra functor and $\mathcal{U}$ is the universal enveloping algebra functor. Commutativity of the lower triangle is the Poincaré-Birkhoff-Witt theorem, i.e. that the universal enveloping algebra of a free Lie algebra is free. Note that the cited article is about free pre-Lie algebras, and does not consider the Connes-Kreimer or Grossman-Larson Hopf algebras. The connection is made thanks to the fact that $\mathcal{T}$ with tree grafting $\curvearrowleft$ coincides with the free pre-Lie algebra functor and the functor of Grossman-Larson primitives, compatibly with their Lie bracket given as antisymmetrisation of $\star$: by Milnor-Moore it then follows that $\mathcal{H}_{\mathsf{GL}} = \mathcal{U}\circ\mathcal{T}$ which is thus equal to $\bigotimes\circ\mathcal{B}$ by commutativity of the above diagram. The Oudom-Guin theorem provides an explicit expression for the product in $\mathcal{H}_{\mathsf{GL}}$. This point of view is related to ours as follows: $\mathcal{B}$ (or at least a particular natural choice for it) coincides with $\mathrm{e}_{\mathsf{GL}}(\mathcal{Q})$. It is enough to show that $\mathcal{H}_{\mathsf{GL}}$ is free over $\mathrm{e}_{\mathsf{GL}}(\mathcal{Q})$, since e_{GL} is valued in $\mathrm{Prim}(\mathcal{H}_{\mathsf{GL}}) = \mathcal{T}$ (the Lie algebra associated to the free associative algebra is the canonical model for the free Lie algebra, so the upper triangle commutes). Then Theorem 3.4 can be read as stating that $\mathcal{H}_{\mathsf{CK}}$ is cofree over the projection $\pi\circ\mathrm{e}_{\mathsf{CK}}$, and thus dually $\mathcal{H}_{\mathsf{GL}}$ is free over the injection $\mathrm{e}_{\mathsf{GL}}\circ\pi^*$, i.e. over $\mathrm{e}_{\mathsf{GL}}(\mathcal{Q})$. We wish to stress the following point, which is central to the scope of this paper. There are uncountably many spaces $\mathcal{B}$ over which a free algebra, in our case $\mathcal{H}_{\mathsf{GL}}$, is free. Requiring that the isomorphism $\mathcal{H}_{\mathsf{GL}}(U) \cong \bigotimes(\mathcal{B}(U))$ be natural in U is a much more restrictive property. In [4, Theorem 2.3] the authors focus on finding such a basis composed of trees (as opposed to linear combinations of them), and in [4, §6] compute the explicit free basis of the truncated $\mathcal{H}^2_{\mathsf{GL}}$: having fixed a frame (ordered basis) of U, take $\mathcal{B}(U) = \{1,\ \bullet_\gamma,\ \overset{\bullet_\alpha}{\bullet_\beta}\}_{\gamma,\alpha\le\beta}$. The resulting isomorphism $\mathcal{H}_{\mathsf{GL}} \to \bigotimes(\mathcal{Q})$, however, is not natural, since any coordinate permutation changes the order of the basis. Moreover, order-2 truncation makes it possible to compare its dual with Hoffman's exponential, and the two do not coincide

$$\mathcal{H}_{\mathsf{CK}}(U) \ni \overset{\bullet_\alpha}{\bullet_\beta} \mapsto \bullet_\alpha\otimes\bullet_\beta + \delta_{\alpha\le\beta}\pi(\bullet_\alpha\bullet_\beta) \neq \bullet_\alpha\otimes\bullet_\beta + \tfrac{1}{2}\pi(\bullet_\alpha\bullet_\beta) = \widetilde{\mathrm{Exp}}(\overset{\bullet_\alpha}{\bullet_\beta})$$

In terms of stochastic integration and rough paths, this means that the geometric rough path associated to the Itô rough path via this isomorphism will not coincide with the Stratonovich rough path, rather its second order terms $\int_{s<u<v<t} \mathrm{d}X_u^\alpha \mathrm{d}X_v^\beta + \delta_{\alpha\leq\beta}[X]^{\alpha\beta}$ will depend on the particular order of the coordinates chosen. In fact, this example, to be compared with the next calculation, shows that bases constituted of single trees will never be natural (as this will immediately fail at order 2).

Example 3.11 ($\mathrm{e}_{\mathrm{GL}}(Q)^3$)**.** We compute the free basis of $\mathcal{H}_{\mathrm{GL}}$ up to level 3: this is a straightforward rearrangement after applying definitions of e_{GL} and π^* (or Example 1.23)

$$\mathrm{e}_{\mathrm{GL}}(Q)^3 = \mathrm{span}\left\{\bullet_\gamma,\ \begin{smallmatrix}\bullet_\beta\\ |\\ \bullet_\alpha\end{smallmatrix} + \begin{smallmatrix}\bullet_\alpha\\ |\\ \bullet_\beta\end{smallmatrix},\ {}^{\beta}\!\vee^{\gamma}_{\alpha} + {}^{\alpha}\!\vee^{\gamma}_{\beta} + {}^{\alpha}\!\vee^{\beta}_{\gamma},\ \begin{smallmatrix}\bullet_\alpha\\ |\\ \bullet_\beta\\ |\\ \bullet_\gamma\end{smallmatrix} + \begin{smallmatrix}\bullet_\beta\\ |\\ \bullet_\gamma\\ |\\ \bullet_\alpha\end{smallmatrix}\right\}$$

We note that this is a strict subspace of $\mathcal{T}^3$. ♦

The following two considerations further motivate our interest in considering natural isomorphisms as they relate to rough analysis: the first involves probability theory, while the second comes from differential geometry.

Remark 3.12 (Preservation of coordinate permutation-invariance in law)**.** Many rough paths considered in the literature (see [22, Part III]) involve a process with i.i.d. components w.r.t. a chosen basis B of the underlying space. If constructed "naturally", one may expect the resulting rough path terms to be invariant, in law, under coordinate permutation. By this we mean, if the rough path is branched, that $\mathbf{X}^{f} \sim \mathbf{X}^{\mathcal{H}(\sigma)(f)}$ where $\mathcal{H}(\sigma) \in \mathrm{Aut}(\mathcal{H}_{\mathrm{CK}})$ is the map induced by some $\sigma \in \mathbb{S}_B$ (*not* for σ a permutation of the vertices of f: the corresponding statement will of course not hold in general unless $\sigma \in \mathbb{S}_f$). Then, when transforming it to a geometric rough path using an isomorphism $\Phi\colon \mathcal{H}_{\mathrm{GL}} \to \bigotimes(\mathcal{B})$, we may wish for this property to continue to hold for the transformed rough path. This will hold if the transformation is natural, since for a word $w \in \bigotimes(\mathcal{P})$

$$\Phi(\mathbf{X})^{\bigotimes\circ\mathcal{P}(\sigma)(w)} = \langle\Phi^*(\bigotimes\circ\mathcal{P}(\sigma)(w)), \mathbf{X}\rangle = \langle\mathcal{H}_{\mathrm{CK}}(\sigma)(\Phi^*(w)), \mathbf{X}\rangle \sim \Phi(\mathbf{X})^{f}$$

but may not if it is not (as illustrated by the order-2 example above). ♦

Remark 3.13 (Itô-Stratonovich transformations on manifolds)**.** In [14] one of the authors introduced a definition of branched rough path on a manifold, and used it to study integrals and differential equations (the latter in the quasi-geometric case), using the framework of bracket extensions. Using the framework of this article, this definition would be available without invoking bracket extensions, cf. Remark 2.15: it involves considering a collection of branched rough paths $\mathbf{X}^i$ and an atlas indexed by the same set, requiring the compatibility condition $(\varphi_j^{-1}\circ\varphi_i)_*\mathbf{X}^i = \mathbf{X}^j$ on the overlap of the two charts, which hinges on defining the *pushforward* of a rough path through a smooth map, done earlier in the paper using a natural construction involving tree grafting (and done in [36] for geometric rough paths using ordered shuffles). If such a manifold-valued branched rough path is defined, we may ask about transforming it to a geometric one using an isomorphism Φ as above. This, however, will only be possible if such an isomorphism commutes with pushforwards

$$f_*\Phi(\mathbf{X}) = \Phi(f_*\mathbf{X}) \quad\Longrightarrow\quad \Phi(\mathbf{X}^j) = (\varphi_j^{-1}\circ\varphi_i)_*\Phi(\mathbf{X}^i)$$

