UNIVERSITY *of* York

This is a repository copy of *A combinatorial approach to geometric rough paths and their controlled paths*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/181761/

Version: Published Version

Article:

Cass, Thomas, Driver, Bruce K., Litterer, Christian et al. (1 more author) (2022) A combinatorial approach to geometric rough paths and their controlled paths. Journal of the London Mathematical Society. pp. 936-981. ISSN: 1469-7750

https://doi.org/10.1112/jlms.12589

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

RESEARCH ARTICLE

A combinatorial approach to geometric rough paths and their controlled paths

Thomas Cass[1] | **Bruce K. Driver[2]** | **Christian Litterer[3]** |
Emilio Rossi Ferrucci[1]

[1]Department of Mathematics, Imperial
College London, London, England

[2]Department of Mathematics, University
of California San Diego, La Jolla,
California, USA

[3]Department of Mathematics, University
of York, York, UK

Correspondence
Emilio Rossi Ferrucci, Department of
Mathematics, Imperial College London,
London, UK.
Email:
emilio.rossi-ferrucci16@imperial.ac.uk

Funding information
EPSRC, Grant/Award Numbers:
EP/S026347/1, EP/V005413/1; Centre for
Doctoral Training in Financial
Computing & Analytics; EPSRC,
Grant/Award Number: EP/L015129/1

Abstract

We develop the structure theory for transformations
of weakly geometric rough paths of bounded $1 < p$-
variation and their controlled paths. Our approach differs from existing approaches as it does not rely on
smooth approximations. We derive an explicit combinatorial expression for the rough path lift of a controlled path, and use it to obtain fundamental identities such as the associativity of the rough integral, the
adjunction between pushforwards and pullbacks and a
change of variables formula for rough differential equations (RDEs). As applications we define rough paths,
rough integration and RDEs on manifolds, extending the
results of Cass, Driver, and Litterer (Proc. Lond. Math.
Soc. (3) **111** (2015), no. 6, 1471–1518) to the case of arbitrary p.

MSC 2020
60L20 (primary)

INTRODUCTION

The theory of rough paths and the fundamental ideas that underlie it are now well established,
with a vast number of applications in areas as diverse as stochastic and numerical analysis,
machine learning and stochastic partial differential equations. Martin Hairer's celebrated work on

regularity structures, for example, can be regarded as a far-reaching generalisation that is inspired by some of the fundamental ideas underlying rough path theory (and indeed rough paths can be identified with a special case of the theory). The basic theory is now somewhat classical and several approaches have been developed to obtain (and in some cases significantly extend) its core results [9, 11–13, 15, 17, 18]. In these approaches, the core results are usually either developed explicitly for the case $2 \leqslant p < 3$ or follow from properties inherited by approximation arguments from the corresponding properties of the lifts of smooth paths.

In this paper we use algebraic and combinatorial methods to explore the basic structure underlying transformations of rough paths of arbitrary roughness under sufficiently regular maps. Our approach allows us to work directly with weakly geometric rough paths and leads to a clean separation of analysis and algebra, yielding explicit combinatorial descriptions of the resulting objects. The structure theory for geometric rough paths is usually deduced from the corresponding properties of the (lifts of) smooth paths by taking closures in a suitable rough path metric. In finite dimensions, this means that identities for geometric rough paths readily extend to the weakly geometric setting. However, this extension is predicated on the close relation of weakly geometric and geometric rough paths established by Friz and Victoir [11] for paths with values in finite-dimensional spaces. It is presently not clear if a similar relation holds in infinite dimensions. Similarly, the smooth approximation arguments are not available when studying more general branched rough paths. Although the combinatorics is, at times, quite complex, we have made it a priority to state all results in a clear, coordinate-free manner, without resorting to the coordinate notation that is used in the proofs.

Another goal of this paper is to unify Lyons's original approach and Gubinelli's 'linearised' version [12] which deals with controlled rough paths, or *controlled paths* as we call them here, to avoid ambiguity. This is widely considered to be the most general and modern approach to rough path theory. Many fundamental results, such as the definition and convergence of controlled-rough integrals, are not present in the literature, stated in this setting. The fact that we work with controlled paths provides further motivation to avoid smooth approximation: controlled paths are only indirectly defined in terms of their reference rough path, and a smooth approximation of the rough path does not automatically yield one of the controlled path.

In [4] the authors derive, without using smooth approximation arguments, the basic structure theory for weakly geometric p-rough paths for the case $2 \leqslant p < 3$, leading to an explicit characterisation of rough paths constrained to an embedded submanifold of $\mathbb{R}^d$. In this paper we generalise these results to weakly geometric rough paths of arbitrary roughness. Even though our results are proved in the finite-dimensional setting thanks to the use of bases, the statements themselves are stated in a coordinate free manner and preserve their meaning when dropping the finite-dimensional assumption. We therefore expect that with some care they may be generalised to the infinite-dimensional setting, and hope that they will inform future work on branched rough paths. Generalising the arguments from the $2 \leqslant p < 3$ case is a non-trivial challenge, as identities and proofs become significantly more combinatorial in nature. As applications we develop the theory of weakly geometric rough paths on manifolds, their controlled paths, and the associated integration theory and rough differential equations (RDEs). For arbitrary p-rough paths the combinatorial identities corresponding to key structural properties of rough paths such as the functoriality of pushforwards are complex and difficult to obtain directly. Fortunately, analogous results for controlled paths are more readily obtained exploiting their more linear structure. Consequently, we first establish identities for controlled paths and then deduce the results for full rough paths. It is a well-known fact going back to [12] that a controlled path can in principle be lifted to a full rough path. An explicit construction for the case $2 \leqslant p < 3$ can, for example, be

found in [9]. For the general lift of a controlled rough path in the branched case Gubinelli has obtained an inductive formula [13, Remark 8.7]. Unfortunately, the case $2 \leqslant p < 3$ is not very representative for the construction of the lift and the description of the lifted rough path in [13] is not very explicit. In this paper we give an closed-form description of the lift for general p. The result may be regarded as a generalisation of the construction of the rough integral carried out in [15]. However, our result does not require any of the symmetries of the one-form integrands, and instead uses a combinatorial property of the ordered shuffle. In addition, generalising an argument from [5], we are able to show that a controlled rough path defined with respect to a weakly geometric rough path lifts again to a weakly geometric path.

The paper is structured as follows: in Section 1 we introduce algebraic preliminaries and notations, in particular, the shuffle and ordered shuffle products, and associated lemmata that will be integral to our constructions. In Section 2 we recall the basic notion of weakly geometric and controlled paths and proceed to construct the lift of a controlled path for general p. Underlying the proof is a generalisation of an argument in [5] that gave a similar result for Lip-γ 1-forms, but avoids the use of symmetries not present in general Gubinelli derivatives. The construction of the lift for general p has considerable combinatorial complexity but enables us to define pushforwards of rough paths using controlled paths. In order to study the properties of those pushforwards we define a change of reference rough path for a controlled rough path and study its structure. The machinery developed for controlled rough paths allows us in Proposition 2.20 to establish the fundamental properties of pushforwards of $p-$rough paths. In particular, we prove the functorial property of the pushfowards. In Theorem 2.21 we demonstrate the associativity of the rough integral, that is, the rigorous formulation of the heuristic statement that 'we can substitute differentials'. Theorem 2.22 establishes a pushforward-pullback adjunction of controlled paths and their reference rough paths under the rough integral pairing. Finally, combining these results in Theorem 2.27 we obtain a change of variable result for the solutions of RDEs. All proofs separate algebra and analysis, in that they do not require approximation by lifts of smooth paths. In Section 3 we study applications to rough paths on manifolds. The structure of pushforwards of rough path gives rise to a natural notion of rough paths on manifolds defined locally using the chart functions (cf. [3]). The change of variable result for RDEs in Theorem 2.27 motivates a definition of RDEs on manifolds. We clarify some of the relations of the intrinsic rough paths considered in this paper with existing notions of rough paths that have been obtained in different geometric settings. The interplay between controlled paths and the change of their reference path allows us to describe a natural class of controlled paths with respect to a rough path on a manifold. We define the rough integral for a suitable class of integrands controlled by a rough path on a manifold, and give meaning to an RDE driven by a manifold-valued rough path, with solution valued in a second manifold. In Remark 3.9 we show how our definitions directly extend those of [4] in the extrinsic framework.

1 | THE SHUFFLE AND TENSOR BIALGEBRAS

We begin with a concise review of bialgebras defined on tensor algebras, for which we refer to [19, Chapter I] [22, Chapter 2]. Given $n_1, \dots, n_m \in \mathbb{N}$ (which may be 0) we define $\mathrm{Sh}(n_1, \dots, n_m)$ to be the subset of the permutation group $\mathfrak{S}_{n_1 + \dots + n_m}$ of $(n_1, \dots, n_m)$-*shuffles*, that is, permutations σ with the property that

$$\sigma(n_1 + \dots + n_{i-1} + 1) < \sigma(n_1 + \dots + n_{i-1} + 2) < \dots < \sigma(n_1 + \dots + n_i) \tag{1}$$

for $i = 1, \dots, m$ (with $n_0 := 0$). We will additionally call σ an $(n_1, \dots, n_m)$-*ordered shuffle* if

$$\sigma(n_1) \leqslant \sigma(n_1 + n_2) \leqslant \dots \leqslant \sigma(n_1 + \dots + n_m) \tag{2}$$

and we denote the set of these with $\overline{\mathrm{Sh}}(n_1, \dots, n_m)$. If $n_i = 0$ for some i we have $\mathrm{Sh}(n_1, \dots, n_m) = \mathrm{Sh}(n_1, \dots, \hat{n}_i, \dots, n_m)$, $\overline{\mathrm{Sh}}(n_1, \dots, n_m) = \overline{\mathrm{Sh}}(n_1, \dots, \hat{n}_i, \dots, n_m)$ (with $\hat{}$ denoting omission).

In this paper the letters $U, V, W, \dots$ will always be finite-dimensional $\mathbb{R}$-vector spaces. Given such a vector space V, a permutation $\sigma \in \mathfrak{S}_n$ induces a linear isomorphism

$$\sigma_* : V^{\otimes n} \to V^{\otimes n}, \quad v_1 \otimes \dots \otimes v_n \mapsto v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(n)}. \tag{3}$$

If $\{e_i\}_{i \in I}$ is a basis of V, we may write $a \in V^{\otimes n}$ uniquely as $a^{(i_1, \dots, i_n)} e_{i_1} \otimes \dots \otimes e_{i_n} =: a^{\boldsymbol{i}} e_{\boldsymbol{i}}$ where there is a sum on the ordered n-tuple of basis indices $\boldsymbol{i} := (i_1, \dots, i_n)$ (following the Einstein convention), and

$$\sigma_*(a) = a^{\boldsymbol{i}} e_{\sigma_* \boldsymbol{i}} = a^{\sigma_*^{-1} \boldsymbol{j}} e_{\boldsymbol{j}},$$

where

$$\rho_* \boldsymbol{i} := (i_{\rho(1)}, \dots, i_{\rho(n)}) \text{ for } \rho \in \mathfrak{S}_n. \tag{4}$$

We let $I^\bullet := \bigcup_{n \in \mathbb{N}} I^n$ be the set of I-valued tuples; this includes the empty tuple (), and we use $I^\bullet_*$ to denote the set of all such non-empty tuples. We will sometimes identify a tuple $(k_1, \dots, k_n)$ with the corresponding tensor $e_{k_1} \otimes \dots \otimes e_{k_n}$ according to the chosen basis. Note that $\sigma_*^{-1} \boldsymbol{i}$ is the tuple obtained by 'permuting $\boldsymbol{i}$ according to σ', for example, if

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 2 & 4 \end{pmatrix} \in \mathrm{Sh}(3, 2), \quad \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix} \in \mathfrak{S}_5,$$

then

$$\sigma_*^{-1}(i_1, i_2, i_3, i_4, i_5) = (i_1, i_4, i_2, i_5, i_3).$$

Note that the composition rule (both for tensors and tuples) is

$$(\sigma \circ \rho)_* = \rho_* \sigma_*. \tag{5}$$

Indeed, denoting $w_k := v_{\sigma(k)}$ we have

$$(\sigma \circ \rho)_*(v_1 \otimes \dots \otimes v_n) = v_{\sigma(\rho(1))} \otimes \dots \otimes v_{\sigma(\rho(n))}$$

$$= w_{\rho(1)} \otimes \dots \otimes w_{\rho(n)}$$

$$= \rho_*(w_1 \otimes \dots \otimes w_n)$$

$$= \rho_* \sigma_*(v_1 \otimes \dots \otimes v_n).$$

For a tuple $\boldsymbol{i} = (i_1, \ldots, i_n)$ we will denote $|\boldsymbol{i}| := n$ its length, and given two tuples $\boldsymbol{i}, \boldsymbol{j}$ we write $\boldsymbol{ij}$ for their concatenation. We will denote $T(V) := \bigoplus_{n=0}^{\infty} V^{\otimes n}$ and for $a \in T(V)$, with a^n its projection onto $V^{\otimes n}$; this has a distinct meaning to the notation $a^{\boldsymbol{i}}$ for tuples $\boldsymbol{i} \in I^{\bullet}$, explained above. When we are considering the tensor products of the dual V^* of a vector space V, or more generally the space of linear maps $\mathcal{L}(V, W)$ from V to another vector space W, we will replace superscripts with subscripts and vice versa.

We denote by $(T(V), \otimes, \Delta_{\sqcup\!\sqcup})$ the tensor bialgebra of V, that is, the product is given by the ordinary tensor product, which in coordinates reads

$$(a \otimes b)^{\boldsymbol{k}} = \sum_{\boldsymbol{ij}=\boldsymbol{k}} a^{\boldsymbol{i}} b^{\boldsymbol{j}}, \tag{6}$$

and the *shuffle coproduct* is defined on elementary tensors (and extended linearly) as

$$\Delta_{\sqcup\!\sqcup} : T(V) \to T(V) \boxtimes T(V)$$

$$v_1 \otimes \cdots \otimes v_n \mapsto \sum_{\substack{k=0,\ldots,n \\ \sigma \in \mathrm{Sh}(k, n-k)}} (v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}) \boxtimes (v_{\sigma(k+1)} \otimes \cdots \otimes v_{\sigma(n)}). \tag{7}$$

Here use the symbol $\boxtimes$ to denote external tensor product, reserving $\otimes$ for the algebra product. In coordinates this reads, for $a \in T(V)$

$$(\Delta_{\sqcup\!\sqcup} a)^{\boldsymbol{i},\boldsymbol{j}} = \sum_{\substack{\sigma \in \mathrm{Sh}(|\boldsymbol{i}|,|\boldsymbol{j}|) \\ \boldsymbol{k} = \sigma_*^{-1}(\boldsymbol{ij})}} a^{\boldsymbol{k}} =: \sum_{\boldsymbol{k} \in \mathrm{Sh}(\boldsymbol{i},\boldsymbol{j})} a^{\boldsymbol{k}}, \tag{8}$$

where $\boldsymbol{i}$ denotes the index of the first $\boxtimes$-factor and $\boldsymbol{j}$ that of the second. In order to give a precise meaning to $\mathrm{Sh}(\boldsymbol{i}, \boldsymbol{j})$ we must introduce multiset notation. Recall that a multiset is like a set (in that the order of its elements is not taken into account), but with the difference that the same element may appear more than once; we will denote multisets with double braces, for example, $\{\!\{1, 2, 2, 2, 3, 3\}\!\}$. If A and B are multisets, we write $A \subseteq B$ if each element of A belongs to B counted with its multiplicity, for example, $\{\!\{2, 2, 3, 3\}\!\} \subseteq \{\!\{1, 2, 2, 2, 3, 3\}\!\}$ but $\{\!\{1, 1, 2, 2, 3, 3\}\!\} \nsubseteq \{\!\{1, 2, 2, 2, 3, 3\}\!\}$, and $A = B$ is defined to mean $A \subseteq B$ and $B \subseteq A$. With this in mind, we are defining

$$\mathrm{Sh}(\boldsymbol{i}, \boldsymbol{j}) := \{\!\{\sigma_*^{-1}(\boldsymbol{ij}) \mid \sigma \in \mathrm{Sh}(|\boldsymbol{i}|, |\boldsymbol{j}|)\}\!\}. \tag{9}$$

This means that the tuple $\boldsymbol{k}$ appears as many times as there are shuffles σ with the property that $\boldsymbol{k} = \sigma_*^{-1}(\boldsymbol{ij})$. Similar multisets will be defined without explicit mention from now on.

The bialgebra that is graded dual to $(T(V), \otimes, \Delta_{\sqcup\!\sqcup})$ is given by $(T(V^*), \sqcup\!\sqcup, \Delta_{\otimes})$. Note that we are using the notion of graded duality for bialgebras (see [10, §1.5]), which is different to ordinary duality: in a nutshell, this just means that we are taking the dual of each (finite-dimensional) direct summand and that the product (coproduct) in one bialgebra is the dual to the product (coproduct) in the other; 'dual' here makes sense because products and coproducts respect the grading. This allows us to avoid considering formal series of tensors, which are unnecessary when considering rough paths without their full signatures, and retain most of the usual properties of

duality. $\sqcup\!\sqcup = \Delta^*_{\sqcup\!\sqcup}$ is the *shuffle product* given by

$$\sqcup\!\sqcup \colon T(V^*) \boxtimes T(V^*) \to T(V^*),$$

$$(v^1 \otimes \cdots \otimes v^n) \boxtimes (v^{n+1} \otimes \cdots \otimes v^{n+m}) \mapsto \sum_{\sigma \in \mathrm{Sh}(n,m)} v^{\sigma^{-1}(1)} \otimes \cdots \otimes v^{\sigma^{-1}(n+m)}, \tag{10}$$

which in coordinates (using subscripts, since we are working in V^*) reads

$$(a \sqcup\!\sqcup b)_{\boldsymbol{k}} = \sum_{\substack{|\boldsymbol{i}|=0,\dots,|\boldsymbol{k}| \\ \sigma \in \mathrm{Sh}(|\boldsymbol{i}|,|\boldsymbol{k}|-|\boldsymbol{i}|) \\ \boldsymbol{ij}=\sigma_*\boldsymbol{k}}} a_{\boldsymbol{i}} b_{\boldsymbol{j}} =: \sum_{(\boldsymbol{i},\boldsymbol{j}) \in \mathrm{Sh}^{-1}(\boldsymbol{k})} a_{\boldsymbol{i}} b_{\boldsymbol{j}}. \tag{11}$$

Note that with this notation $(\boldsymbol{i}, \boldsymbol{j}) \in \mathrm{Sh}^{-1}(\boldsymbol{k}) \Leftrightarrow \boldsymbol{k} \in \mathrm{Sh}(\boldsymbol{ij})$; in particular $\boldsymbol{i}$ or $\boldsymbol{j}$ may be empty. The coproduct $\Delta_\otimes = \otimes^*$ is the *deconcatenation coproduct* given by

$$\Delta_\otimes \colon T(V^*) \to T(V^*) \boxtimes T(V^*)$$

$$v^1 \otimes \cdots \otimes v^n \mapsto \sum_{k=0}^{n} (v^1 \otimes \cdots \otimes v^k) \boxtimes (v^{k+1} \otimes \cdots \otimes v^n) \tag{12}$$

or in coordinates

$$(\Delta_\otimes a)_{i,j} = a_{ij}. \tag{13}$$

Recall that in every bialgebra with coproduct Δ we may define its reduced coproduct $\widetilde{\Delta} a := \Delta a - a \boxtimes 1 - 1 \boxtimes a$, which is also coassociative. Also recall that in a coalgebra (C, Δ) (for us C will always be a tensor algebra) the (reduced) coproduct can be iterated, as coassociativity guarantees that $\Delta^m \colon C \to C^{\boxtimes m}$ has a unique meaning (note that under this convention $\Delta^2 = \Delta$, $\Delta^1 := \mathbb{1}_C$, $\Delta^0 := 1_\mathbb{R}$). Since the above bialgebras are graded and connected, the reduced (iterated) coproduct factors as

$$\widetilde{\Delta}^m = \pi_{\geqslant 1}^{\boxtimes m} \circ \Delta^m \tag{14}$$

where $\pi_{\geqslant 1} \colon T(V) \twoheadrightarrow \bigoplus_{n \geqslant 1} V^{\otimes n}$ is the projection onto tensor products of positive order.

We may define the *ordered shuffle coproduct* $\Delta_{\overline{\sqcup\!\sqcup}}$ and the *ordered shuffle product* $\overline{\sqcup\!\sqcup}$ by requiring shuffles in (7) and (10) to be ordered. This does not, in fact, define a real (co)product, because $\overline{\sqcup\!\sqcup}$ fails to be associative: indeed, it satisfies the alternative relation

$$a \overline{\sqcup\!\sqcup} (b \overline{\sqcup\!\sqcup} c) = (a \sqcup\!\sqcup b + b \sqcup\!\sqcup a) \overline{\sqcup\!\sqcup} c.$$

This property makes $(T(V), \overline{\sqcup\!\sqcup})$ a Zinbiel algebra [8]. Whenever we iterate $\overline{\sqcup\!\sqcup}$ or $\Delta_{\overline{\sqcup\!\sqcup}}$, we will be carrying out composition from left to right, that is, inductively

$$a_1 \overline{\sqcup} \cdots \overline{\sqcup} a_n := (a_1 \overline{\sqcup} a_2) \overline{\sqcup} a_3 \overline{\sqcup} \cdots \overline{\sqcup} a_n$$

$$\Delta_{\sqcup}^m a := \sum_{(a)_{\sqcup}^m} a_{(1)} \boxtimes \cdots \boxtimes a_{(m)}$$

$$:= \sum_{(a)_{\sqcup}^{m-1}} \left(\Delta_{\overline{\sqcup}} a_{(1)} \right) \boxtimes a_{(2)} \boxtimes \cdots \boxtimes a_{(m-1)}. \tag{15}$$

This guarantees that the coordinate expression provided for the unordered shuffle carries over to the ordered case, with $\overline{\mathrm{Sh}}$ instead of Sh, for example,

$$(\widetilde{\Delta}_{\sqcup}^m a)^{k^1,\dots,k^m} = [|k^1|,\dots,|k^m| \geqslant 1] \sum_{\substack{\sigma \in \overline{\mathrm{Sh}}(|k^1|,\dots,|k^m|) \\ k=\sigma_*^{-1}(k^1 \dots k^m)}} a^k$$

$$=: [|k^1|,\dots,|k^m| \geqslant 1] \sum_{k \in \overline{\mathrm{Sh}}(k^1,\dots,k^m)} a^k. \tag{16}$$

Here the square bracket has binary value depending on the truth value of the proposition it contains, and is present because the coproduct is reduced; the set of permutations over which the sum is taken is given by ordered shuffles, reflecting the fact that we are dealing with the ordered shuffle coproduct.

