51st IEEE Conference on Decision and Control
December 10-13, 2012. Maui, Hawaii, USA

Cascaded Analytic Nonlinear Systems Driven by Rough Paths*

Luis A. Duffaut Espinosa†, W. Steven Gray‡, and Makhin Thitsa‡

***Abstract*— It was recently shown that the lack of a suitable probabilistic characterization of the input process for a system of interconnected analytic nonlinear input-output maps is an obstacle to well-posedness. For example, the cascade connection of two such systems is only known to be well-posed when a certain independence property is preserved by the first system in the connection. Hence, it appears that some alternative characterization of an input process is needed in this setting. One possibility is to employ T. Lyons' construction of a rough path. This concept employs the p-variation of a path and Chen's identity in order to extend the notion of integration with respect to paths having finite p-variation for $p \geq 1$. The primary advantage of such an approach in the context of system interconnections is that independence is no longer needed for producing well-posed cascaded analytic nonlinear systems.**

I. INTRODUCTION

In many applications, input-output systems are interconnected to form more complex systems. Describing the nature of the composite system and providing some explicit parametrization for it are generally nontrivial problems when the subsystems are nonlinear. The systems of interest here belong to the class of analytic nonlinear integral operators known as *Fliess operators* [5], [9], [14]. It was recently shown in [2], [3] that the lack of a suitable probabilistic characterization of the input process for interconnections of such systems is an obstacle to well-posedness. For example, the cascade connection of two Fliess systems as shown in Fig. 1 is only known to be well-posed when a certain independence property is preserved by the first system in the connection. Hence, it appears that some alternative characterization of an input process is needed in this setting. One possibility is to employ T. Lyons' construction of a *rough path* [6], [12], [13]. This concept employs p-variation paths and Chen's identity in order to extend the notion of integration with respect to paths having finite p-variation for $p \geq 1$. It will be shown in this paper that the primary advantage of such an approach is that independence is no longer needed for producing well-posed cascaded Fliess operators.

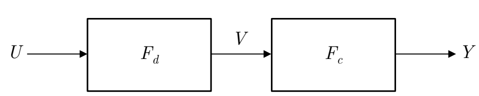

Fig. 1: Cascade connection of two Fliess operators.

The specific goals of the paper are to introduce the *Fliess signature operator*, which maps the input *signature* of a p-variation input to its corresponding output signature, and then to characterize the cascade interconnection for this class of operators. The first step is to introduce the notion of a path's signature as a substitute for its functional form. Lyons' rough path theory is then applied directly to characterize such a signature as a p*-rough path*. The next step is to introduce an algebraic device known as a *transduction* in order to define the Fliess signature operator. Finally, a characterization of the cascade interconnection is made in terms of transductions. In this setting, it can be shown directly that the composite system produces a well-defined output path, thus solving the open problem.

The paper is organized as follows. Section II summarizes the basics of rough path theory and Fliess operator theory used throughout the paper. In Section III, the definitions of a Fliess operator and a Fliess signature operator driven by a p-rough path are presented. It is then shown that the output of a Fliess signature operator is a p-rough path. Finally, Section IV gives a characterization of the cascade of Fliess signature operators in terms of the composition of their corresponding transductions. Section V gives the paper's conclusions.

II. FLIESS OPERATORS WITH ROUGH PATH INPUTS

A finite nonempty set of noncommuting symbols $X = \{x_0, x_1, \ldots, x_m\}$ is called an *alphabet*. Each element of X is called a *letter*, and any finite sequence of letters from X, $\eta = x_{i_1} \cdots x_{i_k}$, is called a *word* over X. The *length* of η, $|\eta|$, is the number of letters in η. The number of times x_i appears in η is designated by $|\eta|_{x_i}$. The set of all words with length k is denoted by X^k. The set of all words including the empty word, $\emptyset$, is written as X^*. Clearly X^* forms a monoid under catenation. Any mapping $c : X^* \to \mathbb{R}^\ell$ is called a *formal power series*. The value of c at $\eta \in X^*$ is written as (c, η). Typically, c is represented as the formal sum $c = \sum_{\eta \in X^*} (c, \eta)\, \eta$. For any language $L \subseteq X^*$, its *characteristic series* is defined as $\mathrm{char}(L) = \sum_{\eta \in L} \eta$. The collection of all formal power series over X is denoted by $\mathbb{R}^\ell\langle\langle X\rangle\rangle$, while the set of polynomials over X is designated by $\mathbb{R}\langle X\rangle$. Each of these sets forms an associative $\mathbb{R}$-algebra under the catenation (Cauchy) product and a commutative and associative $\mathbb{R}$-algebra under the shuffle product, that is, the $\mathbb{R}$-bilinear mapping $\mathbb{R}^\ell\langle\langle X\rangle\rangle \times \mathbb{R}^\ell\langle\langle X\rangle\rangle \to \mathbb{R}^\ell\langle\langle X\rangle\rangle$ uniquely specified by the shuffle product of two words

$$(x_i\eta) \sqcup\!\sqcup (x_j\xi) = x_i(\eta \sqcup\!\sqcup (x_j\xi)) + x_j((x_i\eta) \sqcup\!\sqcup \xi),$$

where $x_i, x_j \in X,\ \eta, \xi \in X^*$ and with $\eta \sqcup\!\sqcup \emptyset = \eta$ [5], [14]. Finally, the *left-shift operator* for any $\xi \in X^*$ is defined as $\xi^{-1} : X^* \to \mathbb{R}\langle X\rangle$ such that $\xi^{-1}(\eta) = \eta'$ when $\eta = \xi\eta'$ with $\eta' \in X^*$, and 0 otherwise.