While we leave an in-depth discussion of these topics for further work, we point out that this can only be expected to hold if Φ is natural. If it is not, the isomorphism will depend on the system of local coordinates chosen, i.e. it will not produce a manifold-valued geometric rough path. ♦

Remark 3.14 (Redundancy of the branched log-signature)**.** In analogy to the geometric case, given the signature Equation (2.8), it is natural to define the *log-signature* by $\log_\star \mathcal{S}(\mathbf{X})$, where

$$\log_\star(h) = \sum_{n\geq1}\frac{(-1)^{n-1}}{n}(h-1)^{\star n}$$

Using the character property of $\mathcal{S}(\mathbf{X})$, we can express log-signature coordinates in terms of signature coordinates using the Eulerian idempotent (which *is* linear):

$$\begin{aligned}\langle h, \log_\star \mathcal{S}(\mathbf{X})\rangle &= \langle \mathrm{e}(h), \mathcal{S}(\mathbf{X})\rangle\\ &= \sum_{n\geq1}\frac{(-1)^{n-1}}{n}\langle h^{(1)}\cdots h^{(n)}, \mathcal{S}(\mathbf{X})\rangle\\ &= \sum_{n\geq1}\frac{(-1)^{n-1}}{n}\langle h^{(1)}, \mathcal{S}(\mathbf{X})\rangle\cdots\langle h^{(n)}, \mathcal{S}(\mathbf{X})\rangle.\end{aligned} \tag{3.6}$$

The difference with the geometric case, however, is that while any truncated element in the free Lie algebra can be realised as the log-signature of a path (a piecewise linear path in fact, by Chow's theorem [22, p.140]), the strictness of the inclusion $\mathrm{e}_{\mathrm{GL}}(Q) \subsetneq \mathcal{T}$ tells us that not the whole branched log-signature is needed to recover $\mathbf{X}$. ♦

4. Applications to the one-dimensional case

4.1. **Kailath-Segall polynomials.** In this section we develop a generalization of the Kailath-Segall polynomials [51] to branched rough paths. The goal is to write the quantities representing the iterated integrals

$$\int_{s<u_1<\ldots<u_n<t} \mathrm{d}X^p_{u_1}\cdots\mathrm{d}X^p_{u_n}$$

with $p \in \mathcal{P}$, as polynomials in increments of X and its "higher-order variations". In fact, we will show that these polynomials are universal in some sense, because they arise as the image of the Eulerian idempotent in the free Hopf algebra over a sequence of symbols behaving like divided powers.

Let us consider a sequence of formal symbols $\mathbf{x} = (x_n : n \geq 1)$, and let $\mathbb{R}[\mathbf{x}]$ be the free polynomial algebra generated by them. We define a coproduct by setting

$$\Delta x_n = \sum_{j=0}^{n} x_j \otimes x_{n-j}$$

with the convention $x_0 \coloneqq 1$. It is clear that it is a commutative, cocommutative bialgebra. We give it a grading by $|x_n| = n$, therefore its also connected. Hence, it is a Hopf algebra.

It is immediate that for any $k > 1$ we have

$$\Delta^{(k-1)} x_n = \sum_{i_1+\cdots+i_k=n} x_{i_1} \otimes \cdots \otimes x_{i_k},$$

whence

$$\mathrm{e}(x_n) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum_{i_1+\cdots+i_k=n} x_{i_1}\cdots x_{i_k}.$$

Given the product is commutative, we may rearrange terms and obtain the perhaps better-known expression

$$P_n \coloneqq \mathrm{e}(x_n) = \sum_{a_1+2a_2+\cdots+na_n=n} (-1)^{a_1+\cdots+a_n-1}\frac{(a_1+\cdots+a_n-1)!}{a_1!\cdots a_n!} x_1^{a_1}\cdots x_n^{a_n} \tag{4.1}$$

Since we are in a cocommutative Hopf algebra, by Proposition 3.2, P_n is primitive and since it is also commutative, again by Proposition 3.2 the collection P_n is a set of algebra generators. In particular, we can rewrite x_n as a polynomial in these variables. In fact, since $\mathrm{id} = \exp(\mathrm{e})$ we may write

$$x_n = \sum_{k=1}^{n} \frac{1}{k!} \sum_{i_1+\cdots+i_k=n} P_{i_1}\cdots P_{i_k}.$$

Again, applying the same reordering as before we may write

$$x_n = \sum_{a_1+2a_2+\cdots+na_n=n} \frac{P_1^{a_1}}{a_1!}\cdots\frac{P_n^{a_n}}{a_n!}. \tag{4.2}$$

Expressions of this form in $\mathcal{H}_{\mathrm{CK}}$ were already known to Foissy [17].

We transfer these identities to $\mathcal{H}_{\mathrm{CK}}$. Working over an arbitrary vector space U as in the previous sections, and given $p \in \mathcal{P}$, let $\mathcal{H}_{\top(p)}$ denote the free commutative algebra generated by the elements $p^{\top n}$, $n \in \mathbb{N}$ in $\mathcal{H}_{\mathrm{CK}}$. A special case is $\mathcal{H}_{\top(\bullet_\gamma)}$ is the sub-Hopf algebra generated by all ladders with nodes decorated with γ. It is immediate that the mapping

$$\mathbb{R}[\mathbf{x}] \to \mathcal{H}_{\top(p)}, \quad x_n \mapsto p^{\top n} \tag{4.3}$$

is a Hopf isomorphism. Combining all of these simple observations immediately yields the following generalisation of [17, §9.1], in which ordinary ladders are replaced by the elements $p^{\top n}$:

Proposition 4.1 (Branched Kailath-Segall polynomials). *For $p \in \mathcal{P}$, $p^{\top n}$ satisfies the expression for x_n (4.2) where P_n is given by (4.1) with $p^{\top k}$ replacing x_k. Evaluating a branched rough path $\mathbf{X}$ on this identity expresses $\langle p^{\top n}, \mathbf{X}_{s,t}\rangle$ as a polynomial in the path increments $\langle P_k, \mathbf{X}_{s,t}\rangle$.*

Continuing to use P_n as in the above proposition, we are able to obtain a classical representation formula for iterated integrals of semimartingales as a special case. We note that this formula continues to hold for discontinuous semimartingales, i.e. processes not explicitly considered in the rest of this article; nonetheless, these same algebraic identities continue to hold. See for example [2,51,52]. Discontinuous semimartingales, such as general Lévy processes, are important examples in this regard, as they allow us to attain the identities in the fullest generality.

Corollary 4.2 (Classical Kailath-Segall polynomials for quasi-geometric rough paths)**.** *Let* $\mathbf{X}$ *be a one-dimensional quasi-geometric rough path. Then*

$$\langle \pi(\mathscr{r}_m)^{\top n}, \mathbf{X}_{s,t}\rangle = (-1)^n \sum_{a_1+2a_2+\dots+na_n=n} \frac{(-1)^{a_1+\dots+a_n}}{a_1!\cdot 2^{a_1}a_2!\cdots n^{a_n}a_n!}\langle \pi(\mathscr{r}_m), \mathbf{X}_{s,t}\rangle^{a_1}\cdots\langle \pi(\mathscr{r}_{mn}), \mathbf{X}_{s,t}\rangle^{a_n}$$

Proof. By Corollary 2.20, defining

$$\widetilde{P}_k := \frac{(-1)^{k-1}}{k}\pi(\mathscr{r}_k) \tag{4.4}$$

we have that $\langle P_k, \mathbf{X}\rangle = \langle \widetilde{P}_k, \mathbf{X}\rangle$ since $P_k = \pi(P_k)$ and by Corollary 2.20. We conclude by replacing P_k with $\widetilde{P}_k$ in (4.2). □

Remark 4.3. The classical Kailath-Segall polynomials are written in terms of the variables $\pi(\mathscr{r}_k)$. One might ask for the branched Kailath-Segall polynomials for $p^{\top n}$ to be expressed explicitly in terms of images of π. Following (2.22), this can be interpreted as saying that for geometric rough paths, $\mathbf{X}^{\ell_n} = X^n/n!$, for quasi-geometric rough paths Corollary 4.2 holds, and for more general branched rough paths, $p^{\top n}$ is a polynomial in further images of π. These can be obtained as in Corollary 4.2, i.e. by observing that $P_n \in \mathcal{P}$, $P_n = \pi(P_n)$, i.e.