Before proceeding, we take a moment to motivate our interest in shuffles and ordered shuffles, although this will become much clearer in Section 2. While it is well known that the former are used to express products of iterated integrals of a (smooth) path X

$$\left(\int_{s<u_1<\dots<u_n<t} \mathrm{d}X_{u_1}^{i_1} \cdots \mathrm{d}X_{u_m}^{i_m} \right) \left(\int_{s<v_1<\dots<v_n<t} \mathrm{d}X_{v_1}^{j_1} \cdots \mathrm{d}X_{v_n}^{j_n} \right)$$

$$= \sum_{k \in \mathrm{Sh}(i,j)} \int_{s<r_1<\dots<r_{n+m}<t} \mathrm{d}X_{r_1}^{k_1} \cdots \mathrm{d}X_{r_{m+n}}^{k_{m+n}},$$

the role of ordered shuffles in the study of iterated integrals of paths is less appreciated. One way to motivate their significance is as follows: let Y be the solution to the ODE

$$\mathrm{d}Y = V(Y)\mathrm{d}X = V(Y)\dot{X}\mathrm{d}t.$$

with X V-valued and Y W-valued. We fix bases on both vector spaces and use Greek indices for V and Latin ones for W — this will be the convention later on as well. Substituting formal Euler expansions, and defining $V_\gamma^k(y) := V_{\gamma_1} \cdots V_{\gamma_n}^k(y)$ for a tuple $\gamma = (\gamma_1,\dots,\gamma_n)$, with the product denoting iterated composition of vector fields (that is, $V_\gamma f(y) := \partial_k f(y) V_\gamma^k(y)$ for a function f of y)

$$\int_{s<u_1<\dots<u_m<t} \mathrm{d}Y_{u_1}^{k_1} \cdots \mathrm{d}Y_{u_m}^{k_m}$$

$$= \int_{s<u_1<\dots<u_m<t} \mathrm{d}\left(V_{\gamma^1}^{k_1}(Y_s) X_{su_1}^{\gamma^1} \right) \cdots \mathrm{d}\left(V_{\gamma^m}^{k_m}(Y_s) X_{su_m}^{\gamma^m} \right)$$

$$= V^{k_1}_{\gamma^1}(Y_s) \cdots V^{k_m}_{\gamma^m}(Y_s) \int_{s<u_1<\ldots<u_m<t} \mathrm{d}\left(\int_{s<r^1_1<\ldots<r^1_{n_1}<u_n} \mathrm{d}X^{\gamma^1_1}_{r^1_1} \cdots \mathrm{d}X^{\gamma^1_{n_1}}_{r^1_{n_1}}\right) \cdots$$

$$\cdots \mathrm{d}\left(\int_{s<r^m_1<\ldots<r^m_{n_m}<u_m} \mathrm{d}X^{\gamma^m_1}_{r^m_1} \cdots \mathrm{d}X^{\gamma^m_{n_m}}_{r^m_{n_m}}\right)$$

$$= V^{k_1}_{\gamma^1}(Y_s) \cdots V^{k_m}_{\gamma^m}(Y_s) \tag{17}$$

$$\cdot \int_{\substack{s<r^1_1<\ldots<r^1_{n_1-1}<r^1_{n_1} \\ \vdots \\ s<r^m_1<\ldots<r^m_{n_m-1}<r^m_{n_m} \\ s<r^1_{n_1}<\ldots<r^m_{n_m}<t}} \mathrm{d}X^{\gamma^1_1}_{r^1_1} \cdots \mathrm{d}X^{\gamma^1_{n_1}}_{r^1_{n_1}} \cdots \cdots \mathrm{d}X^{\gamma^m_1}_{r^m_1} \cdots \mathrm{d}X^{\gamma^m_{n_m}}_{r^m_{n_m}}$$

$$= \sum_{\gamma^1,\ldots,\gamma^m} V^k_{\gamma^1}(Y_s) \cdots V^k_{\gamma^m}(Y_s) \sum_{\gamma\in\overline{\mathrm{Sh}}(\gamma^1,\ldots,\gamma^m)} \int_{s<v_1<\ldots<v_{n:=\sum_i n_i}} \mathrm{d}X^{\gamma_1}_{v_1} \cdots \mathrm{d}X^{\gamma_n}_{v_n}.$$

In other words, ordered shuffles index the sum involved in the expression of the iterated integrals of Y in terms of those of X. The relevance of ordered shuffles in the similar cases of linear RDEs and Lip$(\gamma-1)$ functions was observed in [15, pp. 72–75].

We will now prove a few results that will be used in Section 2. These will be stated in terms of tuples (although they are essentially statements about shuffles), so they can be readily deployed when dealing with rough paths. Unshuffling and concatenating satisfy the following commutativity relation:

$$
\begin{array}{ccc}
T(V) & \xrightarrow{\;\Delta^m_{\sqcup\!\sqcup}\;} & T(V)^{\boxtimes m} \\
\big\uparrow{\scriptstyle\otimes} & & \big\uparrow{\scriptstyle\otimes_{1,m+1}\boxtimes\cdots\boxtimes\otimes_{m,2m}} \\
T(V)^{\boxtimes 2} & \xrightarrow{\;\Delta^m_{\sqcup\!\sqcup}\boxtimes\Delta^m_{\sqcup\!\sqcup}\;} & T(V)^{\boxtimes 2m}
\end{array}
\quad .
$$

The following lemma can be viewed as the counterpart to this statement in the context of ordered shuffles. Since it is essentially combinatorial in nature, we will state it in terms of tuples. Recall that $I^\bullet_*$ denotes the set of all I-valued tuples of any positive order.

Lemma 1.1. *Let $\boldsymbol{k}^1,\ldots,\boldsymbol{k}^m \in I^\bullet_*$. The following identity of multisets holds:*

$$\{\!\{(\boldsymbol{i},\boldsymbol{j}) \in I^\bullet \times I^\bullet \mid \boldsymbol{i}\boldsymbol{j} \in \overline{\mathrm{Sh}}(\boldsymbol{k}^1,\ldots,\boldsymbol{k}^m)\}\!\}$$

$$= \{\!\{(\boldsymbol{i},\boldsymbol{j}) \in I^\bullet \times I^\bullet \mid \boldsymbol{j} \in \overline{\mathrm{Sh}}(\boldsymbol{j}^{l+1},\ldots,\boldsymbol{j}^m);$$

$$\boldsymbol{i} \in \mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{i}^1,\ldots,\boldsymbol{i}^l),\mathrm{Sh}(\boldsymbol{i}^{l+1},\ldots,\boldsymbol{i}^m)); \tag{18}$$

$$\text{where } \boldsymbol{i}^h = \boldsymbol{k}^h \text{ for } h \leqslant l;$$

$$\boldsymbol{i}^h\boldsymbol{j}^h = \boldsymbol{k}^h \text{ and } |\boldsymbol{j}^h| \geqslant 1, \text{ for } h \geqslant l+1;$$

$$\text{with } l = 0,\ldots,m\}\!\}.$$

The following picture is meant to illustrate the idea of the statement: the horizontal lines represent tuples, and the red bullet points represent their terminal elements.
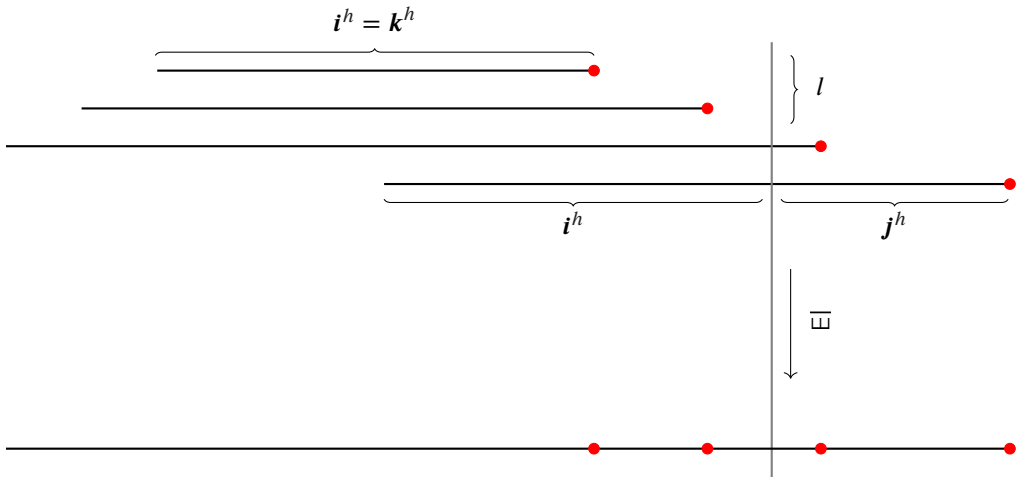
Note that we are taking into account multiplicities in the multiset $\boldsymbol{k} \in \overline{\mathrm{Sh}}(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$, that is, if the same tuple $\boldsymbol{k}$ belongs twice to $\overline{\mathrm{Sh}}(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$, then any pair $(\boldsymbol{i}, \boldsymbol{j})$ such that $\boldsymbol{i}\boldsymbol{j} = \boldsymbol{k}$ appears twice in the multiset; an analogous remark holds for the right-hand side and for all similarly defined multisets. $\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^l), \mathrm{Sh}(\boldsymbol{i}^{l+1}, \dots, \boldsymbol{i}^m))$ here stands for

$$\{\{\boldsymbol{i} \in \mathrm{Sh}(\boldsymbol{a}, \boldsymbol{b}) \mid \boldsymbol{a} \in \overline{\mathrm{Sh}}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^l), \ \boldsymbol{b} \in \mathrm{Sh}(\boldsymbol{i}^{l+1}, \dots, \boldsymbol{i}^m)\}\},$$

the multiset of tuples obtained by shuffling $\boldsymbol{i}^1, \dots, \boldsymbol{i}^m$, the first l with order. When $l = 0$ or m, this reduces respectively to $\mathrm{Sh}(\mathrm{Sh}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^m)) = \mathrm{Sh}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^m)$, $\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^m)) = \overline{\mathrm{Sh}}(\boldsymbol{i}^1, \dots, \boldsymbol{i}^m)$, since the only possible way of shuffling a single is to leave it unchanged. The proof of this lemma is most easily understood when going through its steps with reference to the example that immediately follows it.

Proof of Lemma 1.1. Let $A(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$ denote the first multiset defined above and $B(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$ the second. For tuples $\boldsymbol{\ell}^1, \dots, \boldsymbol{\ell}^n$ and $1 \leqslant a \leqslant b \leqslant n$ define $\boldsymbol{\ell}^{a:b} := \boldsymbol{\ell}^a \dots \boldsymbol{\ell}^b \in I^\bullet$ (juxtaposition) and for a tuple $\boldsymbol{\ell}$ and $1 < c \leqslant d \leqslant |\boldsymbol{\ell}|$ $\boldsymbol{\ell}_{c:d} := (\ell_c, \dots, \ell_d)$, and let $(\boldsymbol{i}, \boldsymbol{j}) \in A(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$. This means that there exists $\sigma \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^1|, \dots, |\boldsymbol{k}^m|)$ with $\boldsymbol{i}\boldsymbol{j} = \sigma_*^{-1} \boldsymbol{k}^{1:m}$. Let

$$l := \begin{cases} 0 & \text{if } |\boldsymbol{i}| < \sigma(|\boldsymbol{k}^1|) \\ m & \text{if } \boldsymbol{j} = () \\ \text{s.t. } \sigma(|\boldsymbol{k}^{1:l}|) \leqslant |\boldsymbol{i}| < \sigma(|\boldsymbol{k}^{1:l+1}|) & \text{otherwise,} \end{cases}$$

which is exists and is unique since σ is an ordered shuffle. We then let $\boldsymbol{i}^h := \boldsymbol{k}^h$ for $h \leqslant l$ and for $h \geqslant l + 1$

$$\boldsymbol{i}^h \boldsymbol{j}^h := \boldsymbol{k}^h, \quad \sigma(|\boldsymbol{k}^{1:h-1}| + |\boldsymbol{i}^h|) \leqslant |\boldsymbol{i}| < \sigma(|\boldsymbol{k}^{1:h-1}| + |\boldsymbol{i}^h| + 1), \tag{19}$$

where $|\boldsymbol{i}^h|$ (and hence $\boldsymbol{i}^h$) is unique since σ is a shuffle. Now, it cannot be the case that for $h \geqslant l+1$ we have $|\boldsymbol{i}^h| = |\boldsymbol{k}^h|$, for this would violate the definition of l: this implies $|\boldsymbol{j}^h| \geqslant 1$. Moreover, we have $\boldsymbol{i} = \sigma_*^{-1} \boldsymbol{i}^{1:m}$ and $\boldsymbol{j} = \sigma_*^{-1} \boldsymbol{j}^{l+1:m}$, where we are defining the right-hand sides using the same expression as before (4), but by considering the numberings on $\boldsymbol{i}^{1:m}$, $\boldsymbol{j}^{l+1:m}$ to be those inherited as subtuples of $\boldsymbol{k}^{1:m}$: this is because $\boldsymbol{i}^h$ occupies the segment of $\boldsymbol{k}^{1:m}$ numbered with $[|\boldsymbol{k}^{1:h-1}| + 1, |\boldsymbol{k}^{1:h-1}| + |\boldsymbol{i}^h|]$, all of which σ maps into $[1, |\boldsymbol{i}|]$, by (19) and again by the shuffle property of σ; similarly, $\boldsymbol{j}^h$ occupies the segment numbered $[|\boldsymbol{k}^{1:h-1}| + |\boldsymbol{i}^h| + 1, |\boldsymbol{k}^{1:h}|]$ which gets mapped above $|\boldsymbol{i}|$. By construction σ (once domains are renumbered) shuffles $\boldsymbol{i}^{1:l}$ and $\boldsymbol{j}^{l+1:m}$ with order, since these are the tuples that contain the functions of $k^h_{|\boldsymbol{k}^h|}$, and $\boldsymbol{i}^{l+1,m}$ without order. If $\rho \in \mathrm{Sh}(n_1, \ldots, n_m)$ and $S \subseteq \{1, \ldots, n_1 + \cdots + n_m\}$, $\rho|_S$ is still a shuffle, with the additional order constraints on those functions of $(n_1 + \cdots + n_q)$ that belong to S: therefore, we have that $\boldsymbol{j} \in \overline{\mathrm{Sh}}(\boldsymbol{j}^{l+1}, \ldots, \boldsymbol{j}^m)$ and $\boldsymbol{i} \in \mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{i}^1, \ldots, \boldsymbol{i}^l), \mathrm{Sh}(\boldsymbol{i}^{l+1}, \ldots, \boldsymbol{i}^m))$. This shows $A(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m) \subseteq B(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$.

Conversely, let $(\boldsymbol{i}, \boldsymbol{j}) \in B(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$, with $l, \boldsymbol{i}^h, \boldsymbol{j}^h$ as in (18). $\boldsymbol{i}\boldsymbol{j}$ is obtained by an ordered shuffle of $\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m$: that the order of each $\boldsymbol{k}^h$, $h \leqslant l$ is preserved is immediate since $\boldsymbol{i}^h = \boldsymbol{k}^h$; that the order of each $\boldsymbol{k}^h$, $h > l$ is preserved is a consequence of the fact that the order of $\boldsymbol{i}^h$ is preserved, that the order of $\boldsymbol{j}^h$ is preserved and that $\boldsymbol{i}$ comes before $\boldsymbol{j}$ in the juxtaposition $\boldsymbol{i}\boldsymbol{j}$; that the shuffle of the $\boldsymbol{k}^h$'s is ordered is a consequence of the fact that the shuffles of $\boldsymbol{i}^1, \ldots, \boldsymbol{i}^l$ and $\boldsymbol{j}^{l+1}, \ldots, \boldsymbol{j}^{l+1}$ are ordered. This shows $B(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m) \subseteq A(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$; also note that in both inclusions multiplicities are indeed counted, since the correspondence between the underlying permutation in $A(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$ and the pair of underlying permutations in $B(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$ is bijective. $\square$

Example 1.2. We illustrate the idea behind this lemma with an example. Let

$$m = 4; \quad |\boldsymbol{k}^1| = 2, \ |\boldsymbol{k}^2| = 3, \ |\boldsymbol{k}^3| = 4, \ |\boldsymbol{k}^4| = 4$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 4 & 6 & 3 & 5 & 8 & 7 & 9 & 11 & 12 & 1 & 2 & 10 & 13 \end{pmatrix} \in \overline{\mathrm{Sh}}(2,3,4,4)$$

$$\sigma_*^{-1}(\boldsymbol{k}^1\boldsymbol{k}^2\boldsymbol{k}^3\boldsymbol{k}^4) = (\underbrace{k^4_1, k^4_2, k^2_1, k^1_1, k^2_2, \textcolor{red}{k^1_2}, k^3_1, \textcolor{red}{k^2_3}}_{=:\boldsymbol{i}}, \underbrace{k^3_2, k^3_3, k^3_4, \textcolor{red}{k^4_3}, \textcolor{red}{k^4_4}}_{=:\boldsymbol{j}})$$

$$(\boldsymbol{i}, \boldsymbol{j}) \in A(\boldsymbol{k}^1, \boldsymbol{k}^2, \boldsymbol{k}^3, \boldsymbol{k}^4),$$

where we have coloured in red the terminal elements of the functions of $\boldsymbol{k}^h$ (and recall that it is necessary to renumber $\boldsymbol{k}^1\boldsymbol{k}^2\boldsymbol{k}^3\boldsymbol{k}^4$ from 1 to 13 before applying (4), and then change the numbering back) and we have

$$(\boldsymbol{i}, \boldsymbol{j}) \in B(\boldsymbol{k}^1, \boldsymbol{k}^2, \boldsymbol{k}^3, \boldsymbol{k}^4) \quad \text{with } l = 2;$$

$$\boldsymbol{i}^1 = (k^1_1, k^1_2), \ \boldsymbol{i}^2 = (k^2_1, k^2_2, k^2_3), \ \boldsymbol{i}^3 = (k^3_1), \ \boldsymbol{i}^4 = (k^4_1, k^4_2);$$

$$\boldsymbol{j}^1 = (k^3_2, k^3_3, k^3_4), \ \boldsymbol{j}^2 = (k^4_3, k^4_4)$$

since $\boldsymbol{k}^h = \boldsymbol{i}^h$ for $h = 1, 2$, $\boldsymbol{k}^h = \boldsymbol{i}^h \boldsymbol{j}^h$ for $h = 3, 4$, and

$$\boldsymbol{j} \in \overline{\mathrm{Sh}}(\boldsymbol{j}^1, \boldsymbol{j}^2); \quad \boldsymbol{i} \in \mathrm{Sh}(\boldsymbol{a}, \boldsymbol{b})$$

$$\text{with } \boldsymbol{a} := (k_1^2, k_1^1, k_2^2, k_2^1, k_3^2) \in \overline{\text{Sh}}(\boldsymbol{i}^1, \boldsymbol{i}^2)$$

$$\boldsymbol{b} := (k_1^4, k_2^4, k_1^3) \in \text{Sh}(\boldsymbol{i}^3, \boldsymbol{i}^4).$$

Note that neither of the two functions of Sh above can be replaced with $\overline{\text{Sh}}$.

The reduced version of this Lemma 1.1 would involve restricting it to non-empty tuples $\boldsymbol{k}^1, \dots, \boldsymbol{k}^m$; we will need the dual of this statement. We will use the notation $((\boldsymbol{i}^1, \dots, \boldsymbol{i}^l), (\boldsymbol{i}^{l+1}, \dots, \boldsymbol{i}^m)) \in (\overline{\text{Sh}}^{-1}, \text{Sh}^{-1})(\text{Sh}^{-1}(\boldsymbol{i}))$ as a shorthand for $(\boldsymbol{i}^1, \dots, \boldsymbol{i}^l) \in \overline{\text{Sh}}^{-1}(\boldsymbol{a})$, $(\boldsymbol{i}^{l+1}, \dots, \boldsymbol{i}^m) \in \text{Sh}^{-1}(\boldsymbol{b}) : (\boldsymbol{a}, \boldsymbol{b}) \in \text{Sh}^{-1}(\boldsymbol{i})$, and m is fixed.

Corollary 1.3. *For I-valued tuples $\boldsymbol{i}, \boldsymbol{j}$ the following identity of multisets holds:*

$$\{\{(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m) \in \overline{\text{Sh}}^{-1}(\boldsymbol{ij}) \mid |\boldsymbol{k}^1|, \dots, |\boldsymbol{k}^m| \geq 1\}\}$$

$$= \{\{(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m) \in (I_*^\bullet)^m \mid (\boldsymbol{j}^{l+1}, \dots, \boldsymbol{j}^m) \in \overline{\text{Sh}}^{-1}(\boldsymbol{j});$$

$$((\boldsymbol{i}^1, \dots, \boldsymbol{i}^l), (\boldsymbol{i}^{l+1}, \dots, \boldsymbol{i}^m)) \in (\overline{\text{Sh}}^{-1}, \text{Sh}^{-1})(\text{Sh}^{-1}(\boldsymbol{i})); \quad (20)$$

$$\boldsymbol{k}^h = \boldsymbol{i}^h, \ h \leq l; \ \boldsymbol{k}^h = \boldsymbol{i}^h \boldsymbol{j}^h, \ |\boldsymbol{j}^h| \geq 1, \ h \geq l+1;$$

$$l = 0, \dots, m\}\}.$$

Proof of Corollary 1.3. Let $C(\boldsymbol{i}, \boldsymbol{j})$ denote the first multiset above and $D(\boldsymbol{i}, \boldsymbol{j})$ the second, and recall the names $A(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m), B(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$ for the sets of Lemma 1.1. We then have (taking into account multiplicities)

$$(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m) \in C(\boldsymbol{i}, \boldsymbol{j}) \iff (\boldsymbol{i}, \boldsymbol{j}) \in A(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$$

$$\iff (\boldsymbol{i}, \boldsymbol{j}) \in B(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)$$

$$\iff (\boldsymbol{k}^1, \dots, \boldsymbol{k}^m) \in D(\boldsymbol{i}, \boldsymbol{j}),$$

thus concluding the proof. □

Next we discuss another combinatorial relation involving ordered and unordered shuffles; similarly to the earlier lemma, we provide a diagram and an example (unrelated to each other) to help explain the idea. $\bigsqcup$ will denote disjoint union of multisets, for example, if the same tuple appears in sets corresponding to two different functions of σ, it should be counted twice.