A. Rough Paths

The rough path theory presented here is based on the treatment in [6], [12], [13]. A path on $J \triangleq [0, T]$ is a function

*The first author was supported by the Australian Research Council (ARC) and Air Force Office of Scientific Research (AFOSR) under agreement number FA2386-09-1-4089.

†School of Engineering and Information Technology, University of New South Wales at ADFA, Canberra, ACT 2600, Australia, l.duffaut@adfa.edu.au

‡Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia 23529-0246, USA, {sgray,mthitsa}@odu.edu

$U : J \to \mathbb{R}^m$. Let $\mathcal{D}_r \triangleq \{0 < t_1 < \cdots < t_r = T\}$ be a partition of J, and $\mathcal{D}(J)$ denotes the set of all finite partitions of J. The *p-variation* of a path U is

$$\|U\|_{p,J} \triangleq \left(\sup_{\mathcal{D}_r \in \mathcal{D}(J)} \sum_{t_l \in \mathcal{D}_r} |U(t_{l+1}) - U(t_l)|^p \right)^{\frac{1}{p}}.$$

Note that $\|U\|_{p,J} = 0$ only when U is constant, so a norm for the vector space $\mathcal{V}^p(J) \triangleq \{U : J \to \mathbb{R}^d : \|U\|_{p,J} < \infty\}$ is $\|U\|_{\mathcal{V}^p,J} \triangleq \|U\|_{p,J} + \sup_{t \in J} |U(t)|$. It is standard that when $p = 1$ and $\|U\|_{\mathcal{V}^p,J} < \infty$, the path U can act as the integrator for a Stieltjes type integral since U defines a function of bounded variation. In fact, Fliess originally introduced his input-output operators as weighted sums of Stieltjes iterated integrals [5]. Let $\Delta_T \triangleq \{(s,t) \in [0,T] \times [0,T] : 0 \le s \le t \le T\}$ and define the continuous map $\omega : \Delta_T \to \mathbb{R}$ such that

$$\omega(s,\tau) + \omega(\tau,t) \le \omega(s,t)$$

and $\omega(\tau,\tau) = 0$ for $s \le \tau \le t \in J$. Throughout this paper, ω will be called a *ω-function*. In rough path theory an ω-function is normally referred to as a *control function*. However, that name is not adopted here to avoid the obvious confusion with system theory terminology. It is known that any convex function of an ω-function is also an ω-function [6, Section 1.2]. If ω satisfies $|U(t) - U(s)| \le \omega(s,t)$ for $U \in \mathcal{V}^p(J)$, then it can be verified that $\|U\|_{p,[s,t]} \le \omega(s,t)^{\frac{1}{p}}$ for all $(s,t) \in \Delta_T$. Moreover, ω provides a reparametrization such that U becomes a $1/p$-Hölder continuous path. In particular, a natural ω-function for $U \in \mathcal{V}^1(J)$ is $\omega(s,t) = \|U\|_{1,[s,t]}$ such that $|U(t) - U(s)| \le \|U\|_{1,[s,t]}$. This mapping is additive in that $\|U\|_{1,[s,t]} = \|U\|_{1,[s,\tau]} + \|U\|_{1,[\tau,t]}$, and $\|U\|_{1,[0,t]}$ as a function of t is continuous. Thus, it has a finite 1-variation. Similarly for $U \in \mathcal{V}^p(J)$, one has $\omega(s,t) = \|U\|^p_{p,[s,t]}$.

Although rough path theory can be developed for paths taking values in a general Banach space, E, the focus here is on the case where $E = \mathbb{R}^{m+1}$ with the usual norm. Moreover, rough paths are normally defined on the tensor space

$$T((E)) = \{c = (c_0, c_1, \ldots) : c_n \in E^{\otimes n}, n \ge 0\} = \bigoplus_{n=0}^{\infty} E^{\otimes n},$$

where $E^{\otimes n} \triangleq \operatorname{span}_{\mathbb{R}}\{e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_n} : i_1, i_2, \ldots, i_n \in I^n, I = \{0,1,\ldots,n\}\}$ is the $\mathbb{R}$-vector space of all multilinear forms of length n, and $e_0, e_1, \ldots, e_m$ is a basis for E. It is known, however, that the space $T((E))$ is isomorphic to $\mathbb{R}\langle\langle X \rangle\rangle$. Thus, for each $c = (c_0, c_1, \ldots) \in T((E))$ there corresponds a unique $c = \sum_{\eta \in X^*} (c,\eta)\eta \in \mathbb{R}\langle\langle X \rangle\rangle$ and vice-versa. Therefore, the entire development will be done using only $\mathbb{R}\langle\langle X \rangle\rangle$. The sum and product in $T((E))$ are identified with the sum and Cauchy product of formal power series. The space $T((E))$ truncated to order n, denoted by $T^{(n)}(E)$, is isomorphic to the $\mathbb{R}$-vector space of all polynomials of degree n, $\mathbb{R}^{(n)}\langle X \rangle$. Define the $\mathbb{R}$-algebra homomorphism $\pi_n : \mathbb{R}\langle\langle X \rangle\rangle \to \mathbb{R}^{(n)}\langle\langle X \rangle\rangle$ such that $c_{(n)} = \pi_n(c)$ is the truncation of c to order n. Then, one can define an associative product of $c, d \in \mathbb{R}^{(n)}\langle\langle X \rangle\rangle$ as $cd \triangleq \pi_n(cd)$.