$$P_n = \sum_{a_1+2a_2+\dots+na_n=n} (-1)^{a_1+\dots+a_n-1}\frac{(a_1+\dots+a_n-1)!}{a_1!\cdots a_n!}\pi((p^{\top 1})^{a_1}\cdots(p^{\top n})^{a_n}). \tag{4.5}$$

and substituting this expression into Proposition 4.1. There are a couple of issues with this substitution, however. First of all, if one wishes for the variables of these polynomials to be $\pi(\ell)$ with $\ell \in \mathcal{F}$, the above expression needs to be re-worked and the polynomials are no longer universal in these variables, i.e. they will depend on p (or rather its "shape"). Moreover, since $\{\pi(\ell) \mid \ell \in \mathcal{F}\}$ is not an independent set, such polynomials are not unique (although there will still be a natural choice that emerges from the above substitution) as they are instead in the quasi-geometric case, thanks to Lemma 2.18. Nevertheless, such polynomials can in principle be computed, in particular when $p = \bullet_\gamma$, in which case they represent the branched corrections to the formula $\mathbf{X}^{\ell_n} = X^n/n!$ valid in the geometric case. ♦

We now return to the general setting to obtain a branched analogue of what is called the Kailath-Segall formula, i.e. a recursive expression for (4.2). Let us recall the definition of the Dynkin operator $\mathcal{D}\colon H \longrightarrow H$ over a graded Hopf algebra H:

$$\mathcal{D}(h) := |h_{(2)}|\mathcal{S}(h_{(1)})h_{(2)} \tag{4.6}$$

where $\mathcal{S}\colon H \longrightarrow H$ is the antipode, which, we recall, satisfies Takeuchi's formula [55]:

$$\mathcal{S}(h) = \sum_{n\geq 1}(-1)^n h^{(1)}\cdots h^{(n)} \tag{4.7}$$

The following lemma relates the Dynkin and Eulerian idempotents in a very special case.

Lemma 4.4. *If* H *is commutative and cocommutative,* $\mathcal{D}(h) = |h|\mathrm{e}(h)$.

Proof. Combining (4.6) and (4.7) we have

$$\begin{aligned}
\mathcal{D}(h) &= \sum_{n\geq 1}(-1)^n|h_{(2)}|(h_{(1)})^{(1)}\cdots(h_{(1)})^{(n)}h_{(2)}\\
&= \sum_{n\geq 1}(-1)^{n-1}|h^{(n)}|h^{(1)}\cdots h^{(n)}\\
&= \sum_{n\geq 1}(-1)^{n-1}\frac{|h^{(1)}|+\ldots+|h^{(n)}|}{n}h^{(1)}\ldots h^{(n)}\\
&= |h|\mathrm{e}(h)
\end{aligned}$$

where we have used that for each n the expression $\sum_{(h)} h^{(\sigma(1))}\cdots h^{(\sigma(n))} = h^{(1)}\cdots h^{(n)}$ does not depend on $\sigma \in \mathbb{S}_n$, which follows from commutativity and cocommutativity combined. □

Formula (4.6) can actually be inverted, since the antipode is the convolutional inverse of the identity map in H, that is, $\mathcal{S} * \mathrm{id} = \mathrm{id} * \mathcal{S} = 1 \circ \varepsilon$ for $|h| > 0$. We readily obtain

$$\mathrm{id} * \mathcal{D} = \mathrm{id} * \mathcal{S} * |\cdot| = |\cdot|$$

In the particular case of $\mathbb{R}[\mathbf{x}]$, after applying this identity and Lemma 4.4, we get the following recursion:

$$nx_n = \sum_{k=1}^{n} kP_k x_{n-k}.$$

Taking again the isomorphism (4.3) (and reasoning as in (4.4)) we have thus proved the following result.

Theorem 4.5 (Branched Kailath-Segall formula)**.** *Let* $\mathbf{X}$ *be a one-dimensional branched rough path*

$$\langle p^{\top n}, \mathbf{X}_{s,t}\rangle = \sum_{k=1}^{n} \frac{k}{n} \langle P_k, \mathbf{X}_{s,t}\rangle \langle p^{\top(n-k)}, \mathbf{X}_{s,t}\rangle \,. \tag{4.8}$$

As special case when $\mathbf{X}$ *is quasi-geometric we recover the usual Kailath-Segall formula [51, Theorem 1]*

$$\langle \pi(\mathscr{v}_m)^{\top n}, \mathbf{X}_{s,t}\rangle = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{n} \langle \pi(\mathscr{v}_{mk}), \mathbf{X}_{s,t}\rangle \langle \pi(\mathscr{v}_m)^{\top(n-k)}, \mathbf{X}_{s,t}\rangle \,. \tag{4.9}$$

Remark 4.6. It is clear by symmetrising that the results of this chapter extend from $\mathcal{H}_{\top(\mathcal{P})}$ to the Hopf algebra $\mathcal{H}_{\top}$ generated as an algebra by the elements

$$\frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} p_{\sigma(1)} \top \ldots \top p_{\sigma(n)}$$

as $p_1, \ldots, p_n$ range in $\mathcal{P}$. $\mathcal{H}_{\top}$ is the largest cocommutative sub-Hopf algebra of $\mathcal{H}_{\mathrm{CK}}$ [19, Lemma 26], and hence the largest whose elements can be written as polynomials in primitives. ♦

4.2. **Construction of general branched rough paths: stochastic examples at order 4.** In this last subsection, we propose a general method for building examples of branched rough paths, and provide stochastic examples. The recipe is as follows:

1 Depending on regularity and number of labels (here we continue to work in the scalar setting, and at order 4), identify a free commutative generating set $\mathcal{A}$ of $\mathcal{H}_{\mathrm{CK}}$ constituted of single words, under the delimiter $\top$, in the alphabet $\mathcal{P}$ (by this we mean single words—similar to Lyndon words—not their linear combinations);
2 Assign to each primitive element p with $|p| \leq \lfloor \rho \rfloor$ and which appears as a letter in one of the elements of $\mathcal{A}$, a path X^p;
3 Using some kind of integration theory, postulate the values of $\mathbf{X}^a$ for $a \in \mathcal{A}$.
4 We can now check, using Corollary 2.20, whether $\mathbf{X}$ is geometric or quasi-geometric. This may involve computing $\mathbf{X}$ on elements not considered in its definition.

Since $\mathcal{H}_{\mathrm{CK}}$ is free-commutative over the span of $\mathcal{A}$, this uniquely identifies a branched rough path. Cofreeness, and in particular the operation $\top$, makes it possible to restrict the definition of $\mathbf{X}$ to a certain number of linear iterated integrals, for which the Chen identity is simpler to state and verify than its branched counterpart. Following these three points, we will be able to exhibit examples of rough paths which are *truly branched*, i.e. fail the quasi-geometricity test. We use this terminology in part to echo the notion of "truly rough" [21, Definition 6.3]: the two share the feature of the rough path needing to be considered as member of a space of "worse-behaved" objects.

We conclude the article by summarising the main results of the previous theorem in the explicit case of one dimensional branched rough paths of regularity $\rho \in (4, 5)$. The choice to consider this case is due to two simple reasons: on the one hand, as already noted in [58], it is possible to explicitly compute a basis of $\mathcal{P}$ and $\mathcal{Q}$ up to order four and three so that we can express all the formulas of the previous sections according to a specific choice of coordinates of any branched rough path. On the other hand, at level four, as described in Corollary 2.20 the branched rough paths strictly contain the quasi-geometric rough paths, and consequently the variable change formulae we will present will be new. Finally, we will be able to relate the results obtained to the construction of a family of branched rough paths over the trajectories of a fractional Brownian motion with Hurst index $1/4$. We begin by an explicit description of the truncated space of primitives. Note we truncate $\mathcal{P}$ and $\mathcal{Q}$ at different degrees since these are the minimal degrees required for the controlledness condition to hold, and thus enough to define rough integration.

Proposition 4.7. *The vector spaces* $\mathcal{P}^4 = \bigoplus_{n=1}^{4} \mathcal{P}^{(n)}$ *and* $\mathcal{Q}^3 = \bigoplus_{n=1}^{3} \mathcal{Q}^{(n)}$ *are generated respectively by the elements*

$$\{\bullet,\ \pi(\bullet\bullet),\pi(\bullet\bullet\bullet),\ \pi(\bullet\bullet\bullet\bullet),\pi(\begin{smallmatrix}\bullet&\bullet\\ |&|\\ \bullet&\bullet\end{smallmatrix})\}, \quad \{\bullet,\ \bullet\bullet,\bullet\bullet\bullet\},$$

which constitute a basis.