Lemma 1.4. *Let $n := n_1 + \cdots + n_m$, $n^l := n_1 + \cdots + n_l$ for $l = 1, \dots, m$, and $\boldsymbol{k}^1, \dots, \boldsymbol{k}^n \in I_*^\bullet$. We have*

$$\text{Sh}(\overline{\text{Sh}}(\boldsymbol{k}^1, \dots, \boldsymbol{k}^{n_1}), \dots, \overline{\text{Sh}}(\boldsymbol{k}^{n^{m-1}+1}, \dots, \boldsymbol{k}^n))$$

$$= \bigsqcup_{\pi \in \text{Sh}(n_1, \dots, n_m)} \overline{\text{Sh}}(\boldsymbol{k}^{\pi^{-1}(1)}, \dots, \boldsymbol{k}^{\pi^{-1}(n)}). \quad (21)$$
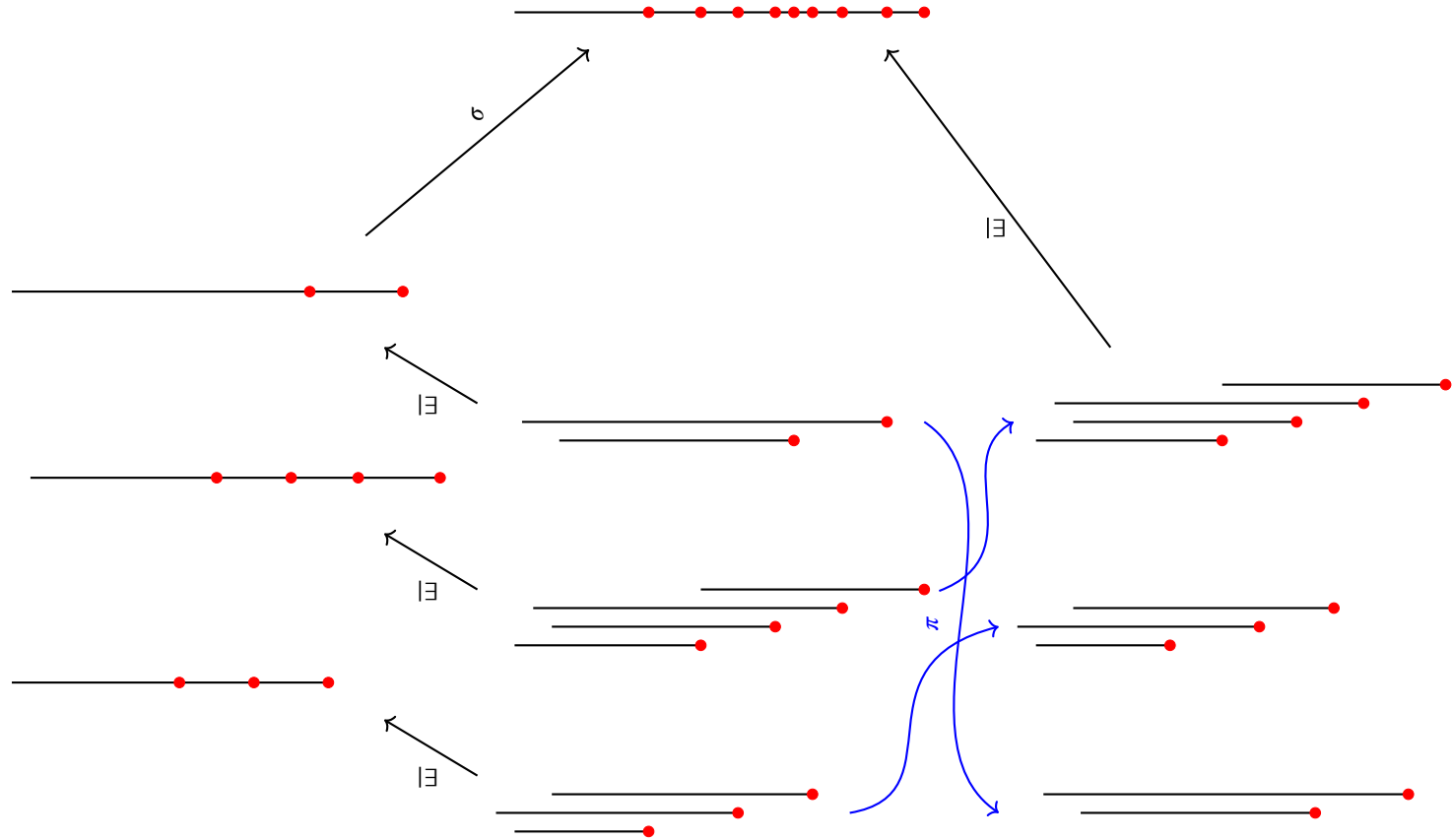
Proof of Lemma 1.4. Let $N_l := |\boldsymbol{k}^{n^{l-1}+1}| + \cdots + |\boldsymbol{k}^{n^l}|$ for $l = 1, \ldots, m$, and $N^l := N_1 + \cdots + N_l$ for $l = 1, \ldots, m$, and $N := N^m$. We have

$$\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^{n_1}), \ldots, \overline{\mathrm{Sh}}(\boldsymbol{k}^{n^{m-1}+1}, \ldots, \boldsymbol{k}^n))$$

$$= \{\{\sigma_*^{-1}(\boldsymbol{h}^1 \ldots \boldsymbol{h}^m) \mid \sigma \in \mathrm{Sh}(N_1, \ldots, N_m);$$

$$\boldsymbol{h}^l \in \overline{\mathrm{Sh}}(\boldsymbol{k}^{n^{l-1}+1}, \ldots, \boldsymbol{k}^{n^l}), \; l = 1, \ldots, m\}\}$$

$$= \{\{\sigma_*^{-1}(\boldsymbol{h}^1 \ldots \boldsymbol{h}^m) \mid \sigma \in \mathrm{Sh}(N_1, \ldots, N_m);$$

$$\boldsymbol{h}^l = \rho_{l*}^{-1}(\boldsymbol{k}^{n^{l-1}+1} \ldots \boldsymbol{k}^{n^l});$$

$$\rho_l \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^{n^{l-1}+1}|, \ldots, |\boldsymbol{k}^{n^l}|), \; l = 1, \ldots, m\}\}.$$

Now, denoting $(\rho_1, \ldots, \rho_m)$ the element of $\mathfrak{S}_N$ which acts on $\{N^{l-1} + 1, \ldots, N^l\}$ with ρ_l, we continue the chain of identities

$$= \{\{\sigma_*^{-1}(\rho_1, \ldots, \rho_m)_*^{-1}(\boldsymbol{k}^1 \ldots \boldsymbol{k}^{n_m}) \mid \sigma \in \mathrm{Sh}(N_1, \ldots, N_m);$$

$$\rho_l \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^{n^{l-1}+1}|, \ldots, |\boldsymbol{k}^{n^l}|), \; l = 1, \ldots, m\}\}$$

$$= \{\{(\sigma \circ (\rho_1, \ldots, \rho_m))_*^{-1}(\boldsymbol{k}^1 \ldots \boldsymbol{k}^{n_m}) \mid \sigma \in \mathrm{Sh}(N_1, \ldots, N_m);$$

$$\rho_l \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^{n^{l-1}+1}|, \ldots, |\boldsymbol{k}^{n^l}|), \; l = 1, \ldots, m\}\}$$

since $(\rho_1^{-1}, \ldots, \rho_m^{-1}) = (\rho_1, \ldots, \rho_m)^{-1}$ and thanks to the composition rule (5). Let π denote the restriction of $\sigma \circ (\rho_1, \ldots, \rho_m)$ to the set

$$T := \{t_1, \ldots, t_n\}, \quad t_l := |\boldsymbol{k}^1| + \cdots + |\boldsymbol{k}^l|.$$

Since the functions of $\boldsymbol{k}^h$ are all non-empty, T is a subset of $\{1, \ldots, N\}$ of cardinality n, so after renumbering it we can consider π as an element of $\mathfrak{S}_n$. Now, since ρ_l is an ordered shuffle, it preserves the ordering of $\{|\boldsymbol{k}^{n^{l-1}+1}|, \ldots, |\boldsymbol{k}^{n^{l-1}+1}| + \cdots + |\boldsymbol{k}^{n^l}|\}$, and since σ is a shuffle, it preserves the ordering $\{\rho_l(|\boldsymbol{k}^{n^{l-1}+1}|), \ldots, \rho_l(|\boldsymbol{k}^{n^l}|)\} \subseteq \{N^{l-1} + 1, \ldots, N^l\}$. These two facts imply $\pi \in \mathrm{Sh}(n_1, \ldots, n_m)$, and we have

$$\{\sigma \circ (\rho_1, \ldots, \rho_m) \mid \rho_l \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^{n^{l-1}+1}|, \ldots, |\boldsymbol{k}^{n^l}|), \; l = 1, \ldots, m\}$$

$$= \{\tau \in \mathrm{Sh}(|\boldsymbol{k}^1|, \ldots, |\boldsymbol{k}^n|) \mid \tau(t_{\pi^{-1}(1)}) < \ldots < \tau(t_{\pi^{-1}(n)})\}$$

$$=: \overline{\mathrm{Sh}^\pi}(|\boldsymbol{k}^1|, \ldots, |\boldsymbol{k}^n|)$$

since any $\tau \in \mathrm{Sh}(|\boldsymbol{k}^1|, \ldots, |\boldsymbol{k}^n|)$ with $\tau(t_{\pi^{-1}(1)}) < \ldots < \tau(t_{\pi^{-1}(n)})$ for some $\pi \in \mathfrak{S}_n$ factors uniquely as $\sigma \circ (\rho_1, \ldots, \rho_m)$ with σ acting on T with π: this is evident from the fact that each ρ_l acts on the segment $[N^{l-1} + 1, N^l]$ and σ acts on the whole segment $[1, N]$ but without altering the order in

each $[N^{l-1} + 1, N^l]$. This implies

$$\{\sigma \circ (\rho_1, \dots, \rho_m) \mid \sigma \in \mathrm{Sh}(N_1, \dots, N_m);$$

$$\rho_l \in \overline{\mathrm{Sh}}(|\boldsymbol{k}^{n^{l-1}+1}|, \dots, |\boldsymbol{k}^{n^l}|), \ l = 1, \dots, m\}$$

$$= \bigsqcup_{\pi \in \mathrm{Sh}(n_1, \dots, n_m)} \overline{\mathrm{Sh}^\pi}(|\boldsymbol{k}^1|, \dots, |\boldsymbol{k}^n|)$$

because as σ ranges over $\mathrm{Sh}(N_1, \dots, N_m)$ all functions of $\mathrm{Sh}(n_1, \dots, n_m) \ni \pi$ are obtained, and the functions of $\overline{\mathrm{Sh}^\pi}$ are mutually disjoint since the ordered shuffle relations imposed by different functions of π are mutually exclusive. Since

$$\overline{\mathrm{Sh}}(\boldsymbol{k}^{\pi^{-1}(1)}, \dots, \boldsymbol{k}^{\pi^{-1}(n)}) = \{\{\tau_*^{-1}(\boldsymbol{k}^1 \dots \boldsymbol{k}^n) \mid \tau \in \overline{\mathrm{Sh}^\pi}(|\boldsymbol{k}^1|, \dots, |\boldsymbol{k}^n|)\}\},$$

the proof is concluded. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Example 1.5. We illustrate the idea behind this lemma with an example. Let

$$n_1 = 2, \ n_2 = 1; \quad |\boldsymbol{k}^1| = 3, \ |\boldsymbol{k}^2| = 2, \ |\boldsymbol{k}^3| = 3 \qquad\qquad (22)$$

and with the notations of the proof let

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix} \in \overline{\mathrm{Sh}}(3, 2), \ \rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \in \overline{\mathrm{Sh}}(3),$$

$$\Rightarrow (\rho_1, \rho_2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 3 & 4 & 2 & 5 & 6 & 7 & 8 \end{pmatrix}$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 6 & 7 & 8 & 1 & 4 & 5 \end{pmatrix} \in \mathrm{Sh}(5, 3)$$

$$\Rightarrow \tau := \sigma \circ (\rho_1, \rho_2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 6 & 7 & 3 & 8 & 1 & 4 & 5 \end{pmatrix} \in \overline{\mathrm{Sh}^\pi}(3, 2, 3)$$

$$\text{with } \pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathrm{Sh}(2, 1),$$

because the restriction of $\sigma \circ (\rho_1, \rho_2)$ to $\{|\boldsymbol{k}^1|, |\boldsymbol{k}^1 + \boldsymbol{k}^2|, |\boldsymbol{k}^1 + \boldsymbol{k}^2 + \boldsymbol{k}^3|\}$ is $\begin{pmatrix} 3 & 5 & 8 \\ 7 & 8 & 5 \end{pmatrix}$, which coincides with π after renumbering domain and codomain. Also note how, given τ and the numbers (22) one can recover σ, ρ_1, ρ_2: firstly obtain ρ_1, ρ_2 considering how τ orders the segments $[1,5]$, $[6,8]$ (and renumbering) and $\sigma = \tau \circ (\rho_1, \dots, \rho_m)$. We therefore have, writing terminal elements in red

$$(\sigma \circ (\rho_1, \rho_2))_*^{-1}(\boldsymbol{k}_1 \boldsymbol{k}_2 \boldsymbol{k}_3) = (k_1^3, k_1^1, k_1^2, k_2^3, \textcolor{red}{k_3^3}, k_2^1, \textcolor{red}{k_3^1}, \textcolor{red}{k_2^2})$$

$$\in \mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{k}^1, \boldsymbol{k}^2), \overline{\mathrm{Sh}}(\boldsymbol{k}^3))$$

$$\in \overline{\mathrm{Sh}}(\boldsymbol{k}^3, \boldsymbol{k}^1, \boldsymbol{k}^2) = \overline{\mathrm{Sh}}(\boldsymbol{k}^{\pi^{-1}(1)}, \boldsymbol{k}^{\pi^{-1}(2)}, \boldsymbol{k}^{\pi^{-1}(3)}),$$

which can also be written as

$$(k_1^3, k_1^1, k_1^2, k_2^3, k_3^3, k_2^1, k_3^1, k_2^2) = \eta_*^{-1}(\boldsymbol{k}^3, \boldsymbol{k}^1, \boldsymbol{k}^2) \in \overline{\mathrm{Sh}}(\boldsymbol{k}^3, \boldsymbol{k}^1, \boldsymbol{k}^2)$$

$$\text{with } \eta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 4 & 5 & 2 & 6 & 7 & 3 & 8 \end{pmatrix} \in \overline{\mathrm{Sh}}(3,3,2).$$

We primarily use Lemma 1.4 in the two cases $n_1 = \cdots = n_m = 1$ with m arbitrary, and $m = 2$ with n_1, n_2 arbitrary; the former admits the following concise reformulation. Given a vector space W, let $\odot$ denote symmetric tensor product, and

$$\odot_m := \frac{1}{m!} \sum_{\pi \in \mathfrak{S}_m} \pi_* \colon W^{\otimes m} \twoheadrightarrow W^{\odot m}, \quad w_1 \otimes \cdots \otimes w_m \mapsto w_1 \odot \cdots \odot w_m \tag{23}$$

denote the symmetrisation map. When referring to the external tensor product we will replace the symbol $\odot$ with $\boxdot$.

Corollary 1.6. *The diagram*

$$\tag{24}$$

commutes.

Proof. The statement in coordinates reads

$$\sum_{\substack{\pi \in \mathfrak{S}_m \\ \boldsymbol{h} \in \overline{\mathrm{Sh}}(\boldsymbol{k}^{\pi(1)}, \ldots, \boldsymbol{k}^{\pi(m)})}} a^{\boldsymbol{h}} = \sum_{\boldsymbol{k} \in \mathrm{Sh}(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)} a^{\boldsymbol{k}}$$

for non-empty tuples $\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m$. Indeed, we have

$$(m! \boxdot \widetilde{\Delta}_{\sqcup\!\sqcup} a)^{\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m} = \left(\sum_{\substack{\pi \in \mathfrak{S}_m \\ (\widetilde{a})_{\sqcup\!\sqcup}^m}} \pi_*(a_{(1)} \boxtimes \cdots \boxtimes a_{(m)}) \right)^{\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m}$$

$$= \sum_{\pi \in \mathfrak{S}_m} \sum_{(\widetilde{a})_{\sqcup\!\sqcup}^m} a_{(\pi(1))}^{\boldsymbol{k}^1} \cdots a_{(\pi(m))}^{\boldsymbol{k}^m}$$

$$= \sum_{\pi \in \mathfrak{S}_m} \sum_{(\widetilde{a})_{\sqcup\!\sqcup}^m} a_{(1)}^{\boldsymbol{k}^{\pi(1)}} \cdots a_{(m)}^{\boldsymbol{k}^{\pi(m)}}$$

$$= \sum_{\pi \in \mathfrak{S}_m} (\widetilde{\Delta}_{\sqcup}^m a)^{\boldsymbol{k}^{\pi(1)},\ldots,\boldsymbol{k}^{\pi(m)}}$$

$$= \sum_{\substack{\pi \in \mathfrak{S}_m \\ \boldsymbol{h} \in \overline{\mathrm{Sh}}(\boldsymbol{k}^{\pi(1)},\ldots,\boldsymbol{k}^{\pi(m)})}} a^{\boldsymbol{h}}.$$

To prove the claim in coordinates we must show the identity of sets

$$\bigsqcup_{\pi \in \mathfrak{S}_m} \overline{\mathrm{Sh}}(\boldsymbol{k}^{\pi(1)}, \ldots, \boldsymbol{k}^{\pi(m)}) = \mathrm{Sh}(\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m)$$

for tuples $\boldsymbol{k}^1, \ldots, \boldsymbol{k}^m$ of positive order. This is precisely Lemma 1.4 with $n_1 = \cdots = n_m = 1$, since $\mathrm{Sh}(n_1, \ldots, n_m) = \mathfrak{S}_m$ and we may replace π^{-1} with π. $\qquad\square$

2 | WEAKLY GEOMETRIC ROUGH PATHS

We denote $T^N(V)$ the vector subspace of $T(V)$ given by all tensors of degree $\leqslant N$, and super/subscripts of $\leqslant N, \geqslant M$ denote truncations of the algebra to tensors of the degrees expressed (for example, for $a \in T(V)$ to belong to $T^N(V)$ it means that $a^n = 0$ for $n > N$, or equivalently $a = a^{\leqslant N}$). We will similarly use $[M, N]$ as a super/subscript to denote tensors of degrees n with $M \leqslant n \leqslant N$. A *control* on $[0, T]$ is a continuous function ω defined on the simplex $\Delta_T := \{(s, t) \in [0, T]^2 \mid s \leqslant t\}$, s.t. $\omega(t, t) = 0$ for $0 \leqslant t \leqslant T$ and is superadditive, that is, $\omega(s, u) + \omega(u, t) \leqslant \omega(s, t)$ for $0 \leqslant s \leqslant u \leqslant t \leqslant T$.

Definition 2.1 (Weakly geometric rough path). Let $T > 0$, $p \geqslant 1$ and ω be a control on $[0, T]$. A *p-weakly geometric rough path* $\boldsymbol{X}$ controlled by ω, defined on $[0, T]$ and with values in V may be defined as a continuous map

$$\boldsymbol{X} \colon \Delta_T \to T^{\lfloor p \rfloor}(V) \tag{25}$$

with $\boldsymbol{X}^0 = 1$ and satisfying the following properties, which we first present in coordinate-free form and subsequently in coordinates with respect to a basis of V:

Regularity. $\sup_{0 \leqslant s < t \leqslant T} \frac{|\boldsymbol{X}_{st}^n|}{\omega(s,t)^{n/p}} < \infty$, or $\sup_{0 \leqslant s < t \leqslant T} \frac{|\boldsymbol{X}_{st}^{\boldsymbol{k}}|}{\omega(s,t)^{|\boldsymbol{k}|/p}} < \infty$ for $n = |\boldsymbol{k}| = 1, \ldots \lfloor p \rfloor$;

Multiplicativity. $(\boldsymbol{X}_{su} \otimes \boldsymbol{X}_{ut})^{\leqslant \lfloor p \rfloor} = \boldsymbol{X}_{st}$, that is, $\boldsymbol{X}_{st}^{\boldsymbol{k}} = \sum_{(\boldsymbol{i},\boldsymbol{j})=\boldsymbol{k}} \boldsymbol{X}_{su}^{\boldsymbol{i}} \boldsymbol{X}_{ut}^{\boldsymbol{j}}$ for $|\boldsymbol{k}| \leqslant \lfloor p \rfloor$ and $0 \leqslant s \leqslant u \leqslant t \leqslant T$;

Integration by parts. $\boldsymbol{X}_{st} \boxtimes \boldsymbol{X}_{st} = \Delta_{\sqcup} \boldsymbol{X}_{st}$, or $\boldsymbol{X}_{st}^{\boldsymbol{i}} \boldsymbol{X}_{st}^{\boldsymbol{j}} = \sum_{\boldsymbol{k} \in \mathrm{Sh}(\boldsymbol{i},\boldsymbol{j})} \boldsymbol{X}_{st}^{\boldsymbol{k}}$ for all $0 \leqslant s \leqslant t \leqslant T$.

Let $\mathscr{C}_\omega^p([0, T], V)$ denote the set of all such maps.

In the following we will refer to such objects as 'rough paths', dropping the 'weakly geometric', since these are the only rough paths that we will be considering in this paper. We will sometimes refer to the third property above as geometricity, since it distinguishes weakly geometric rough paths among the more general branched rough paths, which we do not treat here. We will always denote rough paths in bold. The last condition is usually stated by saying that $\boldsymbol{X}$ takes values in the group $G^{\lfloor p \rfloor}(V)$, defined in [5, Definition 2.9]. We will denote $X := \boldsymbol{X}^1$ the *trace* of X: when

equipped with an initial value X_0 (which will often be provided), this is an element of $C_\omega^p([0, T], V)$ defined as the set of continuous paths $Y : [0, T] \to V$ s.t., denoting $Y_{st} := Y_t - Y_s$ (a notation that will be used for paths throughout)

$$\sup_{0 \leqslant s < t \leqslant T} \frac{|Y_{st}|}{\omega(s, t)^{1/p}} < \infty. \tag{26}$$

It is sufficient to define X to take values in $T^{\lfloor p \rfloor}(V)$, as [17, Theorem 2.2.1] shows that there exists a unique extension of $\hat{X}$ of X to $T(V)$ which satisfies the above three properties, and for $m > \lfloor p \rfloor$ is given by

$$\hat{X}_{st}^m = \lim_{n \to \infty} \left(\bigotimes_{[u,v] \in \pi_n} X_{uv} \right)^m, \tag{27}$$

where $(\pi_n)_n$ is any sequence of partitions on $[s, t]$ with vanishing step size as $n \to \infty$.

The following proposition states that the symmetric part of a weakly geometric rough path is entirely determined by its trace. Given $\ell \in \mathcal{L}(T^N(V), U)$ ($\mathcal{L}$ denotes the space of linear maps) and $a \in T(V)$ we will denote $\langle \ell, a \rangle = \ell(a)$ the evaluation of ℓ on a. We will always identify $\mathcal{L}(\mathbb{R}, U) = U$ by setting $\ell \mapsto \ell(1)$.