Let $U \in \mathcal{V}^1(J)$. For every $\eta \in X^*$, denote by $E_\eta[U]$ the iterated Stieltjes integral of U defined inductively by setting $E_\emptyset[U] = 1$ and

$$E_{x_i\eta'}[U](t_2,t_1) = \int_{t_1}^{t_2} dU_i(\tau) E_{\eta'}[U](\tau,t_1),$$

where $x_i \in X$ and $\eta' \in X^*$. If U is differentiable then $E_{x_i\eta'}[u](t_2,t_1) = \int_{t_1}^{t_2} u_i(\tau) E_{\eta'}[U](\tau,t_1)\, d\tau$, where $u_i = dU_i/dt$.

Definition 1: A **multiplicative functional of degree $n \ge 1$** is a continuous map $P_n[U] : \Delta_T \to \mathbb{R}^{(n)}\langle X \rangle$ defined as $P_n[U](t_2,t_1) \triangleq \sum_{\eta \in X^k, k \le n} E_\eta[U](t_2,t_1)\,\eta$ that satisfies

$$P_n[U](t_2,t_1) = \pi_n(P_n[U](t_2,\tau) P_n[U](\tau,t_1))$$

for $t_1 \le \tau \le t_2$ and is referred to as **Chen's identity**.

A notion of p-variation for $P_n[U]$ is given next.

Definition 2: The map $P_n[U] : \Delta_T \to \mathbb{R}^{(n)}\langle X \rangle$ is said to have **finite total p-variation** if

$$\sup_{\mathcal{D}_r \in \mathcal{D}(J)} \sum_{t_l \in \mathcal{D}_r} |E_\eta[U](t_l,t_{l-1})|^{\frac{p}{|\eta|}} < \infty$$

when $|\eta| \le n$.

Lemma 1: [13, Proposition 3.3.2] Let $P_n[U]$ be a multiplicative functional of order n. If $P_n[U]$ has finite total p-variation then for $(s,t) \in \Delta_T$

$$\omega(s,t) = \sum_{j=1}^{n} \sum_{\eta \in X^j} \sup_{\mathcal{D}_r \in \mathcal{D}(J)} \sum_{t_l \in \mathcal{D}_r} |E_\eta[U](t_l,t_{l-1})|^{\frac{p}{|\eta|}}$$

is a ω-function, and $\|E_\eta[U](t,s)\|_{\mathcal{V}^p(J)} \le \omega(s,t)^{\frac{|\eta|}{p}}$.

For paths with p-variation greater than 1, the idea is to make sense of the differential dU. It is known for ordinary differential equations that by taking the limit of the first order increment processes $E_{x_i}[U](t,s) = \{U_i(t) - U_i(s) : 0 \le s \le t\}$, $i = 0,1,\ldots,m$, as the partition size goes to zero, one can obtain a useful characterization of dU without ever computing them explicitly. Therefore, one may regard the whole collection of first order increments as the *differential* dU when $p = 1$. Employing Chen's identity,

$$E_\eta[U](t_2,t_1) = \lim_{\substack{r \to \infty \\ |\mathcal{D}_r| \to 0}} \sum_{\substack{t_l \in \mathcal{D}_r \\ \eta = \eta_1\eta_2 \\ \eta_1,\eta_2 \ne \emptyset}} E_{\eta_1}[U](t_l,t_{l-1}) E_{\eta_2}[U](t_{l-1},t_1), \quad (1)$$

where $|\mathcal{D}_r|$ denotes the partition size. So all higher order increments are well-defined in terms of the first order increment process, i.e.,

$$\sup_{\mathcal{D}_r \in \mathcal{D}(J)} \sum_{t_l \in \mathcal{D}_r} |E_\eta[U](t_l,t_{l-1})|^{\frac{p}{|\eta|}} < \infty, \quad \forall \eta \in X^*, \quad (2)$$

where $p = 1$. To describe dU when $p > 1$, one needs higher order increments as well. Observe that (2) is not satisfied even for $|\eta| = 1$ and $p = 2$, but the increments do satisfy the weaker condition $\sup_{\mathcal{D}_r \in \mathcal{D}(J)} \sum_{t_l \in \mathcal{D}_r} |U(t_l) - U(t_{l-1})|^{p'} < \infty$ for any $p' \ge p$. Thus, the iterated path integrals defined in (1) do converge in a p-variation metric, and given $\{E_\nu[U] : |\nu| =$

$1, 2, \ldots \lfloor p \rfloor\}$, (1) uniquely determines $E_\eta[U]$ when $\eta \in X^k$, $k \geq \lfloor p \rfloor + 1$.

Theorem 1: [12], [13] Assume $p \geq 1$, and let $P_n[U]$ be a multiplicative functional of order n with finite p-variation so that

$$\|E_\eta[U]\|_{p,[s,t]} \leq \omega(s,t)^{\frac{|\eta|}{p}}$$

for $|\eta| \leq n$ and some ω-function. If $n \geq \lfloor p \rfloor$ then $P_n[U]$ can be extended uniquely to a finite p-variation multiplicative functional $P[U] \in \mathbb{R}\langle\langle X \rangle\rangle$. Moreover, there exist a $K > 0$ such that for *every* $\eta \in X^*$ it follows that

$$\|E_\eta[U]\|_{p,[s,t]} \leq \frac{K\omega(s,t)^{\frac{|\eta|}{p}}}{\left(\frac{|\eta|}{p}\right)!}, \ \ \forall (s,t) \in \Delta_T. \quad (3)$$

This result is known as the *extension theorem*, and it is the first fundamental theorem of rough path theory. In this sense, p indicates how many iterated path integrals (or higher order increments) are needed in order to have a well-posed integration theory. A crucial result used in the proof of the extension theorem is the so-called *neo-classical inequality* given in the next lemma.