Proof. From the definition of the π operator, we can indeed compute $\pi(f)$ for any forest f of cardinality smaller than 4, obtaining that π is non zero over the forests (c.f. [58, Appendix A])

$$\{\bullet,\ \bullet\bullet,\ \bullet\bullet\bullet,\ \bullet\bullet\bullet\bullet,\ [\bullet][\bullet],\ \bullet\bullet[\bullet]\}.$$

Using the result of Lemma 2.18, the identity $\pi([\bullet][\bullet]) = \pi(\bullet\bullet[\bullet])$ and remarking that $\pi([\bullet][\bullet])$ is linearly independent from $\pi(\bullet\bullet\bullet\bullet)$, we can easily conclude. The basis of Q was already computed in Example 1.23. □

We can thus represent elements in $\mathcal{H}^4_{\mathrm{CK}}$ and $\mathcal{H}^3_{\mathrm{GL}}$ in the bases adapted to the (co)algebra structure:

$$\begin{aligned}\mathcal{A}_{\mathcal{P}} = \Big\{&\bullet,\ \pi(\bullet\bullet),\ [\bullet],\ [[\bullet]],\ \pi(\bullet\bullet)\top\bullet,\ \bullet\top\pi(\bullet\bullet),\ \pi(\bullet\bullet\bullet),\ [[[\bullet]]],\ \pi(\bullet\bullet)\top\bullet\top\bullet,\ \bullet\top\pi(\bullet\bullet)\top\bullet,\\ &\bullet\top\bullet\top\pi(\bullet\bullet),\ \pi(\bullet\bullet)\top\pi(\bullet\bullet),\ \pi(\bullet\bullet\bullet)\top\bullet,\ \bullet\top\pi(\bullet\bullet\bullet),\ \pi(\bullet\bullet\bullet\bullet),\ \pi([\bullet][\bullet])\Big\}\\ \mathcal{A}_{Q} = \Big\{&\bullet,\ \bullet\bullet,\ \bullet\star\bullet,\ \bullet\star\bullet\star\bullet,\ \bullet\bullet\bullet,\ \bullet\star\bullet\bullet,\ \bullet\bullet\star\bullet\Big\}.\end{aligned}$$

The set $\mathcal{A}_{\mathcal{P}}$ is a vector space basis of $\mathcal{P}$, but still contains polynomial relations. Although we already have such a basis of the algebra in the form of trees, we wish for the generators to be expressed in terms of $\top$ and π in a way reminiscent of a Lyndon basis.

Proposition 4.8. *Any element of $\mathcal{H}^4_{\mathrm{CK}}$ can be uniquely written as a polynomial w.r.t. the forest product with respect to the new variables*

$$\begin{aligned}\mathcal{A} := \Big\{&\bullet,\ \pi(\bullet\bullet),\ \pi(\bullet\bullet\bullet),\ \bullet\top\pi(\bullet\bullet),\ \pi(\bullet\bullet\bullet)\top\bullet\\ &\pi(\bullet\bullet)\top\pi(\bullet\bullet),\ \bullet\top\pi(\bullet\bullet\bullet),\ \bullet\top\bullet\top\pi(\bullet\bullet)\Big\}.\end{aligned} \tag{4.10}$$

Proof. Denoting by $\mathcal{T}^4$ the set of combinatorial trees of cardinality smaller than 4 and by $\mathcal{L}$ the set in (4.10), to prove the result it is sufficient to show that the free commutative algebra over $\mathcal{T}^4$ coincide with the free commutative algebra over $\mathcal{L}$. Since $\mathcal{T}^4$ and $\mathcal{L}$ have the same cardinality and the elements $\mathcal{L}$ can be expressed as polynomials in $\mathcal{T}^4$ using the operation $\top$, to conclude the identification it is sufficient to show that we can write every element of $\mathcal{T}^4$ as a forest polynomial in elements of $\mathcal{L}$. By using the product formula (1.17) combined with the operation $\top$ we can check the identities

$$\begin{aligned}
[\bullet] &= \frac{1}{2}(\bullet\bullet - \pi(\bullet\bullet))\\
[[\bullet]] &= \frac{1}{6}(\bullet\bullet\bullet - 3\,\pi(\bullet\bullet)\bullet + 2\,\pi(\bullet\bullet\bullet))\\
[\bullet\bullet] &= \frac{1}{3}(\bullet\bullet\bullet - 3\bullet\top\pi(\bullet\bullet) - \pi(\bullet\bullet\bullet))\\
[[[\bullet]]] &= \frac{1}{24}(\bullet\bullet\bullet\bullet + 4\,\pi(\bullet\bullet\bullet)\bullet + \pi(\bullet\bullet)\pi(\bullet\bullet) + 4\,\pi(\bullet\bullet)\top\pi(\bullet\bullet)\\
&\qquad - 6\bullet\bullet\pi(\bullet\bullet) + 4\,\pi(\bullet\bullet\bullet)\top\bullet + 4\bullet\top\pi(\bullet\bullet\bullet))\\
[[\bullet]\bullet] &= \frac{1}{24}(3\bullet\bullet\bullet\bullet - 6\bullet\bullet\pi(\bullet\bullet) - \pi(\bullet\bullet)\pi(\bullet\bullet) - 4\,\pi(\bullet\bullet)\top\pi(\bullet\bullet)\\
&\qquad - 4\,\pi(\bullet\bullet\bullet)\top\bullet - 4\bullet\top\pi(\bullet\bullet\bullet) + 4\,\pi(\bullet\bullet\bullet)\bullet - 24\bullet\top\bullet\top\pi(\bullet\bullet))\\
[[\bullet\bullet]] &= \frac{1}{12}(\bullet\bullet\bullet\bullet - 4\,\pi(\bullet\bullet\bullet)\bullet + 3\pi(\bullet\bullet)\pi(\bullet\bullet) + 12\bullet\top\bullet\top\pi(\bullet\bullet) - 12\bullet\top\pi(\bullet\bullet)\bullet + 12\bullet\top\pi(\bullet\bullet\bullet))
\end{aligned}$$

□

It follows immediately that characters on $\mathcal{H}^4_{\mathrm{CK}}$ are uniquely determined by their values on the basis $\mathcal{B}$.

Corollary 4.9. *Any given function $\mathbf{X}\colon \Delta_T \to \mathbb{R}^{\mathcal{B}}$ satisfying Chen property and the regularity property in Definition 2.1 for some control ω and $p \in (4,5)$ uniquely extends to a p-branched rough path $\mathbf{X}$ controlled by ω.*

Proof. The regularity condition is obviously satisfied by the linear extension $\mathbf{X}\colon \Delta_T \to \mathcal{H}^4_{\mathrm{GL}}$, and the character property follows from the fact that $\mathcal{B}$ is an algebraically independent set. Thus, we only need to check that Chen's property is preserved. This follows from the fact that $\mathbf{X}_{s,t}$ and $\mathbf{X}_{s,u} * \mathbf{X}_{u,t}$ are two characters agreeing on $\mathcal{B}$. □

Remark 4.10. Although not very clear from an algebraic point of view, the possibility of defining a branched rough path $\mathbf{X}$ from the elements in (4.10) allows a branched rough path to be constructed almost as if it were geometric. In fact, starting from two primitive elements $p, q \in \mathcal{P}^4$ we can define $\langle p \top q, \mathbf{X}_{s,t}\rangle$ directly as an iterated integral using stochastic integration as below. ♦