Proposition 2.2. *For $X \in \mathscr{C}_\omega^p([0, T], V)$ we have $n! \odot_n X_{st}^n = X_{st}^{\otimes n}$.*

Proof. We proceed by induction on n. For $n = 0, 1$ there is nothing to prove. For the inductive step we will need the following fact: each $\pi \in \mathfrak{S}_{n+1}$ can be expressed uniquely as $\sigma \circ \rho$ with ρ in the stabiliser of $n + 1$ (a subgroup of $\mathfrak{S}_{n+1}$ isomorphic to $\mathfrak{S}_n$) and $\sigma \in \text{Sh}(n, 1)$: indeed, if $m := \pi(n + 1)$ we may set

$$\sigma(k) := \begin{cases} k & 1 \leqslant k \leqslant m - 1 \\ k + 1 & m \leqslant k \leqslant n \\ m & k = n + 1 \end{cases} \qquad \rho(k) := \begin{cases} \pi(k) & 1 \leqslant \pi(k) \leqslant m - 1 \\ \pi(k) - 1 & m \leqslant \pi(k) \leqslant n \\ n + 1 & k = n + 1 \end{cases} .$$

Uniqueness follows from a counting argument, since there are $n!$ choices for ρ and $n + 1$ for σ. We then compute

$$(n + 1)! \odot_{n+1} X_{st}^{n+1} = \sum_{\pi \in \mathfrak{S}_{n+1}} \pi_* X_{st}^{n+1}$$

$$= \sum_{\rho \in \mathfrak{S}_n} \rho_* \sum_{\sigma \in \text{Sh}(n,1)} \sigma_* X_{st}^{n+1}$$

$$= \sum_{\rho \in \mathfrak{S}_n} \rho_* (\Delta_{\sqcup\!\sqcup} X_{st})^{n,1}$$

$$= \sum_{\rho \in \mathfrak{S}_n} \rho_* X_{st}^n \otimes X_{st}$$

$$= X_{st}^{\otimes(n+1)},$$

where we have used the geometricity axiom and the inductive hypothesis. □

We proceed to define the objects that can be regarded as dual to rough paths, original to [12]. In what follows we will write $\approx_m$ between two real-valued quantities dependent on $0 \leqslant s \leqslant t \leqslant T$ to mean that their difference lies in $O(\omega(s,t)^{m/p})$ as $t \searrow s$, and simply $\approx$ to mean $\approx_{\lfloor p \rfloor + 1}$. We will frequently use the following properties, which are trivial to check:

$$
\begin{aligned}
a_{st} \approx_m b_{st} \approx_n c_{st} &\Rightarrow a_{st} \approx_{n \wedge m} c_{st} \\
a_{st} \approx_m b_{st}, \; c_{st} \approx_n 0 &\Rightarrow a_{st} c_{st} \approx_{m+n} b_{st} c_{st}
\end{aligned}
\tag{28}
$$

from which we deduce more generally

$$
a_{st} \approx_{m_1} b_{st}, \; a_{st}, b_{st} \approx_{n_1} 0, \quad c_{st} \approx_{m_2} d_{st}, \; c_{st}, d_{st} \approx_{n_2} 0
$$
$$
\Rightarrow a_{st} c_{st} \approx_{m_1 + n_2} b_{st} c_{st} \approx_{m_2 + n_1} b_{st} d_{st}
\tag{29}
$$
$$
\Rightarrow a_{st} c_{st} \approx_{(m_1 + n_2) \wedge (m_2 + n_1)} b_{st} d_{st}.
$$

If a continuous map $\widetilde{X} : \Delta_T \to T^{\lfloor p \rfloor}(V)$ satisfies the regularity and integration by parts conditions, and satisfies the multiplicativity condition with a '$\approx$' replacing the '$=$' (*almost multiplicative*), it defines a rough path by [17, Theorem 3.3.1], by taking the limit (27) (with respect to $\widetilde{X}$), and this rough path X is unique with the property that $\widetilde{X}_{st} \approx X_{st}$. The following lemma tells us that this is also true if the integration by parts condition only holds with an $\approx$ (*almost geometric*). In light of this, we will break with the literature in defining an *almost rough path* as an $\widetilde{X}$ that satisfies the regularity condition in Definition 2.1 and is almost multiplicative and almost geometric.

Proposition 2.3 (Almost rough paths). *Let $\widetilde{X}$ be a V-valued almost p-rough path. Then there exists a unique p-rough path X with the property that $X_{st} \approx \widetilde{X}_{st}$.*

Proof. We use that the shuffle algebra is free abelian over the Lyndon words [21, Theorem 6.1] to define an intermediate $\overline{X}$: set $\overline{X}^h := \widetilde{X}^h$ if h is a Lyndon word with $|h| \leqslant \lfloor p \rfloor$, and for a tuple k with $|k| < \lfloor p \rfloor$ expressed (uniquely up to order of factors) as $\sum_\lambda c_\lambda k_\lambda^1 \sqcup \cdots \sqcup k_\lambda^{n_\lambda}$ with the functions of k_λ^j (not necessarily distinct) Lyndon words set

$$
\overline{X}^k := \sum_\lambda c_\lambda \overline{X}^{k_\lambda^1} \cdots \overline{X}^{k_\lambda^{n_\lambda}} = \sum_\lambda c_\lambda \widetilde{X}^{k_\lambda^1} \cdots \widetilde{X}^{k_\lambda^{n_\lambda}} \approx \sum_\lambda c_\lambda \langle k_\lambda^1 \sqcup \cdots \sqcup k_\lambda^{n_\lambda}, \widetilde{X} \rangle = \widetilde{X}^k
$$

since $\widetilde{X}$ is almost geometric. $\overline{X}$ is then $\approx \widetilde{X}$, it satisfies integration by parts (exactly) by construction since if $i = \sum_\mu a_\mu i_\mu^1 \sqcup \cdots \sqcup i_\mu^{m_\mu}$, $j = \sum_\nu b_\nu j_\nu^1 \sqcup \cdots \sqcup j_\nu^{n_\nu}$ with $a_\mu, b_\nu \in \mathbb{Q}$ and i_μ^q, j_ν^r Lyndon

$$
\overline{X}^i \overline{X}^j = \sum_{\mu, \nu} a_\mu b_\nu \overline{X}^{i_\mu^1} \cdots \overline{X}^{i_\mu^{m_\mu}} \overline{X}^{j_\nu^1} \cdots X^{j_\nu^{n_\nu}}
$$

$$
= \left\langle \sum_{\mu, \nu} a_\mu b_\nu i_\mu^1 \sqcup \cdots \sqcup j_\mu^{m_\mu} \sqcup j_\nu^1 \sqcup \cdots \sqcup j_\nu^{n_\nu}, \overline{X} \right\rangle = \langle i \sqcup j, \overline{X} \rangle
$$

and is still almost multiplicative since

$$\overline{X}^k \approx X_{st}^k = \sum_{(i,j)=k} X_{su}^i X_{ut}^j \approx \sum_{(i,j)=k} \overline{X}_{su}^i \overline{X}_{ut}^j.$$

Existence of X then follows immediately by applying the above-referenced result to $\overline{X}$ and uniqueness follows from the fact that if X' is a second p-rough path satisfying the statement of this proposition, then we have $X' \approx \widetilde{X} \approx \overline{X} \Rightarrow X' = X$ again by the same result. $\qquad \square$

Definition 2.4. Let X be as above and U another vector space. An U-valued X-controlled path $\overline{H}$ is an element of $C_\omega^p([0,T], \mathcal{L}(T^{\lfloor p \rfloor - 1}(V), U))$ (where ω is the control for X) s.t. for $n = 0, \ldots, \lfloor p \rfloor - 2$ and each $a \in V^{\otimes n}$

$$\langle \overline{H}_{n;t}, a \rangle \approx_{\lfloor p \rfloor - n} \langle \overline{H}_{[n, \lfloor p \rfloor - 1];s}, X_{st}^{\leqslant \lfloor p \rfloor - 1 - n} \otimes a \rangle. \tag{30}$$

Here $\overline{H}_n$ denotes the nth level of $\overline{H}$. Denote $\mathcal{D}_X(U)$ the vector space of all U-valued X-controlled paths.

We will always denote controlled paths with an overline. The maps $\overline{H}_n := \overline{H}|_{V^{\otimes n}}$ are known as the *Gubinelli derivatives* of H and $H := \overline{H}_0 \in U$ is called the *trace* of $\overline{H}$ (note the discrepancy with rough paths: for these the trace is the order-1 component). Note that the defining condition only involves $X^{\leqslant \lfloor p \rfloor - 1}$, and that it holds automatically at level $\lfloor p \rfloor - 1$ by regularity of $\overline{H}$. In coordinates it reads

$$\overline{H}_{\beta;t}^k \approx_{\lfloor p \rfloor - |\beta|} \sum_{|\alpha|=0}^{\lfloor p \rfloor - 1 - |\beta|} \overline{H}_{(\alpha,\beta);s}^k X_{st}^\alpha, \quad 0 \leqslant |\beta| \leqslant \lfloor p \rfloor - 2. \tag{31}$$

Here the superscript k refers to the value of $\overline{H}$ in U (and will often be omitted when unnecessary), and the sum is not only on the length $|\alpha|$ of the tuple α, but on the tuple itself. For the branched version of this definition, see [14].

An important case is when $U = \mathcal{L}(V, W)$ for another vector space W: by the tensor-hom adjunction we then have

$$\mathcal{L}(T^{\lfloor p \rfloor - 1}(V), \mathcal{L}(V, W)) = \mathcal{L}(T^{\lfloor p \rfloor - 1}(V) \otimes V, W) = \mathcal{L}\left(\bigoplus_{n=1}^{\lfloor p \rfloor} V^{\otimes n}, W \right). \tag{32}$$

We will use angle brackets and coordinate notation for linear maps accordingly, that is, the last slot in a bracket or in a tuple will refer to the copy of V in the target space of the original linear map. We will call controlled paths valued in $\mathcal{L}(V, W)$ W-*valued controlled integrands*, and we may rewrite (30) as

$$\langle \overline{H}_t, b \rangle \approx_{\lfloor p \rfloor - n + 1} \langle \overline{H}_{[n, \lfloor p \rfloor];s}, X_{st}^{[0, \lfloor p \rfloor - n]} \otimes b \rangle \in W, \quad b \in V^{\otimes n}, \quad n = 1, \ldots, \lfloor p \rfloor - 1 \tag{33}$$

or in coordinates

$$\overline{H}_{\beta;t}^k \approx_{\lfloor p \rfloor - |\beta| + 1} \sum_{|\alpha|=0}^{\lfloor p \rfloor - |\beta|} \overline{H}_{(\alpha,\beta);s}^k X_{st}^\alpha, \quad 1 \leqslant |\beta| \leqslant \lfloor p \rfloor - 1. \tag{34}$$

The next example contains a very important example of controlled path.

Example 2.5 (Smooth functions of X). Let $F \in C^\infty(V, U)$, then

$$t \mapsto (F(X_t), DF(X_t), \dots, D^{\lfloor p \rfloor - 1}F(X_t)) \in \mathcal{L}(T^{\lfloor p \rfloor - 1}(V), U) \tag{35}$$

is an X-controlled path, which we denote by $\overline{F}(X)$. Indeed, denoting by $\partial_\gamma F$ the order-$|\gamma|$ partial derivative of F in the directions of the chosen basis determined by the tuple γ, we have, for $0 \leqslant |\beta| \leqslant \lfloor p \rfloor - 2$

$$\overline{F}(X_t)_\beta - \sum_{|\alpha|=0}^{\lfloor p \rfloor - 1 - |\beta|} \overline{F}(X_s)_{(\alpha,\beta)} \boldsymbol{X}_{st}^\alpha$$

$$= \partial_\beta F(X_t) - \sum_{|\alpha|=0}^{\lfloor p \rfloor - 1 - |\beta|} \partial_{\alpha,\beta} F(X_s) \boldsymbol{X}_{st}^\alpha$$

$$= \partial_\beta F(X_t) - \sum_{n=0}^{\lfloor p \rfloor - 1 - |\beta|} \frac{1}{n!} \partial_{\alpha,\beta} F(X_s) X_{st}^{\alpha_1} \cdots X_{st}^{\alpha_n}$$

$$\approx_{\lfloor p \rfloor - |\beta|} 0,$$

where we have used Proposition 2.2 together with the symmetry of higher differentials and Taylor's approximation. Note that the symmetry of Gubinelli derivatives is a special feature of this kind of controlled path, and cannot be expected to hold in general. When $U = \mathcal{L}(V, W)$ we shall call F an W-*valued 1-form*, and we adopt the convention

$$\langle \overline{F}(X), v_1 \otimes \cdots \otimes v_{n+1} \rangle = D^n F(X)(v_1, \dots, v_n)(v_{n+1}) \in W$$

or in coordinates $\overline{F}(X)_{\alpha,\beta} = \partial_\alpha F_\beta(X)$.

The next lemma is necessary for the definition of rough integral.

Lemma 2.6. *Let $X \in \mathscr{C}_\omega^p([0,T], V)$ and $\overline{H} \in \mathscr{D}_X(\mathcal{L}(V, W))$. Define, for $0 \leqslant s \leqslant t \leqslant T$*

$$\Xi_{st} := \langle \overline{H}_s, \boldsymbol{X}_{st}^{\geqslant 1} \rangle \in W, \tag{36}$$

where the evaluation is taken under the identification (32). Then Ξ is almost additive: for all $0 \leqslant s \leqslant u \leqslant t \leqslant T$

$$\Xi_{st} - \Xi_{su} - \Xi_{ut} \approx 0. \tag{37}$$

Proof. Using the multiplicativity axiom, the regularity of $\overline{H}$ and (34) (together with (28)) we may write

$$\Xi_{st} - \Xi_{su} - \Xi_{ut}$$

$$= \sum_{|\gamma|=1}^{\lfloor p \rfloor} (\overline{H}_{\gamma;s} \boldsymbol{X}_{st}^\gamma - \overline{H}_{\gamma;s} \boldsymbol{X}_{su}^\gamma - \overline{H}_{\gamma;u} \boldsymbol{X}_{ut}^\gamma)$$

$$= \sum_{|\gamma|=1}^{\lfloor p \rfloor} (\overline{H}_{\gamma;s}(X_{st}^{\gamma} - X_{su}^{\gamma} - X_{ut}^{\gamma}) - \overline{H}_{\gamma;su}X_{ut}^{\gamma})$$

$$= \sum_{|\gamma|=1}^{\lfloor p \rfloor} \left(\overline{H}_{\gamma;s} \sum_{\substack{(\alpha,\beta)=\gamma \\ |\alpha|,|\beta| \geqslant 1}} X_{su}^{\alpha} X_{ut}^{\beta} - \overline{H}_{\gamma;su}X_{ut}^{\gamma} \right)$$

$$\approx \sum_{\substack{|\alpha|,|\beta| \geqslant 1 \\ |\alpha|+|\beta| \leqslant \lfloor p \rfloor}} \overline{H}_{(\alpha,\beta);s}X_{su}^{\alpha} X_{ut}^{\beta} - \sum_{|\varepsilon|=1}^{\lfloor p \rfloor - 1} \overline{H}_{\varepsilon;su}X_{ut}^{\varepsilon}$$

$$\approx \sum_{\substack{|\alpha|,|\beta| \geqslant 1 \\ |\alpha|+|\beta| \leqslant \lfloor p \rfloor}} \overline{H}_{(\alpha,\beta);s}X_{su}^{\alpha} X_{ut}^{\beta} - \sum_{|\varepsilon|=1}^{\lfloor p \rfloor - 1} \sum_{|\delta|=1}^{\lfloor p \rfloor - |\varepsilon|} \overline{H}_{(\delta,\varepsilon);s}X_{su}^{\delta} X_{ut}^{\varepsilon}$$

$$= 0. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$$

Definition 2.7 (Rough integral). Let $X \in \mathscr{C}_{\omega}^{p}([0,T],V)$, $\overline{H} \in \mathscr{D}_X(\mathcal{L}(V,W))$ be as above. We define

$$\int_0^{\cdot} \overline{H} \mathrm{d}X : [0,T] \to W \tag{38}$$

to be the unique path $I \in C_{\omega}([0,T],W)$ with the property that $I_{st} \approx \Xi_{st}$, which exists by [17, Theorem 3.3.1], and is given by

$$I_{st} = \lim_{n \to \infty} \sum_{[u,v] \in \pi_n} \Xi_{uv}$$

for a sequence of partitions $(\pi_n)_n$ with vanishing step size. We can make $\int \overline{H} \mathrm{d}X$ into an X-controlled path by defining, for $1 \leqslant n \leqslant \lfloor p \rfloor - 1$

$$\left(\overline{\int \overline{H} \mathrm{d}X} \right)_n := \overline{H}_{n-1} \in \mathcal{L}(V^{\otimes n-1}, \mathcal{L}(V,W)) = \mathcal{L}(V^{\otimes n}, W).$$

Note the presence of the bar above the integral sign, which emphasises membership to $\mathscr{D}_X(W)$.

An X-controlled path can be made into a rough path in its own right. We use (17) as a blueprint for the following definition, where we truncate at the correct order to avoid infinite sums.

Definition 2.8 (Lift of a controlled path). Let $X \in \mathscr{C}_{\omega}^{p}([0,T],V)$, $\overline{H} \in \mathscr{D}_X(U)$. Define $1_X\overline{H} : \Delta_T \to T^{\lfloor p \rfloor}(U)$ (notice the partial arrow notation to indicate almost multiplicativity and geometricity) by

$$(1_X\overline{H})_{st}^0 := 1, \quad (1_X\overline{H})_{st}^1 := H_{st} \tag{39}$$

and for $2 \leqslant m \leqslant \lfloor p \rfloor$

$$
\begin{aligned}
(1_X \overline{H})_{st}^m &:= \langle \overline{H}_s^{\boxtimes m}, \widetilde{\Delta}_{\sqcup}^m X_{st} \rangle \\
&= \sum_{\substack{n_1, \ldots, n_m \geqslant 1 \\ n := n_1 + \cdots + n_m \leqslant \lfloor p \rfloor}} \langle \overline{H}_{n_1;s} \boxtimes \cdots \boxtimes \overline{H}_{n_m;s}, (\widetilde{\Delta}_{\sqcup} X)_{st}^{n_1, \ldots, n_m} \rangle.
\end{aligned}
\tag{40}
$$

As it is shown in Theorem 2.9 below, [17, Theorem 3.3.1] applies to this functional, and given any sequence of partitions $(\pi_n)_n$ with vanishing step size

$$
(\uparrow_X \overline{H})_{st} := \lim_{n \to \infty} \left(\bigotimes_{[u,v] \in \pi_n} (1_X \overline{H})_{uv} \right)
\tag{41}
$$

defines an element of $\mathscr{C}_\omega^p([0, T], U)$, which we call the *lift* of $\overline{H}$ to rough path with respect to X.

Note how it was necessary to distinguish the case $m = 1$ above: this is due to the fact that we do not have the $\lfloor p \rfloor^{\text{th}}$ Gubinelli derivative, and therefore (40) would only be accurate at order $\lfloor p \rfloor$ (though in all explicit cases presented here these are known, and (40) is applicable for $m = 1$ too; an example where the case distinction is essential would be Example 2.5 with F only $(\lfloor p \rfloor - 1)$-times differentiable). Equation (40) can be written dually as

$$
(1_X \overline{H})_{st}^m = \langle \overline{H}_{\geqslant 1;s}^{\sqcup m}, X_{st} \rangle = \sum_{\substack{n_1, \ldots, n_m \geqslant 1 \\ n := n_1 + \cdots + n_m \leqslant \lfloor p \rfloor}} \langle \overline{H}_s^{n_1} \sqcup \cdots \sqcup \overline{H}_s^{n_m}, X_{st}^n \rangle
\tag{42}
$$

and in coordinates as

$$
(1_X \overline{H})^{()} = 1, \quad (1_X \overline{H})^k = H^k
$$

$$
(1_X \overline{H})_{st}^{(k_1, \ldots, k_m)} = \sum_{\substack{|\gamma^1|, \ldots, |\gamma^m| \geqslant 1 \\ |\gamma^1| + \cdots + |\gamma^m| \leqslant \lfloor p \rfloor \\ \gamma \in \overline{\mathrm{Sh}}(\gamma^1, \ldots, \gamma^m)}} \overline{H}_{\gamma^1;s}^{k_1} \cdots \overline{H}_{\gamma^m;s}^{k_m} X_{st}^\gamma.
\tag{43}
$$

In explicit calculations we will use $1_X \overline{H}$, for which we have a combinatorial expression, as a proxy for the true lift $\uparrow_X \overline{H}$.

The next result is one of the main theorems in this article. It can be compared with [15, Theorem 4.6], which applies to the special case of integrals of Lip(γ) forms, covered in Example 2.11 below. Their proof makes use of the symmetry of Lip(γ) forms, while the lemma below does not require it. Since some of the indexing is quite complex, it will be helpful to denote

$$
\begin{aligned}
A_1 &:= \bigcup_{m=1}^{\infty} (I_*^{\bullet})^m, \\
A_1^p &:= \bigcup_{m=1}^{\infty} \left\{ (\gamma^1, \ldots, \gamma^m) \in (I_*^{\bullet})^m \mid |\gamma^1| + \cdots + |\gamma^m| \leqslant \lfloor p \rfloor \right\},
\end{aligned}
\tag{44}
$$

so, for example, the last sum in (43) can be written

$$\sum_{\substack{(\gamma^1,\dots,\gamma^m)\in A_1^p \\ \gamma\in\overline{\mathrm{Sh}}(\gamma^1,\dots,\gamma^m)}} \overline{H}^{k_1}_{\gamma^1;s}\cdots\overline{H}^{k_m}_{\gamma^m;s}\boldsymbol{X}^{\gamma}_{st},$$

where m is intended as fixed.