Lemma 2: [10], [12] Let $p \geq 1$, $n \in \mathbb{N}$ and $t_1, t_2 \geq 0$,

$$\frac{1}{p}\sum_{i=0}^{n} \frac{t_1^{\frac{i}{p}}}{\left(\frac{i}{p}\right)!} \frac{t_2^{\frac{n-i}{p}}}{\left(\frac{n-i}{p}\right)!} \leq \frac{(t_1+t_2)^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!}.$$

Let $U = (U_1, \ldots, U_m) \in \mathcal{V}^p(J)$ and $U_0(t) \triangleq t$. The *signature* (or *Chen series*) associated with U is an element of $\mathbb{R}\langle\langle X \rangle\rangle$ defined as

$$P[U](t_2, t_1) = \sum_{\eta \in X^*} E_\eta[U](t_2, t_1)\, \eta.$$

Here all iterated integrals are in the Lyons sense. In particular, if

$$P[U](t_2, t_1) = P[U](t_2, \tau) P[U](\tau, t_1), \ \ t_1 \leq \tau \leq t_2 \quad (4)$$

then $P[U]$ is said to be *multiplicative*. Identity (4) was originally given as a theorem by Chen in [1] when $p = 1$. However, in rough path theory it is a purely algebraic property that a path with arbitrary p-variation must satisfy in order to behave properly as an integrator.

Definition 3: A ***p*-rough path** is a multiplicative functional of degree $\lfloor p \rfloor$ in $\mathbb{R}^m$ having finite p-variation. The space of p-rough paths is denoted by $\Omega_p(\mathbb{R}^m)$.

Thus, a p-rough path is a continuous mapping from Δ_T to $\mathbb{R}^{(\lfloor p \rfloor)}\langle X \rangle$, which is a multiplicative functional of degree $\lfloor p \rfloor$ and has finite total p-variation.

An important class of rough paths is the subset of all rough paths that are limits of 1-rough paths in the p-variation metric. They are known as *geometric rough paths*, and the set of all such paths is denoted by $G\Omega_p(\mathbb{R}^m)$. This set constitutes the input class for Fliess signature operators as defined later in Section III. The goal is then to describe a Fliess signature operator that maps a p-rough path to another p-rough path, thus allowing one to drive a second Fliess signature operator with this signal.

B. Fliess Operators

A Fliess operator is formally defined in terms of the signature of a path as follows.

Definition 4: Let $c \in \mathbb{R}^\ell\langle\langle X \rangle\rangle$ and $U \in G\Omega_p(\mathbb{R}^m)$. The corresponding **Fliess operator** driven by U is

$$F_c[U] = (c, P[U]) \triangleq \sum_{\eta \in X^*} (c, \eta)(P[U], \eta) = \sum_{\eta \in X^*} (c, \eta) E_\eta[U].$$

Note in this setting that the output path generated by a Fliess operator is $Y = \int d(F_{x_0 c}[U])$. Abusing the notation, the output path will be denoted $Y = F_{x_0 c}[U]$. The integral only coincides with Lebesgue integration when $p = 1$. This notation is analogous to the one used when generating a Wiener process by integrating white Gaussian noise. In addition, this reinforces the fact that a Fliess operator output is obtained as an approximation of smooth signals. A condition must be imposed on c in order to give some notion of convergence for a Fliess operator.

Definition 5: A series $c \in \mathbb{R}^\ell\langle\langle X \rangle\rangle$ is called **Gevrey of degree *r*** if

$$|(c, \eta)| \triangleq \max_i |(c_i, \eta)| \leq K M^{|\eta|} (|\eta|!)^r, \ \ \eta \in X^*$$

for some $K, M > 0$ and $r \geq 0$. The set of all such series is denoted by $\mathbb{R}^\ell_{G(r)}\langle\langle X \rangle\rangle$.

Theorem 2: Let $c \in \mathbb{R}^\ell_{G(1/p')}\langle\langle X \rangle\rangle$. If $U \in G\Omega_p(\mathbb{R}^m)$ with $p < p'$ and $Y = F_{x_0 c}[U]$, then $\|F_c[U]\|_{\mathcal{V}^p(J)} < \infty$. Under the same assumptions, if $p = p'$ and $M(m + 1)\,\omega(0,t)^{1/p} < 1$ then $\|F_c[U]\|_{\mathcal{V}^p(J)} < \infty$.

Thus, given an input in $G\Omega_p(\mathbb{R}^m)$, each output component generated by F_c is in $\mathcal{V}^p$. But this means that a cascade interconnection is still not well-posed since being in $\mathcal{V}^p$ does not imply that the output is a p-rough path.