For any $\mathbf{H} \in \mathscr{D}^{\rho}_{\mathbf{X}}$ with $\rho \in (1/5, 1/4)$ all the possible rough integrals are given in coordinates by the following expressions

$$\begin{aligned}
\int_s^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet)} \approx{} & \langle \mathbf{H}_s, 1^*\rangle\langle \bullet, \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet\rangle\langle \begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix}, \mathbf{X}_{s,t}\rangle + \frac{1}{2}\langle \mathbf{H}_s, \bullet\bullet\rangle\langle \pi(\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet \star \bullet\rangle\langle \begin{smallmatrix}\bullet\\|\\\bullet\\|\\\bullet\end{smallmatrix}, \mathbf{X}_{s,t}\rangle \\
& + \frac{1}{6}\langle \mathbf{H}_s, \bullet\bullet\bullet\rangle\langle \pi(\bullet\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle + \frac{1}{2}\langle \mathbf{H}_s, \bullet\bullet \star \bullet\rangle\langle \pi(\bullet\bullet) \top \bullet \top \bullet, \mathbf{X}_{s,t}\rangle \\
& + \frac{1}{2}\langle \mathbf{H}_s, \bullet \star \bullet\bullet\rangle\langle \bullet \top \pi(\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet \star \bullet \star \bullet\rangle\langle \begin{smallmatrix}\bullet\\|\\\bullet\\|\\\bullet\\|\\\bullet\end{smallmatrix}, \mathbf{X}_{s,t}\rangle, \\
\int_s^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} \approx{} & \langle \mathbf{H}_s, 1^*\rangle\langle \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet\rangle\langle \bullet \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle \\
& + \frac{1}{2}\langle \mathbf{H}_s, \bullet\bullet\rangle\langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet \star \bullet\rangle\langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle, \\
\int_s^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet\bullet)} \approx{} & \langle \mathbf{H}_s, 1^*\rangle\langle \pi(\bullet\bullet\bullet), \mathbf{X}_{s,t}\rangle + \langle \mathbf{H}_s, \bullet\rangle\langle \bullet \top \pi(\bullet\bullet\bullet), \mathbf{X}_{s,t}\rangle, \\
\int_s^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet\bullet\bullet)} \approx{} & \langle \mathbf{H}_s, 1^*\rangle\langle \pi(\bullet\bullet\bullet\bullet), \mathbf{X}_{s,t}\rangle, \\
\int_s^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix}\begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix})} \approx{} & \langle \mathbf{H}_s, 1^*\rangle\langle \pi(\begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix}\begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix}), \mathbf{X}_{s,t}\rangle.
\end{aligned}$$

We turn our attention to the construction some explicit examples of stochastic rough paths defined over one-dimensional fractional Brownian motion X of Hurst parameter $H = 1/4$ defined over a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We refer for instance to [41, Chap. 2] for the definition of this process.

As it is known in the literature (see e.g. [50]) X has a.s. infinite quadratic variation, that is, for any $t \in [0, T]$ and any sequence of partitions of Π_n of $[0, t]$ such that $|\Pi_n| \to 0$ the random variable

$$[X]^n_t := \sum_{[u,v] \in \Pi_n} (X_v - X_u)^2$$

diverges as $n \to \infty$ in $L^1(\mathbb{P})$. This property is also shared by any fractional Brownian motion with Hurst parameter $0 < H < 1/2$.

However, in the special case $H = 1/4$, taking a uniform partition of $[0, t]$ one can establish a weaker convergence by considering the renormalised quadratic variation

$$\overline{[X]}^n_t := [X]^n_t - \mathbb{E}[X]^n_t = \sqrt{\frac{t}{n}} \sum_{k=0}^{n-1} \left(\sqrt{\frac{n}{t}} (X_{\frac{(k+1)t}{n}} - X_{\frac{kt}{n}})^2 - 1 \right) \tag{4.11}$$

It can be shown [42, Theorem 1.1] that

$$\overline{[X]}^n_t \xrightarrow{(d)} C_{1/4} W_t \quad (n \to +\infty)$$

for an independent standard Brownian motion W and $C_{1/4} \approx 1535$ a fixed constant.

Even if the resulting convergence is weaker than the usual convergence in rough path metric, from the point of view of rough analysis it is still interesting to consider if the presence of an independent Brownian motion W (we drop the constant) at the level of the quadratic variation allows us to define a family of branched rough path using purely stochastic methods. Using the elementary identity

$$\sum_{[u,v] \in \Pi_n} (X_v - X_u)^2 = X_t^2 - 2 \sum_{[u,v] \in \Pi_n} X_u (X_v - X_u)$$

any limit of the renormalised quadratic variation can be taken as postulating the value of

$$\langle \bullet\, \bullet - 2 \begin{smallmatrix}\bullet\\|\\\bullet\end{smallmatrix}, \mathbf{X}_{0,t}\rangle = \langle \pi(\bullet\bullet), \mathbf{X}_{0,t}\rangle.$$

By Theorem 4.8, this condition is obviously not sufficient to identify uniquely a rough path over X but the possibility to define a canonical notion of stochastic integral

$$\int_s^t (X_r - X_s)\,\mathrm{d}W_r$$

reduces the freedom to only four different possible lifts related to Itô and Stratonovich integration.

Theorem 4.11. *Up to modifications there exists 4 different ρ branched rough paths $\mathbf{X}$ over X with $\rho \in (4,5)$ satisfying the compatibility conditions*

$$\begin{aligned} \langle \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle &= (W_t - W_s), \quad \langle \bullet \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle = \int_s^t (X_r - X_s)\mathrm{d}W_r, \\ \langle \pi(\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle &= (X_t - X_s)(W_t - W_s) - \int_s^t (X_r - X_s)\mathrm{d}W_r. \end{aligned} \tag{4.12}$$

the trivial relations $\langle \pi(\bullet\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle = \langle \bullet \top \pi(\bullet\bullet\bullet), \mathbf{X}_{s,t}\rangle = 0$ and interpreting $\langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle$, $\langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}_{s,t}\rangle$ respectively as an Itô or a Stratonovich integrals. We call them respectively the Stratonovich-Stratonovich, Itô-Itô, Itô-Stratonovich and Stratonovich-Itô branched rough paths, which we will denote by $\mathbf{X}^{S,S}$, $\mathbf{X}^{I,I}$, $\mathbf{X}^{S,I}$, $\mathbf{X}^{I,S}$.

Remark 4.12. We simply explain the reasons that allowed us to set up relations (4.12) and the trivial relations involving $\pi(\bullet\bullet\bullet)$. Due to the fact that X and W are independent, one has zero covariation among them and any Itô type discretisation of $\langle \pi(\bullet\bullet) \top \bullet, \mathbf{X}_{s,t}\rangle$ has to coincide with (4.12). However, from the algebraic identity

$$\pi(\bullet\bullet)\bullet = \pi(\bullet\bullet) \top \bullet + \bullet \top \pi(\bullet\bullet) + \pi(\bullet\bullet\bullet)$$

we deduce that $\langle \pi(\bullet\bullet\bullet), \mathbf{X}_{s,t}\rangle = 0$. The other relations in the statement follows in a similar fashion.

Using the symbols d^* and d to denote Stratonovich and Itô integration the last condition reads in coordinates

$$\begin{aligned} \langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}^{I,I}_{s,t}\rangle &= \langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}^{I,S}_{s,t}\rangle = \int_s^t \frac{(X_r - X_s)^2 - (W_r - W_s)}{2}\,\mathrm{d}W_r, \\ \langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}^{S,I}_{s,t}\rangle &= \langle \bullet \top \bullet \top \pi(\bullet\bullet), \mathbf{X}^{S,S}_{s,t}\rangle = \int_s^t \frac{(X_r - X_s)^2 - (W_r - W_s)}{2}\circ\mathrm{d}W_r, \\ \langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}^{I,S}_{s,t}\rangle &= \langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}^{S,S}_{s,t}\rangle = \int_s^t (W_r - W_s)\circ\mathrm{d}W_r = \frac{(W_t - W_s)^2}{2}, \\ \langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}^{S,I}_{s,t}\rangle &= \langle \pi(\bullet\bullet) \top \pi(\bullet\bullet), \mathbf{X}^{I,I}_{s,t}\rangle = \int_s^t (W_r - W_s)\,\mathrm{d}W_r. \end{aligned}$$

♦

Proof of Theorem 4.11. Since the conditions in the statement define uniquely a map $\mathbf{X}\colon \Delta_T \to \mathbb{R}^{\mathcal{B}}$, thanks to the Corollary 4.9 it is sufficient to show that $\mathbf{X}$ satisfies the regularity assumption a.s. and the Chen property. Concerning the Chen identity, this follows trivially by the fact that the components $\mathbf{X}$ which are not related to primitive elements are given by iterated stochastic integrals which satisfy the usual Chen identities by construction.