Theorem 2.9. $1_X\overline{H}$ *is an almost rough path. Therefore the limit taken in (41) exists and defines a* U-*valued* p-*weakly geometric rough path, controlled by ω on* $[0,T]$, *with trace H.*

Proof. We begin by showing almost multiplicativity, that is, that for $|\boldsymbol{k}| = 0, \dots, \lfloor p \rfloor$ and $0 \leqslant s \leqslant u \leqslant t \leqslant T$

$$\sum_{(\boldsymbol{i},\boldsymbol{j})=\boldsymbol{k}} (1_X\overline{H})^{\boldsymbol{i}}_{su}(1_X\overline{H})^{\boldsymbol{j}}_{ut} \approx (1_X\overline{H})^{\boldsymbol{k}}_{st}.$$

For $|\boldsymbol{k}| = 0$ this is trivial and for $|\boldsymbol{k}| = 1$ it coincides with the statement that H is a path. For $|\boldsymbol{k}| = 2$ (which presupposes $\lfloor p \rfloor \geqslant 2$) we have

$$\begin{aligned}
&\sum_{(\boldsymbol{i},\boldsymbol{j})=(k_1,k_2)} (1_X\overline{H})^{\boldsymbol{i}}_{su}(1_X\overline{H})^{\boldsymbol{j}}_{ut} \\
&= (1_X\overline{H})^{(k_1,k_2)}_{su} + (1_X\overline{H})^{k_1}_{su}(1_X\overline{H})^{k_2}_{ut} + (1_X\overline{H})^{(k_1,k_2)}_{ut} \\
&= \sum_{\substack{(\alpha^1,\alpha^2)\in A_1^p \\ \alpha\in\overline{\mathrm{Sh}}(\alpha^1,\alpha^2)}} (\overline{H}^{k_1}_{\alpha^1;s}\overline{H}^{k_2}_{\alpha^2;s}\boldsymbol{X}^{\alpha}_{su}) + H^{k_1}_{su}H^{k_2}_{ut} + \sum_{\substack{(\beta^1,\beta^2)\in A_1^p \\ \beta\in\overline{\mathrm{Sh}}(\beta^1,\beta^2)}} (\overline{H}^{k_1}_{\beta^1;u}\overline{H}^{k_2}_{\beta^2;u}\boldsymbol{X}^{\beta}_{ut}).
\end{aligned} \quad (45)$$

We continue the calculation by re-expanding all the $\overline{H}$ terms at s and using (28):

$$H^{k_1}_{su} \approx_{\lfloor p \rfloor} \sum_{|\alpha^1|=1}^{\lfloor p \rfloor-1} \overline{H}^{k_1}_{\alpha^1;s}\boldsymbol{X}^{\alpha^1}_{su}$$

$$H^{k_2}_{ut} \approx_{\lfloor p \rfloor} \sum_{|\beta|=1}^{\lfloor p \rfloor-1} \overline{H}_{\beta;u}\boldsymbol{X}^{\beta}_{ut} \approx_{\lfloor p \rfloor} \sum_{\substack{|\beta|=1,\dots,\lfloor p \rfloor-1 \\ |\alpha^2|=0,\dots,\lfloor p \rfloor-1-|\beta|}} \overline{H}^{k_2}_{(\alpha^2,\beta);s}\boldsymbol{X}^{\alpha^2}_{su}\boldsymbol{X}^{\beta}_{ut}.$$

These two identities, the fact that $H_{su}, H_{ut} \approx_1 0$ and (29) imply

$$H^{k_1}_{su}H^{k_2}_{ut} \approx \sum_{\substack{|\beta|=1,\dots,\lfloor p \rfloor-1 \\ |\alpha^1|=1,\dots,\lfloor p \rfloor-1 \\ |\alpha^2|=0,\dots,\lfloor p \rfloor-1-|\beta| \\ \alpha\in\mathrm{Sh}(\alpha^1,\alpha^2)}} \overline{H}^{k_1}_{\alpha^1;s}\overline{H}^{k_2}_{(\alpha^2,\beta);s}\boldsymbol{X}^{\alpha}_{su}\boldsymbol{X}^{\beta}_{ut}.$$

Similarly

$$\overline{H}^{k_1}_{\boldsymbol{\beta}^1;u} \overline{H}^{k_2}_{\boldsymbol{\beta}^2;u} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut} \approx \sum_{\substack{|\boldsymbol{\alpha}^1|=0,\dots,\lfloor p \rfloor - 1 - |\boldsymbol{\beta}^1| \\ |\boldsymbol{\alpha}^2|=0,\dots,\lfloor p \rfloor - 1 - |\boldsymbol{\beta}^2| \\ \boldsymbol{\alpha} \in \mathrm{Sh}(\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2)}} \overline{H}^{k_1}_{(\boldsymbol{\alpha}^1,\boldsymbol{\beta}^1);s} \overline{H}^{k_2}_{(\boldsymbol{\alpha}^2,\boldsymbol{\beta}^2);s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}.$$

Incorporating these computations in (45) we obtain

$$\sum_{(\boldsymbol{i},\boldsymbol{j})=(k_1,k_2)} (\mathbf{1}_{\boldsymbol{X}} \, \overline{H})^{\boldsymbol{i}}_{su} (\mathbf{1}_{\boldsymbol{X}} \, \overline{H})^{\boldsymbol{j}}_{ut}$$

$$\approx \sum_{\substack{(\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2)\in A^p_1 \\ \boldsymbol{\alpha}\in\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2)}} (\overline{H}^{k_1}_{\boldsymbol{\alpha}^1;s} \overline{H}^{k_2}_{\boldsymbol{\alpha}^2;s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su}) + \sum_{\substack{|\boldsymbol{\beta}|=1,\dots,\lfloor p \rfloor - 1 \\ |\boldsymbol{\alpha}^1|=1,\dots,\lfloor p \rfloor - 1 \\ |\boldsymbol{\alpha}^2|=1,\dots,\lfloor p \rfloor - 1 - |\boldsymbol{\beta}| \\ \boldsymbol{\alpha}\in\mathrm{Sh}(\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2)}} (\overline{H}^{k_1}_{\boldsymbol{\alpha}^1;s} \overline{H}^{k_2}_{(\boldsymbol{\alpha}^2,\boldsymbol{\beta});s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut})$$

$$+ \sum_{\substack{(\boldsymbol{\beta}^1,\boldsymbol{\beta}^2)\in A^p_1 \\ |\boldsymbol{\alpha}^1|=0,\dots,\lfloor p \rfloor - 1 - |\boldsymbol{\beta}^1| \\ |\boldsymbol{\alpha}^2|=0,\dots,\lfloor p \rfloor - 1 - |\boldsymbol{\beta}^2| \\ \boldsymbol{\alpha}\in\mathrm{Sh}(\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2) \\ \boldsymbol{\beta}\in\overline{\mathrm{Sh}}(\boldsymbol{\beta}^1,\boldsymbol{\beta}^2)}} (\overline{H}^{k_1}_{(\boldsymbol{\alpha}^1,\boldsymbol{\beta}^1);s} \overline{H}^{k_2}_{(\boldsymbol{\alpha}^2,\boldsymbol{\beta}^2);s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut})$$

$$= \sum_{(\boldsymbol{\gamma}^1,\boldsymbol{\gamma}^2)\in A^p_1} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \overline{H}^{k_2}_{\boldsymbol{\gamma}^2;s} \sum_{\substack{l=0,1,2 \\ \boldsymbol{\alpha}^h=\boldsymbol{\gamma}^h,\ h\leqslant l \\ (\boldsymbol{\alpha}^h,\boldsymbol{\beta}^h)=\boldsymbol{\gamma}^h,\ |\boldsymbol{\beta}^h|\geqslant 1,\ h\geqslant l+1 \\ \boldsymbol{\alpha}\in\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^l),\mathrm{Sh}(\boldsymbol{\alpha}^{l+1},\dots,\boldsymbol{\alpha}^2)) \\ \boldsymbol{\beta}\in\overline{\mathrm{Sh}}(\boldsymbol{\beta}^{l+1},\dots,\boldsymbol{\beta}^2)}} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}$$

$$= \sum_{(\boldsymbol{\gamma}^1,\boldsymbol{\gamma}^2)\in A^p_1} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \overline{H}^{k_2}_{\boldsymbol{\gamma}^2;s} \sum_{\substack{\boldsymbol{\gamma}\in\overline{\mathrm{Sh}}(\boldsymbol{\gamma}^1,\boldsymbol{\gamma}^2) \\ (\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\gamma}}} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}$$

$$= (\mathbf{1}_{\boldsymbol{X}} \, \overline{H})^{(k_1,k_2)}_{st},$$

where we have used Lemma 1.1 in the second last identity and multiplicativity of $\boldsymbol{X}$ in the last. The case of $m := |\boldsymbol{k}| \geqslant 3$ (which presupposes $\lfloor p \rfloor \geqslant 3$) is handled similarly, but has to be distinguished from the previous case since the middle terms are not the same.

$$\sum_{(\boldsymbol{i},\boldsymbol{j})=\boldsymbol{k}} (\mathbf{1}_{\boldsymbol{X}} \, \overline{H})^{\boldsymbol{i}}_{su} (\mathbf{1}_{\boldsymbol{X}} \, \overline{H})^{\boldsymbol{j}}_{ut}$$

$$= \sum_{\substack{(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^m)\in A^p_1 \\ \boldsymbol{\alpha}\in\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^m)}} \overline{H}^{k_1}_{\boldsymbol{\alpha}^1;s} \cdots \overline{H}^{k_m}_{\boldsymbol{\alpha}^m;s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} + \left(\sum_{\substack{(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^{m-1})\in A^p_1 \\ \boldsymbol{\alpha}\in\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^{m-1})}} \overline{H}^{k_1}_{\boldsymbol{\alpha}^1;s} \cdots \overline{H}^{k_{m-1}}_{\boldsymbol{\alpha}^{m-1};s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \right) H^{k_m}_{ut}$$

$$+ \sum_{l=2}^{m-2} \left(\sum_{\substack{(\alpha^1,\dots,\alpha^l)\in A_1^p \\ \alpha\in\overline{\mathrm{Sh}}(\alpha^1,\dots,\alpha^l)}} \overline{H}_{\alpha^1;s}^{k_1} \cdots \overline{H}_{\alpha^l;s}^{k_l} \boldsymbol{X}_{su}^{\alpha} \right) \left(\sum_{\substack{(\beta^{l+1},\dots,\beta^m)\in A_1^p \\ \beta\in\overline{\mathrm{Sh}}(\beta^{l+1},\dots,\beta^m)}} \overline{H}_{\beta^{l+1};u}^{k_{l+1}} \cdots \overline{H}_{\beta^m;u}^{k_m} \boldsymbol{X}_{ut}^{\beta} \right)$$

$$+ H_{su}^{k_1} \left(\sum_{\substack{(\beta^2,\dots,\beta^m)\in A_1^p \\ \beta\in\overline{\mathrm{Sh}}(\beta^2,\dots,\beta^m)}} \overline{H}_{\beta^2;u}^{k_2} \cdots \overline{H}_{\beta^m;u}^{k_m} \boldsymbol{X}_{ut}^{\beta} \right) + \sum_{\substack{(\beta^1,\dots,\beta^m)\in A_1^p \\ \beta\in\overline{\mathrm{Sh}}(\beta^1,\dots,\beta^m)}} \overline{H}_{\beta^1;u}^{k_1} \cdots \overline{H}_{\beta^m;u}^{k_m} \boldsymbol{X}_{ut}^{\beta}$$

$$\approx \sum_{\substack{(\alpha^1,\dots,\alpha^m)\in A_1^p \\ \alpha\in\overline{\mathrm{Sh}}(\alpha^1,\dots,\alpha^m)}} \overline{H}_{\alpha^1;s}^{k_1} \cdots \overline{H}_{\alpha^m;s}^{k_m} \boldsymbol{X}_{su}^{\alpha}$$

$$+ \left(\sum_{\substack{(\alpha^1,\dots,\alpha^{m-1})\in A_1^p \\ \alpha\in\overline{\mathrm{Sh}}(\alpha^1,\dots,\alpha^{m-1})}} \overline{H}_{\alpha^1;s}^{k_1} \cdots \overline{H}_{\alpha^{m-1};s}^{k_{m-1}} \boldsymbol{X}_{su}^{\alpha} \right) \left(\sum_{\substack{|\beta^m|=1,\dots,\lfloor p\rfloor-1 \\ |\alpha^m|=0,\dots,\lfloor p\rfloor-1-|\beta|}} \overline{H}_{(\alpha^m,\beta^m);s}^{k_m} \boldsymbol{X}_{su}^{\alpha^m} \boldsymbol{X}_{ut}^{\beta^m} \right)$$

$$+ \sum_{l=2}^{m-2} \left[\left(\sum_{\substack{(\alpha^1,\dots,\alpha^l)\in A_1^p \\ \alpha\in\overline{\mathrm{Sh}}(\alpha^1,\dots,\alpha^l)}} \overline{H}_{\alpha^1;s}^{k_1} \cdots \overline{H}_{\alpha^l;s}^{k_l} \boldsymbol{X}_{su}^{\alpha} \right) \right.$$

$$\left. \cdot \left(\sum_{\substack{(\beta^{l+1},\dots,\beta^m)\in A_1^p \\ |\alpha^h|=0,\dots,\lfloor p\rfloor-1-|\beta^h|,\ h\geqslant l+1 \\ \alpha\in\mathrm{Sh}(\alpha^{l+1},\dots,\alpha^m) \\ \beta\in\overline{\mathrm{Sh}}(\beta^{l+1},\dots,\beta^m)}} \overline{H}_{(\alpha^{l+1},\beta^{l+1});s}^{k_{l+1}} \cdots \overline{H}_{(\alpha^m,\beta^m);s}^{k_m} \boldsymbol{X}_{su}^{\alpha} \boldsymbol{X}_{ut}^{\beta} \right) \right]$$

$$+ \left(\sum_{|\alpha^1|=1}^{\lfloor p\rfloor-1} \overline{H}_{\alpha^1;s}^{k_1} \boldsymbol{X}_{su}^{\alpha^1} \right) \left(\sum_{\substack{(\beta^2,\dots,\beta^m)\in A_1^p \\ |\alpha^h|=0,\dots,\lfloor p\rfloor-1-|\beta^h|,\ h\geqslant 2 \\ \alpha\in\mathrm{Sh}(\alpha^2,\dots,\alpha^m) \\ \beta\in\overline{\mathrm{Sh}}(\beta^2,\dots,\beta^m)}} \overline{H}_{(\alpha^2,\beta^2);s}^{k_2} \cdots \overline{H}_{(\alpha^m,\beta^m);s}^{k_m} \boldsymbol{X}_{su}^{\alpha} \boldsymbol{X}_{ut}^{\beta} \right)$$

$$
+ \sum_{\substack{(\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m) \in A_1^p \\ |\boldsymbol{\alpha}^h|=0,\dots,\lfloor p \rfloor -1-|\boldsymbol{\beta}^h| \\ \boldsymbol{\alpha} \in \mathrm{Sh}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^m) \\ \boldsymbol{\beta} \in \overline{\mathrm{Sh}}(\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^m)}} \overline{H}^{k_1}_{(\boldsymbol{\alpha}^1,\boldsymbol{\beta}^1);s} \cdots \overline{H}^{k_m}_{(\boldsymbol{\alpha}^m,\boldsymbol{\beta}^m);s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}
$$

$$
= \sum_{(\boldsymbol{\gamma}^1,\dots,\boldsymbol{\gamma}^m)\in A_1^p} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_m}_{\boldsymbol{\gamma}^m;s} \sum_{\substack{l=0,\dots,m \\ \boldsymbol{\alpha}^h = \boldsymbol{\gamma}^h,\ h \leqslant l \\ (\boldsymbol{\alpha}^h,\boldsymbol{\beta}^h)=\boldsymbol{\gamma}^h,\ |\boldsymbol{\beta}^h|\geqslant 1,\ h \geqslant l+1 \\ \boldsymbol{\alpha}\in\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^l),\mathrm{Sh}(\boldsymbol{\alpha}^{l+1},\dots,\boldsymbol{\alpha}^m)) \\ \boldsymbol{\beta}\in\overline{\mathrm{Sh}}(\boldsymbol{\beta}^{l+1},\dots,\boldsymbol{\beta}^m)}} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}
$$

$$
= \sum_{(\boldsymbol{\gamma}^1,\dots,\boldsymbol{\gamma}^m)\in A_1^p} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_m}_{\boldsymbol{\gamma}^m;s} \sum_{\substack{\boldsymbol{\gamma}\in\overline{\mathrm{Sh}}(\boldsymbol{\gamma}^1,\dots,\boldsymbol{\gamma}^m) \\ (\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\gamma}}} \boldsymbol{X}^{\boldsymbol{\alpha}}_{su} \boldsymbol{X}^{\boldsymbol{\beta}}_{ut}
$$

$$
= (\mathbb{1}_X \overline{H})^k_{st}.
$$

We proceed with the proof of geometricity. Again using (29) we have

$$
(\mathbb{1}_X \overline{H})^i_{st} (\mathbb{1}_X \overline{H})^j_{st}
$$

$$
\approx \sum_{\substack{|\boldsymbol{\alpha}^1|,\dots,|\boldsymbol{\alpha}^m|,|\boldsymbol{\beta}^1|,\dots,|\boldsymbol{\beta}^n|\geqslant 1 \\ |\boldsymbol{\alpha}^1|+\cdots+|\boldsymbol{\alpha}^m|+|\boldsymbol{\beta}^1|+\cdots+|\boldsymbol{\beta}^n|\leqslant\lfloor p\rfloor \\ \boldsymbol{\alpha}\in\mathrm{Sh}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^m) \\ \boldsymbol{\beta}\in\mathrm{Sh}(\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^n)}} \overline{H}^{i_1}_{\boldsymbol{\alpha}^1;s} \cdots \overline{H}^{i_m}_{\boldsymbol{\alpha}^m;s} \overline{H}^{j_1}_{\boldsymbol{\beta}^1;s} \cdots \overline{H}^{j_n}_{\boldsymbol{\beta}^n;s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{st} \boldsymbol{X}^{\boldsymbol{\beta}}_{st}
$$

$$
= \sum_{\substack{|\boldsymbol{\alpha}^1|,\dots,|\boldsymbol{\alpha}^m|,|\boldsymbol{\beta}^1|,\dots,|\boldsymbol{\beta}^n|\geqslant 1 \\ |\boldsymbol{\alpha}^1|+\cdots+|\boldsymbol{\alpha}^m|+|\boldsymbol{\beta}^1|+\cdots+|\boldsymbol{\beta}^n|\leqslant\lfloor p\rfloor}} \overline{H}^{i_1}_{\boldsymbol{\alpha}^1;s} \cdots \overline{H}^{i_m}_{\boldsymbol{\alpha}^m;s} \overline{H}^{j_1}_{\boldsymbol{\beta}^1;s} \cdots \overline{H}^{j_n}_{\boldsymbol{\beta}^n;s}
$$

$$
\cdot \sum_{\boldsymbol{\gamma}\in\mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^m),\overline{\mathrm{Sh}}(\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^n))} \boldsymbol{X}^{\boldsymbol{\gamma}}_{st}
$$

$$
= \sum_{\boldsymbol{k}\in\mathrm{Sh}(\boldsymbol{i},\boldsymbol{j})} \sum_{(\boldsymbol{\gamma}^1,\dots,\boldsymbol{\gamma}^{m+n})\in A_1^p} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_{m+n}}_{\boldsymbol{\gamma}^{m+n};s} \sum_{\boldsymbol{\gamma}\in\overline{\mathrm{Sh}}(\boldsymbol{\gamma}^1,\dots,\boldsymbol{\gamma}^{m+n})} \boldsymbol{X}^{\boldsymbol{\gamma}}_{st}
$$

$$
= \sum_{\boldsymbol{k}\in\mathrm{Sh}(\boldsymbol{i},\boldsymbol{j})} (\mathbb{1}_X \overline{H})^k_{st},
$$

where we have used Lemma 1.4 (with $m = 2$) in the second last identity. We may therefore apply Proposition 2.3 to conclude the proof. $\qquad\square$

This construction immediately yields a couple of important examples of rough path.

Example 2.10 (Pushforward of rough paths). Let $\boldsymbol{X}, F$ be as in Example 2.5. We denote

$$
F_* \boldsymbol{X} := \uparrow_X \overline{F}(X) \tag{46}
$$

and call it the *pushforward* of $\boldsymbol{X}$ through F. This is a rough path with trace $F(X)$.

Example 2.11 (Rough integrals as rough paths). Let $X, \overline{H}$ be as in Definition 2.7. We denote $\int \overline{H} \mathrm{d}X := \uparrow_X \overline{\int \overline{H} \mathrm{d}X}$.

We can use these notions to reinterpret the following well-known fact about weakly geometric rough paths. Notice how, in particular, this implies that the rough integral of an exact 1-form is entirely determined by its trace. (Incidentally, arbitrary 1-forms do not require the whole rough path for the integral to be defined either, just the terms $(\odot_{n-1} \otimes \mathbb{1})X^n$ for $n = 1, \dots, \lfloor p \rfloor$.)

Proposition 2.12 (Change of variable formula). *Let X be as above, $F \in C^\infty(V, W)$, then the following identity*

$$\overline{F(X)} = F(X_0) + \overline{\int \overline{DF}(X) \mathrm{d}X} \tag{47}$$

holds in $\mathscr{D}_X(W)$. Therefore, the corresponding identity in $\mathscr{C}_\omega^p([0,T], W)$ holds as well:

$$F_*X = F(X_0) + \int \overline{DF}(X) \mathrm{d}X, \tag{48}$$

where in both cases the constant $F(X_0)$ is only added to the trace of the integral.

Proof. For the trace we have, by Taylor's formula and Proposition 2.2

$$F(X)_{st} \approx \sum_{n=1}^{\lfloor p \rfloor} \frac{1}{n!} \langle D^n F(X_s), X_{st}^{\otimes n} \rangle$$

$$= \sum_{n=1}^{\lfloor p \rfloor} \langle D^n F(X_s), \mathbf{X}_{st}^n \rangle$$

$$\approx \int_s^t \overline{DF}(X) \mathrm{d}X$$

and therefore [17, Theorem 3.3.1] implies $F(X)_{0t} = \int_0^t \overline{DF}(X) \mathrm{d}X$. The other claims follow trivially. $\square$

Remark 2.13. A similar formula would hold for more general controlled paths, that is,

$$\overline{H} = H_0 + \int \overline{H'} \mathrm{d}X \tag{49}$$

(together with its rough path counterpart, given by passing to the $\uparrow_X$ on both sides), provided that we have a $\lfloor p \rfloor$th Gubinelli derivative, needed to define the controlled integrand $\overline{H}'$.

We would now like to show that a path controlled by the lift of a controlled path is controlled by the original rough path in a canonical fashion.

Definition 2.14 (Change of controlling rough path). Let $X \in \mathscr{C}_\omega^p([0,T], V)$, $\overline{H} \in \mathscr{D}_X(U)$, S another vector space, $H := \uparrow_X \overline{H}$ and $\overline{K} \in \mathscr{D}_H(S)$. We then define

$$(\overline{K} * \overline{H})_n := \sum_{m=1}^n \overline{K}_m \circ \overline{H}^{\boxtimes m} \circ \widetilde{\Delta}_{\sqcup}^m|_{V^{\otimes n}}, \tag{50}$$

which for $n = 0$ reduces to $(\overline{K} * \overline{H})_0 := K$. In coordinates this means

$$(\overline{K} * \overline{H})_\gamma^c := \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1,\ldots,\gamma^m) \in \overline{\mathrm{Sh}}^{-1}(\gamma) \\ |\gamma^1|,\ldots,|\gamma^m| \geqslant 1}} \overline{K}_k^c \overline{H}_{\gamma^1}^{k_1} \cdots \overline{H}_{\gamma^m}^{k_m}, \quad |\gamma| \geqslant 1 \tag{51}$$

and $(\overline{K} * \overline{H})_0^c = K^c$.