Let $W = \{w_1, \ldots, w_m\}$ and $X = \{x_1, \ldots, x_{m'}\}$ be two alphabets. When F_c and F_d with $c \in \mathbb{R}^\ell\langle\langle W \rangle\rangle$ and $d \in \mathbb{R}^m\langle\langle X \rangle\rangle$ are interconnected in a cascade fashion, the composite system always has at least a formal Fliess operator representation in terms of the composition product [7]. It is convenient to describe this product using a family of mappings $D_{w_i} : \mathbb{R}\langle\langle X \rangle\rangle \to \mathbb{R}\langle\langle X \rangle\rangle : e \mapsto x_0(d_i \sqcup\!\sqcup e)$, where $i = 0, 1, \ldots, m$ and $d_0 := 1$. Let $D_\emptyset$ be the identity map. Such maps can be composed in an obvious way so that $D_{w_i w_j} := D_{w_i} D_{w_j}$ provides an $\mathbb{R}$-algebra.

Definition 6: [7] The **composition product** of $\eta \in W^*$ and $d \in \mathbb{R}^m\langle\langle X \rangle\rangle$ is

$$\begin{aligned}\eta \circ d &\triangleq (w_{i_k} w_{i_{k-1}} \cdots w_{i_1}) \circ d \\ &= D_{w_{i_k}} D_{w_{i_{k-1}}} \cdots D_{w_{i_1}}(1) \\ &= D_\eta(1).\end{aligned}$$

For any $c \in \mathbb{R}^\ell\langle\langle W \rangle\rangle$, $c \circ d \triangleq \sum_{\eta \in W^*} (c, \eta)\, \eta \circ d$. The composition product is associative and satisfies

$$F_c \circ F_d = F_{c \circ d}.$$

In the current context, the integration in F_c is understood to be with respect to $dV = dF_{x_0 d}[U]$. Therefore, $F_c[V] = F_c \circ F_d[U]$ and *not* $F_c[V] = F_c \circ F_{x_0 d}[U]$ as the notation might suggest.

III. Fliess Signature Operators

The goal of this section is to explicitly describe the mapping between the input signature and the output signature of a Fliess operator. To achieve this objective, a device known as a *transduction* is employed.

Definition 7: [4], [11] Let X and W be two alphabets. Any $\mathbb{R}$-linear mapping $t:\mathbb{R}\langle\langle X\rangle\rangle\to\mathbb{R}\langle\langle W\rangle\rangle$ is called a **transduction**. It is completely specified by

$$t(\eta)=\sum_{\xi\in W^*}(t(\eta),\xi)\xi,\quad \eta\in X^*.$$

With any t one can canonically associate a series in $\mathbb{R}\langle\langle X\otimes W\rangle\rangle$, namely

$$\hat{t}=\sum_{\eta\in X^*}\eta\otimes t(\eta)=\sum_{\eta\in X^*,\,\xi\in W^*}(t(\eta),\xi)\,\eta\otimes\xi.$$

Given the fact that $\hat{t}$ is still a formal power series in the usual sense, transductions have a well-defined notion of Gevrey degree.

Definition 8: Let $X=\{x_0,x_1,\ldots,x_m\}$ and $W=\{w_0,w_1,\ldots,w_\ell\}$. A transduction $t_c:\mathbb{R}\langle\langle X\rangle\rangle\to\mathbb{R}\langle\langle W\rangle\rangle$ is said to be **associated** with $c\in\mathbb{R}^\ell\langle\langle X\rangle\rangle$ if it can be written as

$$t_c(\eta)=\sum_{\xi\in W^*}(\xi\circ c,\eta)\,\xi,\quad \eta\in X^*.$$

The corresponding series $\hat{t}_c\in\mathbb{R}\langle\langle X\otimes W\rangle\rangle$ is then $\hat{t}_c=\sum_{\xi\in W^*}(\xi\circ c)\otimes\xi$.

Definition 9: For $c\in\mathbb{R}^m\langle\langle X\rangle\rangle$ define the **Fliess signature operator** as

$$\begin{aligned} S_c\ &:\ \mathbb{R}\langle\langle X\rangle\rangle\ \to\ \mathbb{R}\langle\langle W\rangle\rangle\\ &:\ P[U]\ \mapsto\ P[Y]=\sum_{\eta\in X^*}t_c(\eta)(P[U],\eta).\end{aligned}$$

Observe then that, as expected,

$$\begin{aligned} S_c(P[U])&=\sum_{\eta\in X^*,\,\xi\in W^*}\xi\,(\xi\circ c,\eta)E_\eta[U]\\ &=\sum_{\xi\in W^*}\xi\,E_\xi[Y]=P[Y],\end{aligned}$$

where $Y=F_{x_0c}[U]$. It is also clear from the above calculation that

$$E_\xi[Y]=\sum_{\eta\in X^*}(\xi\circ c,\eta)E_\eta[U],\ \ \xi\in W^*. \tag{5}$$

Fliess signature operators as defined above are purely algebraic objects. It is necessary therefore to show that $S_c(P[U])$ is a proper p-rough path when U is a p-rough path. But first one needs to make sense of $E_\xi[Y]$, $\xi\in W^*$, by showing that the series (5) converges. Then it will be proved that $P[Y]=S_c(P[U])$ is a multiplicative functional of order $\lfloor p\rfloor$. These two steps provide the means to show that $E_\xi[Y]$, $\xi\in W^*$, satisfies (3). In order to show that the series defining $E_\xi[Y]$ is well-defined, a characterization of the Gevrey degree of the series $\xi\circ c$ is given in the next theorem.