Concerning the regularity in ρ-variation of all four possible models, it is sufficient to prove that the two-parameter process

$$(s,t) \to \left(X_t - X_s, W_t - W_s, \int_s^t (X_r - X_s)\,\mathrm{d}W_r, \int_s^t (X_r - X_s)^2\,\mathrm{d}W_r\right)$$

admits a modification whose components have respectively a.s. finite $1/\rho$, $2/\rho$, $3/\rho$ and $4/\rho$ Hölder norm with $\rho \in (4,5)$. For the first two components, we apply Kolmogorov's continuity theorem to the processes X and W, thanks to the fact that

$$\mathbb{E}|X_t - X_s|^p = C|t-s|^{\frac{p}{4}}, \quad \mathbb{E}|W_t - W_s|^p = C'|t-s|^{\frac{p}{2}}$$

for any $p \geq 1$ and some constants $C, C' > 0$. We obtain modifications of X and W (which we will denote in the same way) with the desired regularity properties and, in addition their Hölder seminorm are also q-integrable for any $q \geq 4$. Passing to the stochastic integrals we only show how to prove that the third component is $3/\rho$-Hölder, since the last component follows similar reasoning. By applying the Burkholder-Davis-Gundy inequality we obtain for any $p \geq 1$ the inequality

$$\mathbb{E}\left|\int_s^t (X_r - X_s)\,\mathrm{d}W_r\right|^{2p} \lesssim \left|\int_s^t \mathbb{E}|X_r - X_s|^2\,\mathrm{d}r\right|^p \lesssim |t-s|^{\frac{3p}{2}},$$

(we adopt the notation $\lesssim$ to denote inequality up to a constant). We now apply [44, Lemma 2.1] to $R_{s,t} := \int_s^t (X_r - X_s)\,\mathrm{d}W_r$ to get an estimate for its Hölder regularity. Indeed, we note that $\delta R_{sut} := R_{s,t} - R_{s,u} - R_{u,t} = (X_u - X_s)(W_t - W_u)$, so that

$$\sup_{(s,t)\in\Delta_T} \frac{|\delta R_{sut}|}{|t-u|^{3/p}} \lesssim \|X\|_{1/p}\|W\|_{2/p}.$$

We then obtain for any $(s,t)\in\Delta_T$, $s\neq t$

$$\frac{|R_{s,t}|}{|t-s|^{3/p}} \lesssim \left(\int_{\Delta_T} \frac{|R_{uv}|^{2p}}{|u-v|^{6p/p+4}}\,\mathrm{d}u\mathrm{d}v\right)^{1/2p} + \|X\|_{1/p}\|W\|_{2/p}$$

where $\|\cdot\|_\alpha$ denotes the α-Hölder seminorm of a path. Since the right-hand side of this inequality does not depend on (s,t), the estimate still holds when taking supremum of the left-hand side over the set $D_N := 2^{-N}\mathbb{N}\times 2^{-N}\mathbb{N}\cap\Delta_T$. By taking p sufficiently large so that

$$\mathbb{E}\int_{\Delta_T} \frac{|R_{uv}|^{2p}}{|u-v|^{6p/p+4}}\,\mathrm{d}u\mathrm{d}v < \infty,$$

which is possible since $1/4 - 1/p > 0$, we obtain that the sequence of random variables

$$Z_N = \sup_{(s,t)\in D_N} \frac{|R_{s,t}|}{|t-s|^{3/p}}$$

is not only integrable but also uniformly integrable ,and converges to a random variable Z_∞ with finite moment of order $2p$. Since D_N is a monotone sequence of increasing sets whose limit D_∞ is a dense and countable subset of Δ_T, up to modifications the process $(s,t)\to\int_s^t (X_r - X_s)\,\mathrm{d}W_r$ admits a modification which is a.s. $3/p$ Hölder with also integrable Hölder seminorm. □

Remark 4.13. From the algebraic identity

$$\pi(\mathrm{I}\,\mathrm{I}) = \frac{1}{3}(\pi(\bullet\bullet\bullet)\top\bullet + \bullet\top\pi(\bullet\bullet\bullet) + \pi(\bullet\bullet)\,\pi(\bullet\bullet) - 2\pi(\bullet\bullet)\top\pi(\bullet\bullet) - \pi(\bullet\bullet\bullet)\,\bullet)$$

it follows that $\mathbf{X}^{I,I}$ and $\mathbf{X}^{S,I}$ satisfy the property

$$\langle\pi(\mathrm{I}\,\mathrm{I}), \mathbf{X}^{S,I}_{s,t}\rangle = \langle\pi(\mathrm{I}\,\mathrm{I}), \mathbf{X}^{I,I}_{s,t}\rangle = \frac{1}{3}\left((W_t - W_s)^2 - 2\int_s^t (W_r - W_s)\,\mathrm{d}W_r\right) = \frac{(t-s)}{3}.$$

Using Corollary 2.20 we immediately deduce that $\mathbf{X}^{I,I}$ and $\mathbf{X}^{S,I}$ are branched rough path which do not arise from quasi-arborification of a quasi-geometric rough path, making them an example of a truly branched rough path arising from a stochastic process. ♦

Combining the definition of the different branched rough paths in Theorem 4.11 we can also identify some rough integrals with other classical objects. In particular it follows from their definition that

$$\int_s^t \mathbf{H}_r\,\mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet\bullet)} = 0$$

for $\mathbf{X} = \mathbf{X}^{I,S}, \mathbf{X}^{S,S}, \mathbf{X}^{I,I}$ and $\mathbf{X}^{S,I}$. Moreover by using the properties of Riemann sums one has also

$$\int_s^t \mathbf{H}_r\,\mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet\bullet\bullet)} = \int_s^t \langle\mathbf{H}_r, 1^*\rangle\,\mathrm{d}r, \qquad \int_s^t \mathbf{H}_r\,\mathrm{d}\mathbf{X}_r^{\pi(\mathrm{I}\,\mathrm{I})} = \frac{1}{3}\int_s^t \langle\mathbf{H}_r, 1^*\rangle\,\mathrm{d}r$$

when $\mathbf{X}$ equals $\mathbf{X}^{I,I}$ or $\mathbf{X}^{S,I}$, and zero in the other two cases. Concerning the integral with respect to $\pi(\bullet\bullet)$ we have a specific identification of the rough integrals with some particular linear combination of stochastic integrals and classical integrals.

Proposition 4.14. *Let $p\in(1/5, 1/4)$. Suppose that $\mathbf{H}\in\mathscr{D}^p_{\mathbf{X}}$ a.s. is adapted to the joint filtration of X and W. The following identities holds a.s.:*

- *when $\mathbf{X} = \mathbf{X}^{I,I}$*

$$\int_0^t \mathbf{H}_r\,\mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \int_0^t \langle\mathbf{H}_r, 1^*\rangle\,\mathrm{d}W_r\,, \tag{4.13}$$

- *when $\mathbf{X} = \mathbf{X}^{I,S}$*

$$\int_0^t \mathbf{H}_r\,\mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \int_0^t \langle\mathbf{H}_r, 1^*\rangle\,\mathrm{d}W_r + \frac{1}{4}\int_0^t \langle\mathbf{H}_r, \bullet\bullet\rangle\,\mathrm{d}r, \tag{4.14}$$

- *when* $\mathbf{X} = \mathbf{X}^{S,I}$
$$\int_0^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \int_0^t \langle \mathbf{H}_r, 1^* \rangle \, \mathrm{d}W_r - \frac{1}{4}\int_0^t \langle \mathbf{H}_r, \bullet \star \bullet \rangle \, \mathrm{d}r, \tag{4.15}$$
- *when* $\mathbf{X} = \mathbf{X}^{S,S}$
$$\int_0^t \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \int_0^t \langle \mathbf{H}_r, 1^* \rangle \, \mathrm{d}W_r - \frac{1}{4}\int_0^t \langle \mathbf{H}_r, \substack{\bullet \\ | \\ \bullet} \rangle \, \mathrm{d}r. \tag{4.16}$$