In Proposition 2.16 below we show that this defines a controlled path. The next example features a case in which the reduced ordered shuffle coproduct can be replaced with its unordered counterpart; this, however, is not the general case.

Example 2.15. Let $X, \overline{H}, H, \overline{K}$ be as above, with $\overline{K} := \overline{F}(H)$ for $F \in C^\infty(U, S)$. Then, since $\overline{K}_m$ is symmetric, we may rewrite (50) by using the unordered shuffle coproduct: by Theorem 1.6

$$(\overline{F}(H) * \overline{H})_n = \sum_{m \geqslant 1} \frac{1}{m!} D^m F(H) \circ \overline{H}^{\boxtimes m} \circ \widetilde{\Delta}_{\sqcup}^m|_{V^{\otimes n}}, \quad n \geqslant 1 \tag{52}$$

or in coordinates

$$(\overline{F}(H) * \overline{H})_\gamma^c := \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1,\ldots,\gamma^m) \in \mathrm{Sh}^{-1}(\gamma) \\ |\gamma^1|,\ldots,|\gamma^m| \geqslant 1}} \frac{1}{m!} \partial_k F^c(H) \overline{H}_{\gamma^1}^{k_1} \cdots \overline{H}_{\gamma^m}^{k_m}, \quad |\gamma| \geqslant 1. \tag{53}$$

When $\overline{H}$ is also given by a smooth function, this is known as the Faà di Bruno formula for the higher derivatives of a composition of functions

$$\partial_\gamma (F \circ G)^c(X) := \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1,\ldots,\gamma^m) \in \mathrm{Sh}^{-1}(\gamma) \\ |\gamma^1|,\ldots,|\gamma^m| \geqslant 1}} \frac{1}{m!} \partial_k F^c(H) \partial_{\gamma^1} G^{k_1} \cdots \partial_{\gamma^m} G^{k_m}(X). \tag{54}$$

We will denote the X-controlled path $\overline{F}(H) * \overline{H} := F_* \overline{H} \in \mathscr{D}_X(S)$ and call it the *pushforward* of $\overline{H}$ through F. Note how this is distinct from $\overline{F}(H) \in \mathscr{D}_H(S)$, and note how $F_* \overline{X} = \overline{F}(X)$ where $\overline{X}$ denotes the controlled path X with zero Gubinelli derivatives. An easy application of Proposition 2.12 shows the following change of variable formula for controlled paths:

$$F_* \overline{H} = \left(F(H_0) + \overline{\int DF(H) \mathrm{d}(\uparrow_X \overline{H})} \right) * \overline{H}. \tag{55}$$

Proposition 2.16. *The map* $\overline{K} * \overline{H} \in \mathcal{L}(T^{\lfloor p \rfloor - 1}(V), S)$ *of Definition 2.14 is an element of* $\mathscr{D}_X(S)$.

Proof. Clearly $\overline{K} * \overline{H} \in C_\omega^p([0, T], \mathcal{L}(T^{\lfloor p \rfloor - 1}(V), U))$, so it remains to show (31). We preliminarily write

$$
\begin{aligned}
\overline{K}_{\boldsymbol{j};st}^c &\approx_{\lfloor p \rfloor - |\boldsymbol{j}|} \sum_{|\boldsymbol{i}|=1}^{\lfloor p \rfloor - |\boldsymbol{j}| - 1} \overline{K}_{(\boldsymbol{i},\boldsymbol{j});s}^c \boldsymbol{H}_{st}^{\boldsymbol{i}} \\
&\approx \sum_{\substack{m=1,\ldots,\lfloor p \rfloor - |\boldsymbol{j}| - 1 \\ (\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^m) \in A_1^p \\ \boldsymbol{\alpha} \in \overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^m)}} \overline{K}_{(\boldsymbol{i},\boldsymbol{j});s}^c \overline{H}_{\boldsymbol{\alpha}^1;s}^{i_1} \cdots \overline{H}_{\boldsymbol{\alpha}^m;s}^{i_m} \boldsymbol{X}_{st}^{\boldsymbol{\alpha}}
\end{aligned}
\tag{56}
$$

and

$$
\overline{H}_{\boldsymbol{\delta};st}^j \approx_{\lfloor p \rfloor - |\boldsymbol{\delta}|} \sum_{|\boldsymbol{\alpha}|=1}^{\lfloor p \rfloor - |\boldsymbol{\delta}| - 1} \overline{H}_{(\boldsymbol{\alpha},\boldsymbol{\delta});s}^j \boldsymbol{X}_{st}^{\boldsymbol{\alpha}}.
\tag{57}
$$

Now, let $1 \leqslant |\boldsymbol{\beta}| \leqslant \lfloor p \rfloor - 2$:

$$
\begin{aligned}
&(\overline{K} * \overline{H})_{\boldsymbol{\beta};t}^c \\
&= \sum_{\substack{n=1,\ldots,|\boldsymbol{\beta}| \\ (\boldsymbol{\beta}^1,\ldots,\boldsymbol{\beta}^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta})}} \overline{K}_{\boldsymbol{j};t}^c \overline{H}_{\boldsymbol{\beta}^1;t}^{j_1} \cdots \overline{H}_{\boldsymbol{\beta}^n;t}^{j_n} \\
&= \sum_{\substack{n=1,\ldots,|\boldsymbol{\beta}| \\ (\boldsymbol{\beta}^1,\ldots,\boldsymbol{\beta}^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta})}} (\overline{K}_{\boldsymbol{j};s}^c + \overline{K}_{\boldsymbol{j};st}^c)(\overline{H}_{\boldsymbol{\beta}^1;s}^{j_1} + \overline{H}_{\boldsymbol{\beta}^1;st}^{j_1}) \cdots (\overline{H}_{\boldsymbol{\beta}^n;s}^{j_n} + \overline{H}_{\boldsymbol{\beta}^n;st}^{j_n}) \\
&\approx_{\lfloor p \rfloor - |\boldsymbol{\beta}|} \sum_{\substack{n=1,\ldots,|\boldsymbol{\beta}| \\ (\boldsymbol{\beta}^1,\ldots,\boldsymbol{\beta}^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta})}} \left(\overline{K}_{\boldsymbol{j};s}^c \overline{H}_{\boldsymbol{\beta}^1;s}^{j_1} \cdots \overline{H}_{\boldsymbol{\beta}^n;s}^{j_n} + \sum_{\substack{\epsilon_0,\ldots,\epsilon_n=0,1 \\ (\epsilon_0,\ldots,\epsilon_n) \neq (0,\ldots,0)}} \xi_{\epsilon_0} \eta_{\epsilon_1}^1 \cdots \eta_{\epsilon_n}^n \right),
\end{aligned}
$$

where $\xi_0 := \overline{K}_{\boldsymbol{j};s}^c$, $\eta_0^l := \overline{H}_{\boldsymbol{\beta}^l;s}^{j_l}$, and ξ_1, η_1^l are given by the right-hand sides of (56) and (57) (with $\boldsymbol{\delta} := \boldsymbol{\beta}^l$), respectively; the $\approx_{\lfloor p \rfloor - |\boldsymbol{\beta}|}$ in the last line above holds since $|\boldsymbol{j}|, |\boldsymbol{\beta}^l| \leqslant |\boldsymbol{\beta}|$ and the functions of ϵ_l are not all zero. We can rewrite this as

$$
\begin{aligned}
&(\overline{K} * \overline{H})_{\boldsymbol{\beta};t}^c \\
&\approx_{\lfloor p \rfloor - |\boldsymbol{\beta}|} \sum_{\substack{n=1,\ldots,|\boldsymbol{\beta}| \\ (\boldsymbol{\beta}^1,\ldots,\boldsymbol{\beta}^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta}) \\ m=0,\ldots,\lfloor p \rfloor - n - 1 \\ (\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^m) \in A_1^p \\ \boldsymbol{\delta} \in \overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^m) \\ \alpha^{m+l}=0,\ldots,\lfloor p \rfloor - |\boldsymbol{\beta}^l| - 1}} (\overline{K}_{(\boldsymbol{i},\boldsymbol{j});s}^c \overline{H}_{\boldsymbol{\alpha}^1;s}^{i_1} \cdots \overline{H}_{\boldsymbol{\alpha}^m;s}^{i_m} \boldsymbol{X}_{st}^{\boldsymbol{\delta}}).
\end{aligned}
$$

$$\cdot (\overline{H}^{j_1}_{(\boldsymbol{\alpha}^{m+1},\boldsymbol{\beta}^1);s} \boldsymbol{X}^{\boldsymbol{\alpha}^{m+1}}_{st}) \cdots (\overline{H}^{j_n}_{(\boldsymbol{\alpha}^{m+n},\boldsymbol{\beta}^n);s} \boldsymbol{X}^{\boldsymbol{\alpha}^{m+n}}_{st})$$

$$\approx_{\lfloor p \rfloor - |\boldsymbol{\beta}|} \sum_{\substack{q=1,\ldots,\lfloor p \rfloor -1 \\ n=1,\ldots,|\boldsymbol{\beta}| \\ |\boldsymbol{\alpha}^1|,\ldots,|\boldsymbol{\alpha}^{q-n}| \geqslant 1 \\ |\boldsymbol{\alpha}^1|+\cdots+|\boldsymbol{\alpha}^q| \leqslant \lfloor p \rfloor - |\boldsymbol{\beta}|-1 \\ \boldsymbol{\alpha} \in \mathrm{Sh}(\overline{\mathrm{Sh}}(\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^{q-n}),\mathrm{Sh}(\boldsymbol{\alpha}^{q-n+1},\ldots,\boldsymbol{\alpha}^q)) \\ (\boldsymbol{\beta}^1,\ldots,\boldsymbol{\beta}^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta}) \\ \boldsymbol{\gamma}^l = \boldsymbol{\alpha}^l, \; l \leqslant q-n; \; \boldsymbol{\gamma}^l = (\boldsymbol{\alpha}^l,\boldsymbol{\beta}^l), \; l \geqslant q-n+1}} \overline{K}^c_{\boldsymbol{k};s} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_q}_{\boldsymbol{\gamma}^q;s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{st}$$

$$= \sum_{|\boldsymbol{\alpha}|=0}^{\lfloor p \rfloor - |\boldsymbol{\beta}|-1} \overline{K}^c_{\boldsymbol{k};s} \left(\sum_{\substack{q=1,\ldots,|\boldsymbol{\alpha}|+|\boldsymbol{\beta}| \\ l=0,\ldots,q \\ ((\boldsymbol{\alpha}^1,\ldots,\boldsymbol{\alpha}^l),(\boldsymbol{\alpha}^{l+1},\ldots,\boldsymbol{\alpha}^q)) \in (\overline{\mathrm{Sh}}^{-1},\mathrm{Sh}^{-1})(\mathrm{Sh}^{-1}(\boldsymbol{\alpha})) \\ (\boldsymbol{\beta}^{l+1},\ldots,\boldsymbol{\beta}^q) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta}) \\ \boldsymbol{\gamma}^h = \boldsymbol{\alpha}^h, \; |\boldsymbol{\alpha}^h| \geqslant 1, \; h \leqslant l; \; \boldsymbol{\gamma}^h = (\boldsymbol{\alpha}^h,\boldsymbol{\beta}^h), \; |\boldsymbol{\beta}^h| \geqslant 1, \; h \geqslant l+1}} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_q}_{\boldsymbol{\gamma}^q;s} \right) \boldsymbol{X}^{\boldsymbol{\alpha}}_{st}$$

$$= \sum_{|\boldsymbol{\alpha}|=0}^{\lfloor p \rfloor - |\boldsymbol{\beta}|-1} \left(\sum_{(\boldsymbol{\gamma}^1,\ldots,\boldsymbol{\gamma}^q) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\alpha},\boldsymbol{\beta})} \overline{K}^c_{\boldsymbol{k};s} \overline{H}^{k_1}_{\boldsymbol{\gamma}^1;s} \cdots \overline{H}^{k_q}_{\boldsymbol{\gamma}^q;s} \right) \boldsymbol{X}^{\boldsymbol{\alpha}}_{st}$$

$$= \sum_{|\boldsymbol{\alpha}|=0}^{\lfloor p \rfloor - |\boldsymbol{\beta}|-1} (\overline{K} * \overline{H})^c_{(\boldsymbol{\alpha},\boldsymbol{\beta});s} \boldsymbol{X}^{\boldsymbol{\alpha}}_{st}$$

as needed. The second-last identity above is given by Corollary 1.3. The proof of (31) for $(\overline{K} * \overline{H})_0 = K$ is a much simplified version of the proof above. $\qquad\square$

Another application of the change of controlling path construction is a Leibniz rule for controlled paths.

Definition 2.17 (Leibniz rule for controlled paths). Let $\boldsymbol{X} \in \mathscr{C}^p_\omega([0,T],V)$, U_i be vector spaces for $i = 1,2,3$, $\overline{H} \in \mathscr{D}_{\boldsymbol{X}}(\mathcal{L}(U_1,U_2))$, $\overline{K} \in \mathscr{D}_{\boldsymbol{X}}(\mathcal{L}(U_2,U_3))$. Define $(\overline{K} \cdot \overline{H})_0 := K \circ H$ and for $n = 1,\ldots,\lfloor p \rfloor - 1$

$$(\overline{K} \cdot \overline{H})_n := \times \circ (\overline{K} \boxtimes \overline{H}) \circ \Delta_\sqcup, \tag{58}$$

where $\times : \mathcal{L}(U_2,U_3) \boxtimes \mathcal{L}(U_1,U_2) \to \mathcal{L}(U_1,U_3)$ is ordinary composition of linear maps.

In coordinates

$$(\overline{K} \cdot \overline{H})^{(i)}_{\boldsymbol{\gamma}} = \sum_{(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathrm{Sh}^{-1}(\boldsymbol{\gamma})} \overline{K}^{(i)}_{\boldsymbol{\alpha}\,(k)} \overline{H}^{(k)}_{\boldsymbol{\beta}\,(j)}. \tag{59}$$

Here we are using the notation $\binom{h}{l}$ to denote indices in spaces of linear maps, the upper index referring to the codomain and the lower to the domain; this allows us to use the Einstein convention on such superscripts. When the domain coincides with V (that is, it is a controlled integrand), we will often place it after the bottom tuple, for example, $\overline{H}^{(k)}_{\gamma} = \overline{H}^{k}_{(\gamma,\delta)}$, as done previously. The presence of the unreduced $\Delta_{\sqcup}$ in the previous definition might seem strange at first, but it is easily justified as follows.

Proposition 2.18. $\overline{K} \cdot \overline{H}$ *defines an element of* $\mathscr{D}_X(\mathcal{L}(U_1, U_3))$ *which coincides with* $\overline{\times}(K, H) *$ $(\overline{K}, \overline{H})$. *In particular, if* $\overline{H} = \overline{A}(X)$, $\overline{K} = \overline{B}(X)$ *for smooth functions* A, B, *then*

$$\overline{H} \cdot \overline{K} = \overline{A(\cdot)B(\cdot)}(X) \qquad (60)$$

with $A(\cdot)B(\cdot)$ *denoting the function* $x \mapsto A(x)B(x)$.

Proof. We apply Example 2.15 with the function $\times$ and the controlled path $(\overline{K}, \overline{H}) \in$ $\mathscr{D}_X(\mathcal{L}(U_2, U_3) \oplus \mathcal{L}(U_1, U_2))$. Denoting by $\binom{h}{\cdot}$ coordinates for the second direct summand and with $\binom{\cdot}{k}$ those for the first, we have

$$\times^{(i)}_{(j)}(\kappa, \eta) = \eta^{(i)}_{(l)} \kappa^{(l)}_{(j)}$$

$$\partial_{\binom{h}{p}} \times^{(i)}_{(j)}(\kappa, \eta) = \delta^{ih} \kappa^{(p)}_{(j)}, \quad \partial_{\binom{q}{k}} \times^{(i)}_{(j)}(\kappa, \eta) = \delta^{jk} \eta^{(i)}_{(q)}$$

$$\partial_{\binom{h}{p},\binom{p}{k}} \times^{(i)}_{(j)}(\kappa, \eta) = \delta^{jk} \delta^{hi} = \partial_{\binom{p}{k},\binom{h}{p}} \times^{(i)}_{(j)}(\kappa, \eta)$$

and all other derivatives vanish. Therefore

$$(\overline{\times}(K, H) * (\overline{K}, \overline{H}))^{(i)}_{(j)\gamma} = \overline{K}^{(i)}_{(l)\gamma} H^{(l)}_{(j)} + K^{(i)}_{(l)} \overline{H}^{(l)}_{(j)\gamma} + \sum_{\substack{(\alpha,\beta) \in \mathrm{Sh}^{-1}(\gamma) \\ |\alpha|,|\beta| \geqslant 1}} \overline{K}^{(i)}_{(l)\alpha} \overline{H}^{(l)}_{(k)\beta}.$$

The factor $1/2!$ is not present in the sum, since each non-vanishing second derivative is counted twice, as emphasised above. This expression coincides with (59). The last statement holds since $A(\cdot)B(\cdot) = \times(A, B)$. $\qquad \square$

Next we define a notion of pullback for controlled integrands.

Definition 2.19. Let X be as above, $F \in C^{\infty}(V, W)$, $\overline{H} \in \mathscr{D}_{F_*X}(\mathcal{L}(W, U))$. Let

$$F^*\overline{H} := (\overline{H} * \overline{F}(X)) \cdot \overline{DF}(X) \in \mathscr{D}_X(\mathcal{L}(V, U)). \qquad (61)$$

We will not need the coordinate expression of the pullback of a controlled path, although it can still be derived as done in other cases. The next proposition reassures us of the compatibility and associativity of some of the operations defined up to now.

Proposition 2.20.

(1) *Let* $\boldsymbol{X}, \overline{H}, \boldsymbol{H}, \overline{K}$ *be as in Example 2.15, then*

$$\uparrow_{\boldsymbol{X}}(\overline{K} * \overline{H}) = \uparrow_{\boldsymbol{H}}\overline{K} \in \mathscr{C}_{\omega}^{p}([0, T], S). \tag{62}$$

In particular lifting commutes with pushforwards

$$F_*(\uparrow_{\boldsymbol{X}}\overline{H}) = \uparrow_{\boldsymbol{X}}(F_*\overline{H}) \tag{63}$$

and furthermore $F_*(G_*\boldsymbol{X}) = (F \circ G)_*\boldsymbol{X}$ *for appropriately valued smooth functions* F, G;

(2) $(\overline{J} * \overline{K}) * \overline{H} =: \overline{J} * \overline{K} * \overline{H} := \overline{J} * (\overline{K} * \overline{H})$ *for appropriately valued controlled paths* $\overline{H}, \overline{K}, \overline{J}$, *and in particular* $F_* G_* \overline{H} = (F \circ G)_*\overline{H}$ *for appropriately valued smooth functions* F, G;

(3) $(\overline{J} \cdot \overline{K}) * \overline{H} = (\overline{J} * \overline{H}) \cdot (\overline{K} * \overline{H})$ *for appropriately valued controlled paths* $\overline{H}, \overline{K}, \overline{J}$; *in particular, taking* $\overline{J} = \overline{A}(H), \overline{K} = \overline{B}(H)$ *we have* $(A(\cdot)B(\cdot))_*\overline{H} = A_*\overline{H} \cdot B_*\overline{H}$;

(4) $(\overline{J} \cdot \overline{K}) \cdot \overline{H} =: \overline{J} \cdot \overline{K} \cdot \overline{H} := \overline{J} \cdot (\overline{K} \cdot \overline{H})$ *for appropriately valued controlled paths* $\overline{H}, \overline{K}, \overline{J}$;

(5) $F^*(G^*\overline{H}) = (G \circ F)^*\overline{H}$ *for appropriately valued smooth maps* F, G.

Proof. We begin with (1):

$$\uparrow_{\boldsymbol{X}}(\overline{K} * \overline{H})_{st}^{(c_1, \dots, c_m)}$$

$$= \sum_{\substack{(\gamma^1, \dots, \gamma^m) \in A_1^p \\ \gamma \in \overline{\mathrm{Sh}}(\gamma^1, \dots, \gamma^m)}} (\overline{K} * \overline{H})_{\gamma^1;s}^{c_1} \cdots (\overline{K} * \overline{H})_{\gamma^m;s}^{c_m} \boldsymbol{X}_{st}^{\gamma}$$

$$= \sum_{\substack{(\gamma^1, \dots, \gamma^m) \in A_1^p \\ \gamma \in \overline{\mathrm{Sh}}(\gamma^1, \dots, \gamma^m) \\ n_l = 1, \dots, |\gamma^l| \\ (\gamma^{l1}, \dots, \gamma^{ln_l}) \in \overline{\mathrm{Sh}}^{-1}(\gamma^l)}} (\overline{K}_{\boldsymbol{k}^1;s}^{c_1} \overline{H}_{\gamma^{11};s}^{k_1^1} \cdots \overline{H}_{\gamma^{1n_1};s}^{k_{n_1}^1}) \cdots (\overline{K}_{\boldsymbol{k}^m;s}^{c_m} \overline{H}_{\gamma^{m1};s}^{k_1^m} \cdots \overline{H}_{\gamma^{mn_m};s}^{k_{n_m}^m}) \boldsymbol{X}_{st}^{\gamma}$$

$$= \sum_{\substack{(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m) \in A_1^p \\ \boldsymbol{h} \in \overline{\mathrm{Sh}}(\boldsymbol{k}^1, \dots, \boldsymbol{k}^m)}} \overline{K}_{\boldsymbol{k}^1;s}^{c_1} \cdots \overline{K}_{\boldsymbol{k}^m;s}^{c_m} \sum_{\substack{(\delta^1, \dots, \delta^q) \in A_1^p \\ \delta \in \overline{\mathrm{Sh}}(\delta^1, \dots, \delta^q)}} \overline{H}_{\delta^1;s}^{h_1} \cdots \overline{H}_{\delta^q;s}^{h_q} \boldsymbol{X}_{st}^{\delta},$$

$$= (\uparrow_{\boldsymbol{H}}\overline{K})^{c_1, \dots, c_m}$$

and the statement follows by [17, Theorem 3.3.1]. As for the second statement

$$F_*(\uparrow_{\boldsymbol{X}}\overline{H}) = \uparrow_{\boldsymbol{H}}\overline{F}(H) = \uparrow_{\boldsymbol{X}}(\overline{F}(H) * \overline{H}) = \uparrow_{\boldsymbol{X}}(F_*\overline{H})$$

and

$$F_*(G_*\boldsymbol{X}) = F_*(\uparrow_{\boldsymbol{X}}\overline{G}(X)) = \uparrow_{\boldsymbol{X}}(F_*\overline{G}(X)) = \uparrow_{\boldsymbol{X}}(F_* G_*\overline{X}) = \uparrow_{\boldsymbol{X}}((F \circ G)_*\overline{X})$$

(where $\overline{X}$ is the $\boldsymbol{X}$-controlled path with trace X and zero Gubinelli derivatives) and the conclusion is implied by (2) below.