Theorem 3: Let $c\in\mathbb{R}^\ell_{G(1/p)}\langle\langle X\rangle\rangle$ with growth constants $K_c,M_c>0$ and $p\geq 1$. Then for any $\xi\in W^*$ it follows that $\xi\circ c\in\mathbb{R}_{G(1/p)}\langle\langle X\rangle\rangle$. Specifically,

$$|(\xi\circ c,\theta)|\leq(K_c^kM_c^{-|\xi|})(2^kM_c)^{|\theta|}(|\theta|!)^{\frac{1}{p}}$$

for all $\theta\in X^*$ such that $|\theta|\geq|\xi|$ and where $k=|\xi|-|\xi|_{w_0}$. Otherwise, $(\xi\circ c,\theta)=0$.

Proof: First observe that any $\xi\in W^*$ can be written as $\xi_k=w_0^{n_k}w_{i_k}w_0^{n_{k-1}}w_{i_{k-1}}\cdots w_0^{n_0}$ with each $i_j\neq 0$. The claim above is trivial when $k=0$. For the $k=1$ case, it noted that in general $|(\xi\circ c,\theta)|\leq(\xi\circ\bar{c},\theta)$ for every $\theta\in X^*$, where $(\bar{c},\theta)=K_cM_c^{|\theta|}(|\theta|)^{1/p}$ [8, Lemma 1]. Thus, it follows when $|\theta|\geq|\xi_1|$ that

$$\begin{aligned} &|(\xi_1\circ c,\theta)|\\ &\leq((w_0^{n_1}w_{i_1}w_0^{n_0})\circ\bar{c},\theta)\\ &=(\bar{c}_{i_1}\sqcup\!\sqcup\, x_0^{n_0},x_0^{-(n_1+1)}(\theta))\\ &=\sum_{i=0}^{m_1}\sum_{\substack{\eta\in X^i\\ \xi\in X^{m_1-i}}}(\bar{c}_{i_1},\eta)(x_0^{n_0},\xi)(\eta\sqcup\!\sqcup\,\xi,x_0^{-(n_1+1)}(\theta))\\ &=(K_cM_c^{|\theta|-|\xi_1|})((|\theta|-|\xi_1|)!)^{\frac{1}{p}}\binom{m_1}{n_0}\\ &\leq(K_cM_c^{-|\xi_1|})(2M_c)^{|\theta|}(|\theta|!)^{\frac{1}{p}},\end{aligned}$$

where $m_1=|\theta|-(n_1+1)$. Thus, the inequality is true for $k=1$. Next, assume it is true up to some arbitrary fixed $k\geq 0$. Let $m_{k+1}=|\theta|-(n_{k+1}+1)$ and define $q\geq 1$ to satisfy $1/p+1/q=1$. Observe then for $|\theta|\geq|\xi_{k+1}|$ that

$$\begin{aligned} &|(\xi_{k+1}\circ c,\theta)|\\ &\leq((w_0^{n_{k+1}}w_{i_{k+1}}\xi_k)\circ\bar{c},\theta)\\ &=(\bar{c}_{i_{k+1}}\sqcup\!\sqcup\,(\xi_k\circ\bar{c}),x_0^{-(n_{k+1}+1)}(\theta))\\ &=\sum_{i=0}^{m_{k+1}}\sum_{\substack{\eta\in X^i\\ \xi\in m_{k+1}-i}}(\bar{c}_{i_{k+1}},\eta)(\xi_k\circ\bar{c},\xi)\cdot\\ &\qquad(\eta\sqcup\!\sqcup\,\xi,x_0^{-(n_{k+1}+1)}(\theta))\\ &\leq\sum_{i=0}^{m_{k+1}}K_cM_c^i(i!)^{\frac{1}{p}}(K_c^kM_c^{-|\xi_k|})(2^kM_c)^{m_{k+1}-i}\cdot\\ &\qquad((m_{k+1}-i)!)^{\frac{1}{p}}\binom{m_{k+1}}{i}\\ &=(K_c^{k+1}M_c^{m_{k+1}-|\xi_k|})(2^k)^{m_{k+1}-i}(m_{k+1}!)^{\frac{1}{p}}\sum_{i=0}^{m_{k+1}}\binom{m_{k+1}}{i}^{\frac{1}{q}}\\ &\leq(K_c^{k+1}M_c^{m_{k+1}-|\xi_k|})(2^{k+1})^{m_{k+1}}(m_{k+1}!)^{\frac{1}{p}}\\ &\leq(K_c^{k+1}M_c^{-|\xi_{k+1}|})(2^{k+1}M_c)^{|\theta|}(|\theta|!)^{\frac{1}{p}}.\end{aligned}$$

Thus, the theorem is proved. ■

A straightforward consequence of this theorem is the following corollary, which says that $E_\xi[Y]$, $\xi\in W^*$, is well-defined.

Corollary 1: Let $c\in\mathbb{R}^\ell_{G(1/p')}\langle\langle X\rangle\rangle$. If $U\in G\Omega_p(\mathbb{R}^m)$ with $p\leq p'$ and $Y=F_{x_0c}[U]$ then

$$\|E_\xi[Y]\|_{\mathcal{V}^p(J)}<\infty,\ \ \xi\in W^*.$$

The next lemma is key to showing that $P[Y]$ is a multiplicative functional of order $\lfloor p\rfloor$.