Proof. It is sufficient to show the result when $t = 1$. We begin by proving the identification (4.13) using an argument similar to [21, Prop 5.1]. By definition of rough integral when $\mathbf{X} = \mathbf{X}^{I,I}$ this equals the a.s. limit
$$\int_0^1 \mathbf{H}_r \, \mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \lim_{n \to +\infty} \sum_{[s,t] \in \Pi_n} \mathbf{Z}_{s,t}^{I,I}$$
where $\mathbf{Z}_{s,t}^{I,I}$ is explicitly given by
$$\begin{aligned}\mathbf{Z}_{s,t}^{I,I} &= \langle \mathbf{H}_s, 1^* \rangle (W_t - W_s) + \langle \mathbf{H}_s, \bullet \rangle \int_s^t (X_r - X_s) \, \mathrm{d}W_r \\ &\quad + \frac{1}{2}\langle \mathbf{H}_s, \bullet\bullet - \bullet \star \bullet \rangle \int_s^t (W_r - W_s) \, \mathrm{d}W_r + \frac{1}{2}\langle \mathbf{H}_s, \bullet \star \bullet \rangle \int_s^t (X_r - X_s)^2 \, \mathrm{d}W_r\end{aligned}$$
and Π_n is a sequence of partitions of $[0,1]$ whose size $|\Pi_n|$ converges to 0. Using the result in [49, Chapter 4] we know that
$$\int_0^t \langle 1^*, \mathbf{H}_r \rangle \mathrm{d}W_r = \overset{\mathbb{P}}{\lim_{n \to +\infty}} \sum_{[s,t] \in \Pi_n} \langle \mathbf{H}_s, 1^* \rangle (W_t - W_s).$$
Therefore the result follows by showing that
$$\overset{\mathbb{P}}{\lim_{n \to +\infty}} \sum_{[s,t] \in \Pi_n} \mathbf{Z}_{s,t}^{I,I} = \langle \mathbf{H}_s, 1^* \rangle (W_t - W_s)$$
We will prove this convergence into two steps: First, we will show that all components of $\mathbf{H}$ are uniformly bounded by the sequence of random variables
$$\begin{aligned} Z_n := \sum_{[s,t] \in \Pi_n} \langle \mathbf{H}_s, \bullet \rangle \int_s^t (X_r - X_s) \mathrm{d}W_r + \frac{1}{2}\langle \mathbf{H}_s, \bullet\bullet - \bullet \star \bullet \rangle \int_s^t (W_r - W_s) \mathrm{d}W_r \\ + \frac{1}{2}\langle \mathbf{H}_s, \bullet \star \bullet \rangle \int_s^t (X_r - X_s)^2 \mathrm{d}W_r m \end{aligned} \tag{4.17}$$
which converges to 0 in L^2. If $\Pi_n = \{0 = t_0 < t_1^n < \dots t_N^n = 1\}$ one has the estimate
$$\begin{aligned} &\|Z_n\|_{L^2}^2 \lesssim \\ &\left\| \sum_{k=0}^{N-1} \int_{t_k^n}^{t_{k+1}^n} (X_r - X_{t_k^n}) \mathrm{d}W_r \right\|_{L^2}^2 + \left\| \sum_{k=0}^{N-1} \int_{t_k^n}^{t_{k+1}^n} (X_r - X_{t_k^n})^2 \mathrm{d}W_r \right\|_{L^2}^2 + \left\| \sum_{k=0}^{N-1} \int_{t_k^n}^{t_{k+1}^n} (W_r - W_{t_k^n}) \mathrm{d}W_r \right\|_{L^2}^2 . \end{aligned}$$
Thanks to the Itô isometry and elementary computations on the variance of X and W the random variables inside each sum are centred and uncorrelated, therefore the above sum coincides with
$$\sum_{k=0}^{N-1} \int_{t_k^n}^{t_{k+1}^n} \mathbb{E}(X_r - X_{t_k^n})^2 + \mathbb{E}(X_r - X_{t_k^n})^4 + \mathbb{E}(W_r - W_{t_k^n})^2 \mathrm{d}r \lesssim \sum_{k=0}^{N-1} |t_{k+1}^n - t_k^n|^{3/2} \lesssim |\Pi_n|^{1/2},$$
thereby obtaining the desired convergence. To treat the general case we simply introduce the stopping time
$$\tau_M = \inf\{t \in [0,1] : |\langle \mathbf{H}_t, \bullet \rangle| \wedge |\langle \mathbf{H}_t, \bullet\bullet \rangle| \wedge |\langle \mathbf{H}_t, \bullet \star \bullet \rangle| \geq M\}$$
and consider the sequence Z_n^M defined as Z_n in (4.17) where the components of $\mathbf{H}_t$ are stopped with respect to the stopping time τ_M. Applying the previous part to the process Z_n^M we can use the fact that $\tau_M \to +\infty$ a.s. as $M \to +\infty$ to obtain that Z_n converges to 0 in probability. Passing to the other other three cases, namely (4.14), (4.15) and (4.16), the result follows easily from the first one by simply checking that one has the identities
$$\mathbf{Z}_{s,t}^{I,S} = \mathbf{Z}_{s,t}^{I,I} + \frac{1}{4}\langle \mathbf{H}_s, \bullet\bullet \rangle (t - s)$$
$$\mathbf{Z}_{s,t}^{S,I} = \mathbf{Z}_{s,t}^{I,I} - \frac{1}{4}\langle \mathbf{H}_s, \bullet \star \bullet \rangle (t - s)$$
$$\mathbf{Z}_{s,t}^{S,S} = \mathbf{Z}_{s,t}^{I,I} - \frac{1}{4}\langle \mathbf{H}_s, \substack{\bullet \\ | \\ \bullet} \rangle (t - s)$$

as direct consequence of the Itô formula applied to $(W_t - W_s)^2$. □

Remark 4.15. Combining the explicit representations of the rough integrals with the simple Itô formula (2.18) we can recover some previous results already present in the literature. For example, by simply plugging the Stratonovich-Stratonovich branched rough path into it we can simplify the deterministic integral due to the identity

$$\langle \mathrm{D}^2\phi(X_t), \overset{\bullet}{\underset{\bullet}{|}} \rangle = 0$$

thereby obtaining the following change of variable formula

$$\varphi(X_t) - \varphi(X_s) = \int_s^t \mathrm{D}\varphi(X_r)\,\mathrm{d}(\mathbf{X}^{S,S})_r^{\pi(\bullet)} + \int_s^t \varphi''(X_r)\,\mathrm{d}W_r. \tag{4.18}$$

In addition, by evaluating (4.15) with $\mathbf{H} = \varphi(X)$ we have

$$\int_0^t \varphi(X_r)\,\mathrm{d}\mathbf{X}_r^{\pi(\bullet\bullet)} = \int_0^t \varphi(X_r)\,\mathrm{d}W_r + \frac{1}{4}\int_0^t \varphi''(X_r)\,\mathrm{d}r. \tag{4.19}$$

Using different techniques in [42] the authors proved two identities in law very similar to those. In the first place, the authors considered the following weighted Riemann sums

$$S_n^W := \sum_{k=0}^{n-1} \varphi(X_{\frac{k}{n}})((X_{\frac{(k+1)}{n}} - X_{\frac{k}{n}})^2 - 1)$$

and showed the convergence

$$S_n^W \overset{(d)}{\to} C_{1/4}\int_0^t \varphi(X_r)\,\mathrm{d}W_r + \frac{1}{4}\int_0^t \varphi''(X_r)\,\mathrm{d}r. \tag{4.20}$$

Then, by considering the midpoint Riemann sum

$$S_n^M := \sum_{k=0}^{\lfloor n/2 \rfloor} \varphi'(X_{\frac{2k-1}{n}})(X_{\frac{2k}{n}} - X_{\frac{2k-2}{n}})$$

they show that this sum converges in law to

$$S_n^M \overset{(d)}{\to} \varphi(X_1) - \varphi(X_0) - \frac{\kappa}{2}\int_0^1 \varphi''(X_r)\mathrm{d}W_r \tag{4.21}$$

for some universal constant $\kappa \approx 1290$. Similar Ito formulas in law were deduced for a stochastic process with similar properties [7]. Therefore up to a redefinition of some underlying constants, we can combine the identities (4.18) and (4.19) to obtain the convergences in law

$$S_n^W \overset{(d)}{\to} \int_0^1 \mathrm{D}\varphi(X_r)\mathrm{d}(\mathbf{X}^{S,S})_r^{\pi(\bullet)}, \quad S_n^M \overset{(d)}{\to} \int_0^1 \mathrm{D}\varphi(X_r)\mathrm{d}(\mathbf{X}^{S,S})_r^{\pi(\bullet)} \tag{4.22}$$

From this equalities in law, it would be interesting to understand how Riemann sums like (4.21) and (4.20) could be used to define a proper convergence in law between rough paths. ♦

References

[1] M. Aguiar and A. Lauve, *The characteristic polynomial of the Adams operators on graded connected Hopf algebras*, Algebra Number Theory **9** (2015), no. 3, 547–583.