The proof of (2) is straightforward (with the second claim deduced from the Faà di Bruno formula (54)).

As for (3), we have

$$((\overline{J} \cdot \overline{K}) * \overline{H}))_{\gamma}^{\binom{a}{b}} = \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1,\ldots,\gamma^m) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\gamma)}} (\overline{J} \cdot \overline{K})_{\boldsymbol{k}}^{\binom{a}{b}} \overline{H}_{\gamma^1}^{k_1} \cdots \overline{H}_{\gamma^m}^{k_m}$$

$$= \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1,\ldots,\gamma^m) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\gamma) \\ (\boldsymbol{i},\boldsymbol{j}) \in \mathrm{Sh}^{-1}(\boldsymbol{k})}} \overline{J}_{\boldsymbol{i}}^{\binom{a}{c}} \overline{K}_{\boldsymbol{j}}^{\binom{c}{b}} \overline{H}_{\gamma^1}^{k_1} \cdots \overline{H}_{\gamma^m}^{k_m}$$

$$= \sum_{\substack{l+q=1,\ldots,|\gamma| \\ (\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathrm{Sh}^{-1}(\gamma) \\ (\alpha^1,\ldots,\alpha^l) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\alpha}) \\ (\beta^1,\ldots,\beta^q) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\boldsymbol{\beta})}} \overline{J}_{\boldsymbol{i}}^{\binom{a}{c}} \overline{K}_{\boldsymbol{j}}^{\binom{c}{b}} \overline{H}_{\alpha^1}^{i_1} \cdots \overline{H}_{\alpha^l}^{i_l} \overline{H}_{\beta^1}^{j_1} \cdots \overline{H}_{\beta^q}^{j_q}$$

$$= (\overline{J} * \overline{H}) \cdot (\overline{K} * \overline{H}),$$

where the second last identity follows from Lemma 1.4 (with $m = 2$).

As for (4), it is easy to show, using associativity of composition and of $\Delta_{\sqcup\!\sqcup}$, that both sides coincide with $\times^3 \circ (\overline{J} \boxtimes \overline{K} \boxtimes \overline{H}) \circ \Delta_{\sqcup\!\sqcup}^3$ where $\times^3$ denotes composition of three linear maps.

Finally, (5) is shown as follows:

$$F^* G^* \overline{H} = (G^* \overline{H} * \overline{F}(X)) \cdot \overline{DF}(X)$$

$$= (((\overline{H} * \overline{G}(F(X))) \cdot \overline{DG}(F(X))) * \overline{F}(X)) \cdot \overline{DF}(X)$$

$$= (\overline{H} * \overline{G}(F(X)) * \overline{F}(X)) \cdot (\overline{DG}(F(X)) * \overline{F}(X)) \cdot \overline{DF}(X)$$

$$= ((\overline{H} * \overline{G \circ F}(X)) \cdot \overline{D(G \circ F)}(X)$$

$$= (G \circ F)^* \overline{H}.$$

Here, we have used the previous points (2), (3) and (4) in the proposition, as well as the fact that

$$\overline{G}(F(X)) * \overline{F}(X) = G_* \overline{F}(X) = G_* F_* \overline{X} = (G \circ F)_* \overline{X} = \overline{G \circ F}(X)$$

and similarly that

$$(\overline{DG}(F(X)) * \overline{F}(X)) \cdot \overline{DF}(X) = \overline{DG \circ F}(X) \cdot \overline{DF}(X)$$

$$= \overline{DG \circ F(\cdot) DF(\cdot)}(X)$$

$$= \overline{D(G \circ F)}(X),$$

where we have used Proposition 2.18. □

In the next theorem we prove the property, well known in both ordinary and stochastic calculus, which allows to 'substitute the differential'. This will be especially convenient when manipulating RDEs. We have only introduced the theory necessary to handle weakly geometric rough paths, and the theorem is therefore stated in this context, but one can expect this type of result to also hold true in other settings, such as Itô calculus and branched rough paths.

Theorem 2.21 (Associativity of the rough integral). *Let $X \in \mathscr{C}_\omega^p([0,T], V)$, $\overline{H} \in \mathscr{D}_X(\mathcal{L}(V, W))$, $I := \int \overline{H}\mathrm{d}X$, $\overline{I}$ and I, respectively, the canonical controlled and rough paths above I, $\overline{K} \in \mathscr{D}_I(\mathcal{L}(W, U))$. Then*

$$\left(\overline{\int \overline{K}\mathrm{d}I}\right) * \overline{I} = \overline{\int (\overline{K} * \overline{I}) \cdot \overline{H}\mathrm{d}X} \tag{64}$$

and therefore

$$\int \overline{K}\mathrm{d}I = \int (\overline{K} * \overline{I}) \cdot \overline{H}\mathrm{d}X. \tag{65}$$

Proof. In this proof we will denote, for a tuple γ, $\gamma^{\cdot}$ its last entry and γ^- the tuple obtained by removing $\gamma^{\cdot}$, so $\gamma = (\gamma^-, \gamma^{\cdot})$. Moreover, we will interchangeably use the two indexing notations for controlled integrands, for example, $\overline{H}_\gamma^k = \overline{H}_{\gamma^-}^{\binom{k}{\gamma^{\cdot}}}$. For $|\gamma| \geqslant 1$ we then have

$$\left(\overline{\int (\overline{K} * \overline{I}) \cdot \overline{H}\mathrm{d}X}\right)_\gamma^c = ((\overline{K} * \overline{I}) \cdot \overline{H})_\gamma^c$$

$$= \sum_{(\alpha, \beta) \in \mathrm{Sh}^{-1}(\gamma^-)} (\overline{K} * \overline{I})_\alpha^{\binom{c}{h}} \overline{H}_\beta^{\binom{h}{\gamma^{\cdot}}}$$

$$= \sum_{\substack{(\alpha, \beta) \in \mathrm{Sh}^{-1}(\gamma^-) \\ (\alpha^1, \ldots, \alpha^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\alpha)}} \overline{K}_k^{\binom{c}{h}} \overline{H}_{\alpha^1}^{k_1} \cdots \overline{H}_{\alpha^n}^{k_n} \overline{H}_\beta^{\binom{h}{\gamma^{\cdot}}}$$

$$= \sum_{\substack{(\alpha, \beta) \in \mathrm{Sh}^{-1}(\gamma^-) \\ (\alpha^1, \ldots, \alpha^n) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\alpha)}} \overline{K}_{(k,h)}^c \overline{H}_{\alpha^1}^{k_1} \cdots \overline{H}_{\alpha^n}^{k_n} \overline{H}_{(\beta,\gamma^{\cdot})}^h$$

$$= \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1, \ldots, \gamma^m) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\gamma)}} \overline{K}_k^c \overline{H}_{\gamma^1}^{k_1} \cdots \overline{H}_{\gamma^m}^{k_m}$$

$$= \sum_{\substack{m=1,\ldots,|\gamma| \\ (\gamma^1, \ldots, \gamma^m) \in A_1 \cap \overline{\mathrm{Sh}}^{-1}(\gamma)}} \left(\overline{\int \overline{K}\mathrm{d}I}\right)_k^c \overline{I}_{\gamma^1}^{k_1} \cdots \overline{I}_{\gamma^m}^{k_m}$$

$$= \left(\left(\overline{\int \overline{K}\mathrm{d}I}\right) * \overline{I}\right)_\gamma^c.$$

At the trace level, we have, through a similar argument

$$\left(\overline{\int}(\overline{K}*\overline{I})\cdot\overline{H}\mathrm{d}\boldsymbol{X}\right)^c_{\gamma;st} \approx \sum_{|\gamma|=1}^{\lfloor p\rfloor}((\overline{K}*\overline{I})\cdot\overline{H})^c_{\gamma;s}\boldsymbol{X}^\gamma_{st}$$

$$= \sum_{\substack{|\gamma|=1,\ldots\lfloor p\rfloor\\ m=1,\ldots,|\gamma|\\ (\gamma^1,\ldots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma)}} \overline{K}^c_{k;s}\overline{H}^{k_1}_{\gamma^1;s}\cdots\overline{H}^{k_m}_{\gamma^m;s}\boldsymbol{X}^\gamma_{st}$$

$$\approx \int\overline{K}^c\mathrm{d}\boldsymbol{I}.$$

As for the statement at the level of rough paths, we have

$$\int\overline{K}\mathrm{d}\boldsymbol{I} = \uparrow_I\overline{\int}\overline{K}\mathrm{d}\boldsymbol{I}$$

$$= \uparrow_X\left[\left(\overline{\int}\overline{K}\mathrm{d}\boldsymbol{I}\right)*\overline{I}\right]$$

$$= \uparrow_X\left[\overline{\int}(\overline{K}*\overline{I})\cdot\overline{H}\mathrm{d}\boldsymbol{X}\right]$$

$$= \int(\overline{K}*\overline{I})\cdot\overline{H}\mathrm{d}\boldsymbol{X},$$

where we have used (1) in Proposition 2.20 and the previous statement. □

The next result, for which geometricity is essential, tells us that F_* and F^* behave as adjoint operators under the rough integral pairing. Its proof is an immediate consequence of Proposition 2.12 and Theorem 2.21.

Theorem 2.22 (Pushforward–pullback adjunction). *Let $\boldsymbol{X},\overline{H},F$ be as above, then*

$$\left(\overline{\int}\overline{H}\mathrm{d}F_*\boldsymbol{X}\right)*\overline{F}(X) = \overline{\int}F^*\overline{H}\mathrm{d}\boldsymbol{X} \tag{66}$$

and therefore

$$\int\overline{H}\mathrm{d}F_*\boldsymbol{X} = \int F^*\overline{H}\mathrm{d}\boldsymbol{X}. \tag{67}$$

Next we move on to the topic of RDEs. We will introduce two equivalent notions of solution to an RDE. Given a field of linear maps $F\in C^\infty(W,\mathcal{L}(V,W))$ and a smooth map $g\in C^\infty(W,U)$ we define, for $y\in W$

$$Fg(y) := Dg(y)\circ F(y) \in \mathcal{L}(V,U) \tag{68}$$

and inductively

$$F^n g(y) := F(\eta \mapsto F^{n-1} g(\eta))|_{\eta=y} \in \mathcal{L}(V, \mathcal{L}(V^{\otimes n-1}, U)) = \mathcal{L}(V^{\otimes n}, U). \tag{69}$$

In coordinates we denote

$$F_\gamma g^c(y) := Fg(y)^c_\gamma = \partial_k g(y) F^k_\gamma(y) \Rightarrow F^n g(y)^c_{(\gamma_1,...,\gamma_n)} = F_{\gamma_1} \cdots F_{\gamma_n} g^c(y).$$

We will also use the compact notation $F_\gamma g^c(y)$ for the latter.

Remark 2.23. Note that $F_{\gamma_1} \cdots F_{\gamma_n} g^c(y)$ can be read right to left as well as left to right, that is, it is equal to $F_{\gamma_1} \cdots F_{\gamma_{n-1}} (F_{\gamma_n} g^c)(y)$ for $n \geqslant 2$. This can be seen by induction on n (with the quantifier $\forall g$ inside the inductive hypothesis). For $n = 2$ the statement is tautological. For the inductive step we have

$$\begin{aligned}
F_{\gamma_1} \cdots F_{\gamma_{n+1}} g^c(y) &= F_{\gamma_1}(F_{\gamma_2} \cdots F_{\gamma_{n+1}} g^c)(y) \\
&= F_{\gamma_1}(F_{\gamma_2} \cdots F_{\gamma_n}(F_{\gamma_{n+1}} g^c))(y) \\
&= F_{\gamma_1} \cdots F_{\gamma_n}(F_{\gamma_{n+1}} g^c)(y),
\end{aligned}$$

where in the second identity we have used the inductive hypothesis.

Definition 2.24 (Davie solution to an RDE). Let $F \in C^\infty(W, \mathcal{L}(V, W))$. A *solution* to the RDE

$$dY = F(Y)dX, \quad Y_0 = y_0 \tag{70}$$

is a path $Y \in C([0, T], W)$ starting at y_0 with the property that for all $g \in C^\infty(W)$.

$$g(Y)_{st} \approx \sum_{n=1}^{\lfloor p \rfloor} \langle F^n g(Y_s), X^n_{st} \rangle. \tag{71}$$

Proposition 2.25 (Gubinelli solution to an RDE). *Let Y be a solution to (70). Then*

$$\overline{Y} := (Y, F(Y) = F\mathbb{1}(Y), ..., F^{\lfloor p \rfloor - 1} \mathbb{1}(Y)) \in \mathscr{D}_X(W) \tag{72}$$

and moreover

$$\overline{Y} = y_0 + \overline{\int} F_* \overline{Y} dX. \tag{73}$$

Conversely, if an X-controlled path satisfying the above identity, its trace satisfies Definition 2.24.

In order to prove this proposition we will make use of the following

Lemma 2.26. *The X-controlled path $\overline{Y}$ has the form (72) if and only if $\overline{Y}_n = (F_*\overline{Y})_{n-1}$ for $n \geqslant 1$. Moreover, in this case*

$$((Fg)_*\overline{Y})_{n-1} = F^n g(Y), \quad n \geqslant 1$$

with g as above.

Proof. If we prove

$$(Fg)_*(Y, F(Y), \dots, F^{\lfloor p \rfloor -1}\mathbb{1}(Y))_{n-1} = F^n g(Y),$$

we will have shown both the second statement and the 'only if' part of the first (just choose $g = \mathbb{1}$). In order to show this, we first establish the identity

$$F_\gamma g^c(y) = \sum_{\substack{(\gamma^1,\dots,\gamma^m)\in\overline{\mathrm{Sh}}^{-1}(\gamma) \\ |\gamma^1|,\dots,|\gamma^m|\geqslant 1}} \partial_k g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma^m}\mathbb{1}^{k_m}(y). \tag{74}$$

Note that this is not a closed-form formula for $F^n g$, since iterated compositions of the vector fields F_γ also appear on the right-hand side, but it will be useful for us nonetheless. We proceed by induction on $|\gamma|$. For $|\gamma| = 1$ there is nothing to show. For the inductive step, using Remark 2.23 we have

$$F_{(\gamma_1,\dots,\gamma_{n+1})}g^c(y)$$

$$= F_{\gamma_1}(F_{\gamma_2,\dots,\gamma_{n+1}}g^c)(y)$$

$$= F_{\gamma_1} \sum_{(\gamma^1,\dots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma_2,\dots,\gamma_{n+1})} \partial_k g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma^m}\mathbb{1}^{k_m}(y)$$

$$= \sum_{(\gamma^1,\dots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma_2,\dots,\gamma_{n+1})} \left(\partial_{h,k} g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma^m}\mathbb{1}^{k_m} \right.$$

$$\left. + \sum_{l=1}^m \partial_k g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots \partial_h(F_{\gamma^l}\mathbb{1}^{k_l}) \cdots F_{\gamma^m}\mathbb{1}^{k_m} \right) F_{\gamma_1}^h(y)$$

$$= \sum_{(\gamma^1,\dots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma_2,\dots,\gamma_{n+1})} \left(\partial_{h,k} g^c F_{\gamma_1}^h F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma^m}\mathbb{1}^{k_m} \right.$$

$$\left. + \sum_{l=1}^m \partial_k g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma_1,\gamma^l}\mathbb{1}^{k_l} \cdots F_{\gamma^m}\mathbb{1}^{k_m} \right)(y)$$

$$= \sum_{(\gamma^1,\dots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma_1,\dots,\gamma_{n+1})} \partial_k g^c F_{\gamma^1}\mathbb{1}^{k_1} \cdots F_{\gamma^m}\mathbb{1}^{k_m}(y).$$

Now, for $n = |\gamma| \geqslant 1$ we have

$$((Fg)_*\overline{Y})^c_\gamma = \sum_{(\gamma^1,\ldots,\gamma^m)\in\overline{\mathrm{Sh}}^{-1}(\gamma_1,\ldots,\gamma_{n-1})} \partial_k(F_{\gamma_n}g^c)F_{\gamma^1}\mathbb{1}^{k_1}\cdots F_{\gamma^m}\mathbb{1}^{k_m}(Y)$$

$$= F_{(\gamma_1,\ldots,\gamma_{n-1})}(F_{\gamma_n}g^c)(y)$$

$$= F_\gamma g^c(y),$$

where we have used Remark 2.23 and (74).

We now show the 'if' implication of the first statement. Namely, we need to show that if $\overline{Y} \in \mathscr{D}_X(W)$ has the property that $\overline{Y}_n = (F_*\overline{Y})_{n-1}$ for $n = 1,\ldots,\lfloor p\rfloor$, then $\overline{Y}_n = F^n\mathbb{1}(Y)$. We show this by induction on n. For $n = 1$ the assertion is obvious. For the inductive step we have

$$\overline{Y}^h_\gamma = \sum_{(\gamma^1,\ldots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma)} \partial_k F^h(Y)\overline{Y}^{k_1}_{\gamma^1}\cdots\overline{Y}^{k_1}_{\gamma^1}$$

$$= \sum_{(\gamma^1,\ldots,\gamma^m)\in A_1\cap\overline{\mathrm{Sh}}^{-1}(\gamma)} \partial_k F^h F_{\gamma^1}\mathbb{1}^{k_1}\cdots F_{\gamma^1}\mathbb{1}^{k_1}(Y)$$

$$= F_\gamma\mathbb{1}^h(Y),$$

where we have used (74) and the inductive hypothesis. $\square$

Proof of Proposition 2.25. Let Y be a Davie solution to the RDE. Taking g in (71) to be $\mathbb{1}, F,\ldots,F^{\lfloor p\rfloor-2}\mathbb{1}$ proves that $\overline{Y}$ defined in (72) is indeed an element of $\mathscr{D}_X(W)$. By Lemma 2.26 we then have $F_*\overline{Y} = (F(Y),\ldots,F^{\lfloor p\rfloor}\mathbb{1}(Y))$ and by Definition 2.7

$$\int_s^t F_*\overline{Y}\mathrm{d}X \approx \langle(F_*\overline{Y})_s, X^{\geqslant 1}_{st}\rangle \approx Y_{st}$$

again by the Davie definition. Since both the left- and right-hand sides are increments of paths, we conclude by [17, Theorem 3.3.1] that identity must hold. Therefore, since $Y_0 = y_0$, (73) holds at the trace level, and for $n \geqslant 1$

$$\left(y_0 + \overline{\int F_*\overline{Y}\mathrm{d}X}\right)_n = (F_*\overline{Y})_{n-1} = \overline{Y}_n.$$

Conversely, assume that there exists some $\overline{Y} \in \mathscr{D}_X(W)$ s.t. (73) holds: this implies that for $m \geqslant 0$

$$(F_*\overline{Y})_m = \left(y_0 + \overline{\int F_*\overline{Y}\mathrm{d}X}\right)_{m+1} = \overline{Y}_{m+1}$$

and therefore by Lemma 2.26 $\overline{Y}$ must have the form (72). Finally, $Y_0 = y_0$ and for $g \in C^\infty(W)$ and $Y := \uparrow_X\overline{Y}$

$$g(Y)_{st} = \int_s^t \overline{Dg}(Y)\mathrm{d}Y$$

$$= \int_s^t (\overline{Dg}(Y) * \overline{Y}) \cdot (\overline{F}(Y) * \overline{Y}) \mathrm{d}\boldsymbol{X}$$

$$= \int_s^t (\overline{Dg}(Y) \cdot \overline{F}(Y)) * \overline{Y} \mathrm{d}\boldsymbol{X}$$

$$= \int_s^t \overline{(Dg(\cdot)F(\cdot))(Y)} * \overline{Y} \mathrm{d}\boldsymbol{X}$$

$$= \int_s^t (Fg)_* \overline{Y} \mathrm{d}\boldsymbol{X}$$

$$\approx \sum_{n=1}^{\lfloor p \rfloor} \langle F^n g(Y_s), \boldsymbol{X}_{st}^n \rangle,$$

where we have used Theorem 2.21, Proposition 2.20 and Lemma 2.26. This concludes the proof. $\qquad\square$

The above proposition tells us that once we have the solution in the sense of Definition 2.24, we can obtain an $\boldsymbol{X}$-controlled path, and thus by Definition 2.8 a rough path. If we want to emphasise the existence of these superstructures, we will write

$$\mathrm{d}\overline{Y} = F(Y)\mathrm{d}\boldsymbol{X} \quad \text{and} \quad \mathrm{d}\boldsymbol{Y} = F(Y)\mathrm{d}\boldsymbol{X}, \quad Y_0 = y_0 \tag{75}$$

that is, $\boldsymbol{Y} := \uparrow_{\boldsymbol{X}} \overline{Y}$. Notice that the initial condition only involves the trace.

The next result will be instrumental in defining RDEs on manifolds in a coordinate-invariant manner.

Theorem 2.27 (Change of variable formula for RDE solutions). *Let $\boldsymbol{X}, F, \boldsymbol{Y}$ be as above, $g \in C^\infty(W, U)$. Then $(\boldsymbol{Y}, g_* \boldsymbol{Y})$ jointly solve the RDE*

$$\mathrm{d}\begin{pmatrix} \boldsymbol{Y} \\ \boldsymbol{Z} \end{pmatrix} = \begin{pmatrix} F(Y) \\ Dg(Y)F(Y) \end{pmatrix} \mathrm{d}\boldsymbol{X}. \tag{76}$$

In particular, if g is invertible, defining $C^\infty(U, \mathcal{L}(V, U)) \ni F_g(z) := Dg(g^{-1}(Z))F(g^{-1}(Z))$, $g_* \boldsymbol{Y}$ coincides with the rough path solution to

$$\mathrm{d}\boldsymbol{Z} = F_g(Z)\mathrm{d}\boldsymbol{X}. \tag{77}$$

Proof. Using Proposition 2.12, Theorem 2.21 and Proposition 2.20 we have

$$\mathrm{d}(g_* \boldsymbol{Y}) = y_0 + \int \overline{Dg}(Y)\mathrm{d}\boldsymbol{Y}$$

$$= y_0 + \int Dg_* \overline{Y} \cdot F_* \overline{Y} \mathrm{d}\boldsymbol{X}$$

$$= y_0 + \int (Dg(\cdot)F(\cdot))_* \overline{Y} \mathrm{d}\boldsymbol{X}.$$

This proves the first claim; as for the second, we continue

$$d(g_*Y) = \int (Dg(\cdot)F(\cdot))_* \overline{Y} dX$$

$$= \int (Dg(g^{-1}(\cdot))F(g^{-1}(\cdot)))_* g_* \overline{Y} dX,$$

where we have again used Proposition 2.20. This concludes the proof. □

The following theorem is proved in [4, Corollary 2.17, Theorem 4.2] in the case of $2 \leqslant p < 3$ and the proof carries over to the general case:

Theorem 2.28 (Local existence and uniqueness). *Precisely one of the following two possibility holds with respect to (70).*

(1) *A solution on $[0, T]$ exists.*
(2) *There exists an $S \leqslant T$ and a solution on $[0, S)$, with $Y_{[0,S)}$ not contained in any compact set of $\mathbb{R}^e$.*

Moreover, in either case, the solution is unique on the interval on which it is defined.