Lemma 3: Let $c\in\mathbb{R}^\ell_{G(1/p')}\langle\langle X\rangle\rangle$. If $U\in G\Omega_p(\mathbb{R}^m)$ with $p\leq p'$ and $Y=F_{x_0c}[U]$ then

$$E_\xi[Y](t_2,t_1)=\sum_{\xi=\xi_1\xi_2}E_{\xi_1}[Y](t_2,\tau)E_{\xi_2}[Y](\tau,t_1), \tag{6}$$

where $\xi \in W^*$ and $t_1 \leq \tau \leq t_2$.

Proof: The proof is by induction on $|\xi| = k$. The claim is trivial for $k = 0$. Suppose (6) holds up to some fixed $k \geq 0$. Then for any $\xi \in W^k$ and $w_i \neq w_0$, it follows that

$$\begin{aligned} &E_{w_i\xi}[Y](t_2,t_1)\\ &= \sum_{\nu\in X^*} ((w_i\xi)\circ c,\nu)E_\nu[U](t_2,t_1)\\ &= \sum_{\nu\in X^*} (x_0(c_i \sqcup\!\sqcup (\xi\circ c)),\nu)E_\nu[U](t_2,t_1)\\ &= \sum_{\nu\in X^*} (c_i \sqcup\!\sqcup (\xi\circ c),\nu)E_{x_0\nu}[U](t_2,t_1)\\ &= \sum_{\nu\in X^*} \int_{t_1}^{t_2} (c_i \sqcup\!\sqcup (\xi\circ c),\nu)E_\nu[U](\tau,t_1)\,d\tau. \end{aligned} \tag{7}$$

Since U is a p-rough path then $\|E_\eta[U]\|_{\mathcal{V}^p} \propto 1/(|\eta|!)^{\frac{1}{p}}$. Therefore, given that $c \in \mathbb{R}_{G(1/p')}\langle\langle X\rangle\rangle$ for $p \leq p'$, the series

$$\begin{aligned} &\sum_{\nu\in X^*} (c_i \sqcup\!\sqcup (\xi\circ c),\nu)E_\nu[U](\tau,t_1)\\ &= \underbrace{\sum_{\nu\in X^*} (c_i,\nu)E_\nu[U](\tau,t_1)}_{F_{c_i}[U](\tau,t_1)} \underbrace{\sum_{\nu\in X^*} (\xi\circ c,\nu)E_\nu[U](\tau,t_1)}_{E_\xi[Y](\tau,t_1)} \end{aligned}$$

is convergent. Thus, it is safe for the integral and summation in (7) to be exchanged, which gives

$$E_{w_i\xi}[Y](t_2,t_1) = \int_{t_1}^{t_2} \sum_{\nu\in X^*} (\xi\circ c,\nu)E_\nu[U](\tau,t_1)\,dY_i(\tau).$$

For any $t \in [t_1,t_2]$ and applying the induction hypothesis,

$$\begin{aligned} &E_{w_i\xi}[Y](t_2,t_1)\\ &= \int_{t_1}^{t} \sum_{\nu\in X^*} (\xi\circ c,\nu)E_\nu[U](\tau,t_1)\,dY_i(\tau)\\ &\quad + \int_{t}^{t_2} \underbrace{\sum_{\nu\in X^*} (\xi\circ c,\nu)E_\nu[U](\tau,t_1)}_{E_\xi[Y](\tau,t_1)}\,dY_i(\tau)\\ &= E_{w_i\xi}[Y](t,t_1)\\ &\quad + \int_{t}^{t_2} \sum_{\xi=\xi_1\xi_2} E_{\xi_1}[Y](\tau,t)E_{\xi_2}[Y](t,t_1)\,dY_i(\tau)\\ &= E_\emptyset[Y](t_2,t)E_{w_i\xi}[Y](t,t_1)\\ &\quad + \sum_{\xi=\xi_1\xi_2} E_{w_i\xi_1}[Y](t_2,t)E_{\xi_2}[Y](t,t_1)\\ &= \sum_{w_i\xi=\xi_1\xi_2} E_{\xi_1}[Y](t_2,t)E_{\xi_2}[Y](t,t_1). \end{aligned}$$

This concludes the proof. ∎

Theorem 4: Let $c \in \mathbb{R}^{\ell}_{G(1/p')}\langle\langle X\rangle\rangle$. If $U \in G\Omega_p(\mathbb{R}^m)$ with $p \leq p'$ and $Y = F_{x_0c}[U]$ then $P_{\lfloor p\rfloor}[Y]$ is a multiplicative functional of order $\lfloor p\rfloor$.

Lemma 4: Let $c \in \mathbb{R}^{\ell}_{G(1/p')}\langle\langle X\rangle\rangle$. If $U \in G\Omega_p(\mathbb{R}^m)$ with $p \leq p'$ and $Y = F_{x_0c}[U]$ then

$$\|E_\xi[Y]\|_{p,[t_1,t_2]} \leq \frac{\omega(t_1,t_2)^{\frac{|\xi|}{p}}}{\left(\frac{|\xi|}{p}\right)!},\ \ 0 \leq t_1 \leq t_2,\ \ \xi \in W^*.$$

It is thus trivial using Theorem 4 and Lemma 4 to conclude that $P[Y]$ is indeed a p-rough path. Furthermore, the next lemma gives a bound for the summation of iterated integrals of the same order with arbitrary p-variation inputs.