[2] F. Avram and M. S. Taqqu, *Symmetric polynomials of random variables attracted to an infinitely divisible law*, Probability theory and related fields **71** (1986), no. 4, 491–500.

[3] C. Bellingeri, *Quasi-geometric rough paths and rough change of variable formula*, Ann. Inst. Henri Poincaré Probab. Stat. **59** (2023), no. 3, 1398–1433.

[4] H. Boedihardjo and I. Chevyrev, *An isomorphism between branched and geometric rough paths*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 2, 1131–1148.

[5] D. J. Broadhurst and D. Kreimer, *Towards cohomology of renormalization: Bigrading the combinatorial hopf algebra of rooted trees*, Commun. Math. Phys. **215** (2000), 217–236.

[6] Y. Bruned, C. Curry, and K. Ebrahimi-Fard, *Quasi-shuffle algebras and renormalisation of rough differential equations*, Bull. Lond. Math. Soc. **52** (2020), no. 1, 43–63.

[7] K. Burdzy and J. Swanson, *A change of variable formula with Itô correction term*, Ann. Probab. **38** (2010), no. 5, 1817–1869.

[8] T. Cass and M. Weidner, *Tree algebras over topological vector spaces in rough path theory*, (2017).

[9] F. Chapoton, *Free pre-Lie algebras are free as Lie algebras*, Canad. Math. Bull. **53** (2010), no. 3, 425–437.

[10] F. Chapoton and M. Livernet, *Pre-Lie algebras and the rooted trees operad*, Internat. Math. Res. Notices **2001** (2001), no. 8, 395–408.

[11] A. Connes and D. Kreimer, *Hopf algebras, renormalization and noncommutative geometry*, Commun. Math. Phys **199** (1998), no. 1.

[12] A. M. Davie, *Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation*, Applied Mathematics Research eXpress **2008** (2008), abm009.

[13] K. Ebrahimi-Fard, S. J. A. Malham, F. Patras, and A. Wiese, *Flows and stochastic taylor series in itô calculus*, Journal of physics. A, Mathematical and theoretical **48** (2015), no. 49, 495202 (eng).

[14] E. Ferrucci, *A transfer principle for branched rough paths*, 2022.

[15] L. Foissy, *An introduction to hopf algebras of trees*, http://loic.foissy.free.fr/pageperso/p11.pdf, 2013.
[16] L. Foissy, F. Patras, and J.-Y. Thibon, *Deformations of shuffles and quasi-shuffles*, Ann. Inst. Fourier (Grenoble) **66** (2016), no. 1, 209–237.
[17] L. Foissy, *Finite-dimensional comodules over the Hopf algebra of rooted trees*, J. Algebra **255** (2002), no. 1, 89–120.
[18] ______, *Les algèbres de hopf des arbres enracinés décorés, I*, Bull. Sci. Math. **126** (2002), no. 3, 193–239.
[19] ______, *Algebraic structures associated to operads*, 2017, arXiv:1702.05344 [math.RA].
[20] ______, *Algebraic structures on typed decorated rooted trees*, Symmetry, Integrability and Geometry: Methods and Applications (SIGMA) **17** (2021), no. 086.
[21] P. K. Friz and M. Hairer, *A course on rough paths*, second ed., Universitext, Springer, Cham, [2020] ©2020, With an introduction to regularity structures.
[22] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications.
[23] R. Grossman and R. G. Larson, *Hopf-algebraic structure of families of trees*, J. Algebra **126** (1989), no. 1, 184–210.
[24] M. Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140.
[25] ______, *Ramification of rough paths*, J. Differential Equations **248** (2010), no. 4, 693–721.
[26] M. Hairer and D. Kelly, *Geometric versus non-geometric rough paths*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 1, 207–251.
[27] M. E. Hoffman, *Quasi-shuffle products*, J. Algebraic Combin. **11** (2000), no. 1, 49–68.
[28] M. E. Hoffman, *Combinatorics of rooted trees and Hopf algebras*, Transactions of the American Mathematical Society **355** (2003), no. 9, 3795–3811 (eng).
[29] M. E. Hoffman and K. Ihara, *Quasi-shuffle products revisited*, J. Algebra **481** (2017), 293–326.
[30] K. Itô, *On a formula concerning stochastic differentials*, Nagoya Mathematical Journal **3** (1951), no. none, 55 – 65.
[31] D. Kelly, *Itô corrections in stochastic calculus*, Ph.D. thesis, University of Warwick, 2012.
[32] J.-L. Loday and M. Ronco, *On the structure of cofree Hopf algebras*, J. Reine Angew. Math. **592** (2006), 123–155.
[33] J.-L. Loday and B. Vallette, *Algebraic operads*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 346, Springer, Heidelberg, 2012.
[34] J. Loday and B. Vallette, *Algebraic operads*, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 2012.
[35] T. Lyons, *Differential equations driven by rough signals (I): an extension of an inequality of l. c. young*, Mathematical research letters **1** (1994), no. 4, 451–464 (eng).
[36] T. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths ecole d'eté de probabilités de saint-flour xxxiv-2004*, 2007.
[37] T. Lyons and N. Victoir, *An extension theorem to rough paths*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **24** (2007), no. 5, 835–847.
[38] T. J. Lyons, *Differential equations driven by rough signals.*, Rev. Mat. Iberoam **14** (1998), no. 2, 215–310.
[39] D. Manchon, *Hopf algebras in renormalisation*, Handbook of Algebra (M. Hazewinkel, ed.), vol. 5, North-Holland, 2008, pp. 365–427.
[40] J. W. Milnor and J. C. Moore, *On the structure of Hopf algebras*, Ann. of Math. (2) **81** (1965), 211–264.
[41] I. Nourdin, *Selected aspects of fractional Brownian motion*, Bocconi & Springer Series, vol. 4, Springer, Milan; Bocconi University Press, Milan, 2012.
[42] I. Nourdin and A. Réveillac, *Asymptotic behavior of weighted quadratic variations of fractional Brownian motion: the critical case* $H = 1/4$, Ann. Probab. **37** (2009), no. 6, 2200–2230.
[43] J.-C. Novelli, F. Patras, and J.-Y. Thibon, *Natural endomorphisms of quasi-shuffle Hopf algebras*, Bull. Soc. Math. France **141** (2013), no. 1, 107–130.
[44] D. Nualart and S. Tindel, *A construction of the rough path above fractional Brownian motion using Volterra's representation*, Ann. Probab. **39** (2011), no. 3, 1061–1096.
[45] J.-M. Oudom and D. Guin, *On the lie enveloping algebra of a pre-lie algebra*, J. K-Theory **2** (2008), no. 1, 147–167.
[46] F. Panaite, *Relating the Connes–Kreimer and Grossman–Larson Hopf algebras built on rooted trees*, Letters in Mathematical Physics **51** (2000).
[47] F. Patras, *La décomposition en poids des algèbres de hopf*, Ann. Inst. Fourier **43** (1993), no. 4, 1067–1087.
[48] C. Reutenauer, *Free Lie algebras*, London Mathematical Society Monographs. New Series, vol. 7, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications.
[49] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999.
[50] L. C. G. Rogers, *Arbitrage with fractional Brownian motion*, Math. Finance **7** (1997), no. 1, 95–105.
[51] A. Segall and T. Kailath, *Orthogonal functionals of independent-increment processes*, IEEE Trans. Info. Theory **22** (1976), 287–298.
[52] J. L. Solé and F. Utzet, *On the orthogonal polynomials associated with a lévy process*, Ann. Probab. **36** (2008), 765–795.
[53] M. E. Sweedler, *Hopf algebras*, Mathematics lecture note series, W. A. Benjamin, 1969.
[54] M. TAKEUCHI, *Free Hopf algebras generated by coalgebras*, Journal of the Mathematical Society of Japan **23** (1971), no. 4, 561 – 582.
[55] M. Takeuchi, *Free Hopf algebras generated by coalgebras*, J. Math. Soc. Japan **23** (1971), 561–582.
[56] R. G. Underwood, *Fundamentals of Hopf algebras*, Universitext, Springer, Cham, 2015.
[57] J. Unterberger, *Hölder-continuous rough paths by Fourier normal ordering*, Comm. Math. Phys. **298** (2010), no. 1, 1–36.
[58] M. G. Varzaneh, S. Riedel, A. Schmeding, and N. Tapia, *The geometry of controlled rough paths*, 2022.
[59] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Mathematica **67** (1936), no. none, 251 – 282.