Example 2.29 (Non-autonomous RDEs). We can define RDEs that also depend on the driving signal, by 'doubling the variables' that is,

$$dY = F(Y, X)dX \stackrel{\text{def}}{\iff} d\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \mathbb{1} \\ F(Y, X) \end{pmatrix} dX. \tag{78}$$

This will be important when defining RDEs on manifolds, driven by a manifold-valued rough path.

3 | WEAKLY GEOMETRIC ROUGH PATHS ON MANIFOLDS

In this section we show how our algebraic framework for weakly geometric rough paths can be deployed to transfer the theory to the manifold setting. In this section M will be a smooth m-dimensional manifold and TM its tangent bundle. The following definition is similar in spirit to the one provided in [3].

Definition 3.1 (Manifold-valued rough path). Given a smooth atlas $\{\varphi : A_\varphi \to \mathbb{R}^m\}_\varphi$ of M, an M-valued *p-weakly geometric rough path* controlled by ω on $[0, T]$, $X \in \mathscr{C}_\omega^p([0, T], M)$ consists of a collection of rough paths ${}^\varphi X|_{[a_\varphi, b_\varphi]} \in \mathscr{C}_\omega^p([a_\varphi, b_\varphi], \mathbb{R}^m)$ for all possible collections of intervals $[a_\varphi, b_\varphi]$ (indexed by the chart φ in the atlas) s.t. $\text{Im}({}^\varphi X|_{[a_\varphi, b_\varphi]}) \subseteq \text{Range}(\varphi)$, and that for each pair of intervals $[a_\varphi, b_\varphi], [a_\psi, b_\psi]$

$$(\psi \circ \varphi^{-1})_* {}^\varphi X = {}^\psi X \in \mathscr{C}_\omega^p([a_\varphi, b_\varphi] \cap [a_\psi, b_\psi], \mathbb{R}^m). \tag{79}$$

The *trace* of X is the path $t \mapsto X_t := \varphi^{-1}({}^\varphi X_t) \in M$ whenever $t \in [a_\varphi, b_\varphi]$ (independently of φ), $X \in C^p([0, T], M)$.

It makes sense to allow the mappings $\varphi \mapsto {}^{\varphi}X$ and $\varphi \mapsto [a_\varphi, b_\varphi]$ to be multi-valued, so that the same chart can be used multiple times (for example, if the trace X goes back and forth between charts). The definition only depends on the smooth structure of M (that is, an equivalence class of atlases), since a rough path according to a particular atlas is uniquely extended to the maximal atlas for the smooth structure by reading (79) as a definition of the right-hand side. Also note that the definition is already fixed once, we have ${}^{\varphi}X|_{[a_\varphi, b_\varphi]}$ on a set of intervals, one for each chart, s.t. $\bigcup_\varphi (a_\varphi, b_\varphi) = (0, T)$: if the atlas if finite we call this a *finite representation* of X. The following example shows how our theory applies to the case of Stratonovich calculus of manifold-valued semimartingales.

Example 3.2 (Stratonovich rough path). Let X be an M-valued continuous semimartingale, that is, $f(X)$ is a real-valued semimartingale for all $f \in C^\infty(M)$. Since semimartingales are almost surely (a.s.) bounded p-variation for any $p > 2$, we only need to define a rough path above X up to level 2: we then define its *Stratonovich rough path* in coordinates by

$$\boldsymbol{X}_{st}^{\alpha\beta} := \int_s^t X_{su}^\alpha \circ dX_u^\beta, \tag{80}$$

where the integral is intended in the Stratonovich sense. This is well known to a.s. define a rough path in the linear setting, but since the above coordinate expression is taken according to a chart, we need the following lemma to establish that it defines a (stochastic) rough path in the sense of Definition 3.1.

Lemma 3.3. *Let X be an $\mathbb{R}^m$-valued continuous semimartingale, $f \in C^\infty(\mathbb{R}^m, \mathbb{R}^n)$ and $\boldsymbol{X}$ be defined as in (80). Then*

$$(f_* \boldsymbol{X})_{st}^{ij} = \int_s^t f^i(X) \circ df^j(X). \tag{81}$$

Proof. Using that, in the linear setting, Stratonovich integrals a.s. coincide with rough integrals against the Stratonovich rough path [9, Theorem 9.1]

$$\int_s^t f^i(X) \circ df^j(X) = \int_s^t f^i(X) \partial_\gamma f^j(X) \circ dX$$

$$= \int_s^t \overline{f}^i \overline{\partial}_\gamma \overline{f}^j(X) d\boldsymbol{X}^\gamma$$

$$\approx f^i \partial_\gamma f^j(X_s) X_{st}^\gamma + (\partial_\alpha f^i \partial_\beta f^j + f^i \partial_{\alpha\beta} f^j)(X_s) \boldsymbol{X}_{st}^{\alpha\beta}$$

$$\approx f^i(X_s) f^j(X)_{st} + \partial_\alpha f^i \partial_\beta f^j(X_s) \boldsymbol{X}_{st}^{\alpha\beta}.$$

We then have

$$\int_s^t f^i(X)_{su} df^j(X_u) = \int_s^t f^i(X) df^j(X) - f^i(X_s) f^j(X)_{st} \approx \partial_\alpha f^i \partial_\beta f^j(X_s) \boldsymbol{X}_{st}^{\alpha\beta}$$

concluding the proof. $\qquad\square$

We can define a path valued in a fixed vector space V and controlled by a manifold-valued rough path as follows.

Definition 3.4. Let $X \in \mathscr{C}_\omega^p([0,T], M)$. We define an V-valued X-controlled path $\overline{H} \in \mathscr{D}_X(V)$ to be a collection ${}^\varphi\overline{H} \in \mathscr{D}_{\varphi X}(V)$ with $\varphi, a_\varphi, b_\varphi$ as in Definition 3.1 and

$$ {}^\varphi\overline{H} * \overline{(\varphi \circ \psi^{-1})}({}^\psi X) = {}^\psi\overline{H}. $$

As for rough path, this only depends on the smooth structure of M. If we wish to define a controlled integrand, the trace of H must lie in the bundle $\mathcal{L}(TM, W)$ for some vector space W, 'above' X: for this we need a separate definition.

Definition 3.5 (Controlled integrand). Let $X \in \mathscr{C}_\omega^p([0,T], M)$. We define a W-valued X-controlled integrand $\overline{H} \in \mathscr{D}_X(\mathcal{L}(TM, W))$ to be a collection ${}^\varphi\overline{H} \in \mathscr{D}_{\varphi X}(\mathcal{L}(\mathbb{R}^m, W))$ with $\varphi, a_\varphi, b_\varphi$ as in Definition 3.1 and

$$ (\varphi \circ \psi^{-1})^* {}^\varphi\overline{H} = {}^\psi\overline{H}. \tag{82} $$

The *trace* of $\overline{H}$ is the path $H := {}^\varphi H \circ T_X\varphi$, which is valued in the fibre of X of the bundle $\mathcal{L}(TM, W)$ (if $W = \mathbb{R}^e$, then this bundle is $(T^*M)^e$.

We can now define rough integration on manifolds.

Definition 3.6 (Rough integral on manifolds). Let $X \in \mathscr{C}_\omega^p([0,T], M)$ and $\overline{H} \in \mathscr{D}_X(\mathcal{L}(TM, W))$. We define the *rough integral*

$$ \int_0^\cdot \overline{H} \mathrm{d}X := \sum_{[s_\varphi, t_\varphi]} \int_{s_\varphi}^{t_\varphi} {}^\varphi\overline{H} \mathrm{d}^\varphi X \in C_\omega^p([0,T], W), \tag{83} $$

where we are summing over a finite partition of $[0, \cdot]$ whose intervals $[s_\varphi, t_\varphi]$ are indexed by charts φ with the property that each $X_{[s_\varphi, t_\varphi]}$ is contained in the domain of φ.

This definition does not depend on the subdivision (since the integral taken with respect to two different subdivisions coincides with that take with respect to their common refinement) and does not depend on the charts used, since, for another chart ψ and $X_{[s,t]}$ contained in the domains of φ and ψ

$$ \int_s^t {}^\psi\overline{H} \mathrm{d}^\psi X = \int_s^t (\varphi \circ \psi^{-1})^* {}^\varphi\overline{H} \mathrm{d}(\psi \circ \varphi^{-1})_* {}^\varphi X $$

$$ = \int_s^t (\psi \circ \varphi^{-1})^* (\varphi \circ \psi^{-1})^* {}^\varphi\overline{H} \mathrm{d}^\varphi X $$

$$ = \int_s^t {}^\varphi\overline{H} \mathrm{d}^\varphi X $$

by Theorem 2.22 and Proposition 2.20.

We proceed to the topic of RDEs on manifolds: following the approach of [7], and for maximum generality, we will consider both the driving rough path and the solution to be manifold-valued; for this purpose we let N be an n-dimensional manifold.

Definition 3.7 (RDEs on manifolds). Let $\boldsymbol{X} \in \mathscr{C}_{\omega}^{p}([0,T], M)$ and F a section of the vector bundle $\mathcal{L}(TM, TN)$ (that is, the vector bundle over $N \times M$ with fibres $\mathcal{L}(TM, TN)_{y,x} := \mathcal{L}(T_x M, T_y N)$). We will say that $\boldsymbol{Y} \in \mathscr{C}_{\omega}^{p}([0,T], M)$ is a solution to the RDE

$$\mathrm{d}\boldsymbol{Y} = F(Y, X)\mathrm{d}\boldsymbol{X}, \quad Y_0 = y_0 \tag{84}$$

if for all charts φ and ψ on M and N, respectively, we have $Y_0 = y_0$ and on all intervals $[s,t]$ s.t. $X_{[s,t]}$ is contained in the domain of φ and $Y_{[s,t]}$ is contained in the domain of ψ

$$\mathrm{d}^{\psi}\boldsymbol{Y} = F(^{\psi}Y, {}^{\varphi}X)\mathrm{d}(^{\varphi}\boldsymbol{X}) \quad \text{on } [s,t]$$

in the sense of Example 2.29, with initial condition $^{\psi}Y_s$.

This does not depend on the chart: given other chats $\overline{\varphi}, \overline{\psi}$ we have

$$\mathrm{d}\begin{pmatrix} \overline{\varphi}\boldsymbol{X} \\ \overline{\psi}\boldsymbol{Y} \end{pmatrix} = \mathrm{d}\begin{pmatrix} (\overline{\varphi} \circ \varphi^{-1})_* {}^{\varphi}\boldsymbol{X} \\ (\overline{\psi} \circ \psi^{-1})_* {}^{\psi}\boldsymbol{Y} \end{pmatrix} = \begin{pmatrix} \mathbb{1} \\ F(^{\overline{\psi}}Y, {}^{\overline{\varphi}}X) \end{pmatrix} \mathrm{d}(^{\overline{\varphi}}\boldsymbol{X})$$

by Theorem 2.27 applied to the change of variable $(\overline{\varphi} \circ \varphi^{-1}, \overline{\psi} \circ \psi^{-1})$. Local existence can be inferred from Theorem 2.28, which in particular implies global existence if M is compact. More general conditions that guarantee global existence can be found in [6, 22].

We end with a few brief remarks which link the topic of this paper to the existing literature, without elaborating on the details.

Remark 3.8 (Stratonovich calculus). We can further expand on Example 3.2 to include the following: if H is a $\mathcal{L}(TM, W)$-valued semimartingale above the M-valued semimartingale X (that is, $H_t \in \mathcal{L}(T_{X_t}M, W)$), the Stratonovich integral $\int H\mathrm{d}X$ coincides a.s. with the rough integral $\int \overline{H}\mathrm{d}\boldsymbol{X}$ against the Stratonovich rough path $\boldsymbol{X}$ for any choice of $\overline{H}$ with trace H, and the solution to the Stratonovich differential equation $\mathrm{d}Y = F(Y, X)\mathrm{d}X$ coincides a.s. with the RDE (84).

Remark 3.9 (The extrinsic viewpoint). In [4] the topic of manifold-valued theory of rough paths, rough integration (specifically of 1-forms) and RDEs was treated from the extrinsic point of view. Here $\boldsymbol{X} \in \mathscr{C}_{\omega}^{p}([0,T], \mathbb{R}^d)$ is defined in [4, Definition 3.17] to be *constrained* to a smoothly embedded manifold M if its trace is M-valued and for all 1-forms $F \in \Gamma\mathcal{L}(T\mathbb{R}^d, W)$ (Γ denoting the space of sections)

$$\forall x \in M \; F(x)|_{T_x M} = 0 \; \Rightarrow \; \int \overline{F}(X)\mathrm{d}\boldsymbol{X} = 0 \in \mathscr{C}_{\omega}^{p}([0,T], W). \tag{85}$$

In [4, Corollary 3.32, Proposition 3.35] this is shown to be equivalent to the trace X being M-valued and $(I \otimes Q(X_s))\boldsymbol{X}_{st}^2 \approx 0$, or equivalently to $(P(X_s) \otimes P(X_s))\boldsymbol{X}_{st}^2 \approx \boldsymbol{X}_{st}^2$, where for $x \in M$ $P(x)$ is

the orthogonal projection $T_x\mathbb{R}^d \to T_xM$ and $Q := \mathbb{1} - P$. Moreover, one may replace $\boldsymbol{X}^2_{st}$ with its antisymmetric part $(\wedge\boldsymbol{X}^2)_{st}$ in these identities (because $\odot\boldsymbol{X}^2$ is already fixed by the trace).

This approach carries over to the case of higher p considered here. If $\boldsymbol{X} \in \mathscr{C}^p_\omega([0,T],\mathbb{R}^d)$, we may say that it is *constrained* to the smoothly embedded manifold M if $\pi_*\boldsymbol{X} = \boldsymbol{X}$, where π is the Riemannian projection of a tubular neighbourhood U of M onto M (that is, it maps a point in U to the unique point on M closest to it — this is well defined and smooth on a thin enough tubular neighbourhood). This extends the definition of [4] since $D_x\pi = P(x)$. In order to generalise the equivalent condition $(I \otimes Q(X_s))(\wedge\boldsymbol{X})_{st}$ we can take the log of our original condition, that is, $\log\pi_*\boldsymbol{X} = \log\boldsymbol{X}$: this has the advantage of eliminating all the redundancies of the former (as explained in [16, p. 767]), and its precise coordinate expression can be derived by using [11, Definition 7.20], but at higher orders cannot be described in terms of antisymmetric tensors. The Chen–Strichartz formula [2, Theorem 1.1], however, expresses $\log\boldsymbol{X}$ as a Lie polynomial; the task of expressing it in a basis of the Lie algebra is more complex still [20].

Care must be taken when defining the rough integral of a controlled integrand, since it is no longer the case that for an $\boldsymbol{X}$-controlled integrand $\overline{H}$, $\int \overline{H}\mathrm{d}\boldsymbol{X}$ (defined in the ordinary sense, where $\boldsymbol{X}$ is considered an element of $\mathscr{C}^p_\omega([0,T],\mathbb{R}^d)$), does not always only depend on the trace H of $\boldsymbol{X}$ restricted to TM, although this is indeed the case when $\overline{H}$ is given by a 1-form: this is because if $F \in \Gamma\mathcal{L}(T\mathbb{R}^d, W)$ (Γ denoting the space of sections) vanishes on TM (that is, $F(x)P(x) = 0$ for $x \in M$) by Theorem 2.22, we have

$$\int \overline{F}(X)\mathrm{d}\boldsymbol{X} = \int \overline{F}(X)\mathrm{d}\pi_*\boldsymbol{X} = \int \pi^*\overline{F}(X)\mathrm{d}\boldsymbol{X} = 0 \tag{86}$$

$$\text{since} \quad \pi^*\overline{F}(X) = (\overline{F}(X) * \overline{\pi}(X)) \cdot \overline{D\pi}(X) = \overline{F}(X) \cdot \overline{P}(X) = \overline{F(\cdot)P(\cdot)}(X) = 0,$$

where we have used Proposition 2.18. This then implies that if $F, G \in \Gamma\mathcal{L}(T\mathbb{R}^d, W)$ restrict to the same element of $\Gamma\mathcal{L}(TM, W)$, then $\int \overline{F}(X)\mathrm{d}\boldsymbol{X} = \int \overline{G}(X)\mathrm{d}\boldsymbol{X}$. Similarly, we have that if $\overline{H}, \overline{K} \in D_{\boldsymbol{X}}(\mathcal{L}(TM, W))$ are such that $\pi^*\overline{H} = \pi^*\overline{K}$, then $\int \overline{H}\mathrm{d}\boldsymbol{X} = \int \overline{K}\mathrm{d}\boldsymbol{X}$, but this involves conditions on all levels of $\overline{H}$, not just the trace (for a simple counterexample where this identity fails when only assuming $HP(X) = KP(X)$ see [1, Example 4.3]).

Finally, the original definition of constrained rough path given by integration (85) also carries over to higher p. The fact that this is implied by $\pi_*\boldsymbol{X} = \boldsymbol{X}$ was shown in (86). For the converse, we rewrite the identity as $(\mathbb{1} - \pi)_*\boldsymbol{X} = 0$: the trace level is implied by the fact that X is M-valued, and at orders $\geqslant 1$ the identity $\int \overline{Q}(X)\mathrm{d}\boldsymbol{X} = (\mathbb{1} - \pi)_*\boldsymbol{X}$ is straightforward to check.

Remark 3.10 (Parallel transport and Cartan development). If M has a connection ∇ (with Christoffel symbols Γ), given $\boldsymbol{X} \in \mathscr{C}^p_\omega([0,T],M)$ we can define parallel transport of vectors above its trace X as the solution to the RDE driven by $\boldsymbol{X}$, valued in TM, defined by the horizontal lift based at X: in local coordinates the parallel frame path $t \mapsto A_t \in T_{X_t}M$ satisfies

$$\mathrm{d}A^\gamma = -\Gamma^\gamma_{\alpha\beta}(X)A^\beta\mathrm{d}\boldsymbol{X}^\alpha. \tag{87}$$

The Cartan development $\boldsymbol{Y}$ of a $\boldsymbol{Z} \in \mathscr{C}^p_\omega([0,T], T_oM)$ (with $o \in M$ a fixed basepoint) can be viewed as the projection onto M of an FM-valued RDE driven by $\boldsymbol{Z}$, defined the fundamental horizontal vector fields (the full solution additionally consists of a parallel frame above the developed path):

in local coordinates

$$\begin{cases} \mathrm{d}\boldsymbol{Y}^k = A_\gamma^k \mathrm{d}\boldsymbol{Z}^\gamma \\ \mathrm{d}A_\gamma^k = -\Gamma_{ij}^k(Y) A_\alpha^i A_\gamma^j \mathrm{d}\boldsymbol{Z}^\alpha \end{cases}. \tag{88}$$

Here we have emphasised how development actually defines a (possibly explosive) M-valued rough path $\boldsymbol{Y}$, and not just its trace.

JOURNAL INFORMATION

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. J. Armstrong, D. Brigo, T. Cass, and E. Rossi Ferrucci, *Non-geometric rough paths on manifolds*, arXiv:2007.06970, 2020, https://arxiv.org/abs/2007.06970.
2. F. Baudoin, *An introduction to the geometry of stochastic flows*, Published by Imperial College Press, London, and distributed by World Scientific Publishing Co., 2004.
3. Y. Boutaib and T. Lyons, *A new definition of rough paths on manifolds*, arXiv:1510.07833v2, 2015, https://arxiv.org/abs/1510.07833v2.
4. T. Cass, B. K. Driver, and C. Litterer, *Constrained rough paths*, Proc. Lond. Math. Soc. (3) **111** (2015), no. 6, 1471–1518.
5. T. Cass, B. K. Driver, N. Lim, and C. Litterer, *On the integration of weakly geometric rough paths*, J. Math. Soc. Japan **68** (2016), no. 4, 1505–1524.
6. B. Driver, *Global existence of geometric rough flows*, arXiv:1810.03708, 2018, https://arxiv.org/abs/1810.03708.
7. M. Émery, *Stochastic calculus in manifolds*, Universitext, Springer, Berlin, 1989. With an appendix by P.-A. Meyer.
8. K. Ebrahimi-Fard and F. Patras, *Cumulants, free cumulants and half-shuffles*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **471** (2015), no. 2176, 20140843.
9. P. K. Friz and M. Hairer, *A Course on rough paths with an introduction to regularity structures*, Universitext, Springer International Publishing, Cham, 2014.
10. L. Foissy, *An introduction to hopf algebras of trees*, 2013.
11. P. K. Friz and N. B. Victoir, *Theory and applications*, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010.
12. M. Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140.
13. M. Gubinelli, *Ramification of rough paths*, J. Differential Equations **248** (2010), no. 4, 693–721.
14. M. Hairer and D. Kelly, *Geometric versus non-geometric rough paths*, Ann. Inst. H. Poincaré Probab. Statist. **51** (2015), no. 1, 207–251.
15. T. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths ecole d'eté de probabilités de saint-flour xxxiv-2004*, 2007.
16. T. J. Lyons and N. Sidorova, *On the radius of convergence of the logarithmic signature*, Illinois J. Math. **50** (2006), no. 1–4, 763–790.

17. T. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 215–310.
18. T. Lyons, *System control and rough paths*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2002.
19. D. Manchon, *Hopf algebras, from basics to applications to renormalization*, arXiv:math/0408405v2, 2006. https://arxiv.org/abs/math/0408405.
20. J. Reizenstein, *Calculation of iterated-integral signatures and log signatures*, arXiv:1712.02757, 2017. https://arxiv.org/abs/1712.02757.
21. C. Reutenauer, *Free lie algebras*, London Mathematical Society monographs, New Series 7, Clarendon, Oxford, 1993.
22. M. Weidner, *A geometric view on rough differential equations*, Ph.D. thesis, Imperial College London, 2018. https://spiral.imperial.ac.uk:8443/handle/10044/1/62658.