Lemma 5: Let $c \in \mathbb{R}^{\ell}_{G(1/p')}\langle\langle X\rangle\rangle$. If $U \in G\Omega_p(\mathbb{R}^m)$ with $p \leq p'$ and $Y = F_{x_0c}[U]$ then for any $k \geq 1$

$$\left\|\sum_{\xi\in W^k} E_\xi[Y]\right\|_{p,[t_1,t_2]} \leq (m+1)^{k-1}\frac{\omega(t_1,t_2)^{\frac{k}{p}}}{k!},\ \ 0 \leq t_1 \leq t_2.$$

Proof: First observe that

$$\begin{aligned} \sum_{\xi\in W^k} E_\xi[Y] &= \sum_{\eta\in X^*} E_\eta[U] \sum_{\xi\in W^k} (\xi\circ c,\eta)\\ &= \sum_{\eta\in X^*} E_\eta[U]\left(\mathrm{char}\left(W^k\right)\circ c,\eta\right). \end{aligned}$$

It is not difficult to show that the composition product satisfies $d_1^{\sqcup\!\sqcup\, j}\circ d_2 = (d_1\circ d_2)^{\sqcup\!\sqcup\, j}$ for any $d_1,d_2 \in \mathbb{R}\langle\langle X\rangle\rangle$. The fact that $\mathrm{char}\left(W^k\right) = \mathrm{char}(W)^{\sqcup\!\sqcup\, k}/\,k!$ gives

$$\begin{aligned} &\sum_{\xi\in W^k} E_\xi[Y]\\ &= \frac{1}{k!}\sum_{\eta\in X^*} E_\eta[U]\left((\mathrm{char}\,(W)\circ c)^{\sqcup\!\sqcup\, k},\eta\right)\\ &= \frac{1}{k!}\sum_{\eta\in X^*} E_\eta[U]\left(\left(x_0\left(1+\sum_{i=1}^m c_i\right)\right)^{\sqcup\!\sqcup\, k},\eta\right)\\ &= \frac{1}{k!}\left(\left(x_0\left(1+\sum_{i=1}^m c_i\right)\right)^{\sqcup\!\sqcup\, k},P[U]\right)\\ &= \frac{1}{k!}\,F_{\left(x_0\left(1+\sum_{i=1}^m c_i\right)\right)^{\sqcup\!\sqcup\, k}}[U]\\ &= \frac{1}{k!}\,\left(F_{x_0\left(1+\sum_{i=1}^m c_i\right)}[U]\right)^k\\ &= \frac{1}{k!}\,(Y_0+\cdots+Y_m)^k, \end{aligned}$$

where Y_i is the i-th component of the output path obtained as a smooth approximation from $\Omega_1(\mathbb{R})$ to $G\Omega_p(\mathbb{R})$ via integration of $F_c[U]$. From Jensen's inequality, one has for any convex function φ and scalars $a_i > 0$ that

$$\varphi\left(\sum_i \frac{a_i\psi_i}{a_i}\right) \leq \left(\frac{\sum_i a_i\varphi(\psi_i)}{\sum_j a_j}\right).$$

Hence, $(Y_0+\cdots+Y_m)^k \leq (m+1)^{k-1}(|Y_0|^k+\cdots+|Y_m|^k)$. Finally, it follows that

$$\left\|\sum_{\xi\in W^k} E_\xi[Y]\right\|_{p,[s,t]} = \left\|\frac{1}{k!}\,(Y_0+\cdots+Y_m)^k\right\|_{p,[s,t]}$$

$$\leq \frac{(m+1)^{k-1}\omega(s,t)^{\frac{k}{p}}}{k!},$$

where $w(s,t) \triangleq \sum_{i=0}^{m} w_i(s,t)$ with ω_i being the ω-function for Y_i. ■

Finally, the main result of the section is given below.

Theorem 5: Let $c \in \mathbb{R}^{\ell}_{G(1/p')}\langle\langle X\rangle\rangle$. If $U \in G\Omega_p(\mathbb{R}^m)$ with $p \leq p'$ then the Fliess signature operator S_c maps $G\Omega_p(\mathbb{R}^m)$ into $G\Omega_p(\mathbb{R}^\ell)$.

Proof: As a consequence of Lemma 3, $P[Y]$ is a multiplicative functional of order $\lfloor p\rfloor$ with bounded iterated integrals up to order $\lfloor p\rfloor$, which means that it is a p-rough path. The extension theorem says then that $P[Y]$ can be extended to a multiplicative functional satisfying (3) for arbitrary order. Moreover, a Fliess signature operator is the p-variation continuous extension of $S_c : \Omega_1(\mathbb{R}^m) \to \Omega_1(\mathbb{R}^\ell)$ into $S_c : G\Omega_p(\mathbb{R}^m) \to G\Omega_p(\mathbb{R}^\ell)$ since $\Omega_1(\mathbb{R}^m)$ is dense in $G\Omega_p(\mathbb{R}^m)$. This follows from the fact that Y is obtained by a smooth approximation, and that $E_\xi[Y]$ is continuous in the p-variation metric. ■

IV. Cascade Interconnections with Rough Path Inputs

In light of Theorem 5, the output of a Fliess signature operator driven by a p-rough path input U can be fed into a second Fliess signature operator as long as its generating series is in $\mathbb{R}^m_{G(1/p)}\langle\langle X\rangle\rangle$. The cascade is best described algebraically by a composition of transductions.

Definition 10: Let $X = \{x_0, x_1, \ldots, x_m\}$. For $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ and $d \in \mathbb{R}^m\langle\langle X\rangle\rangle$, the **composition of transductions** t_c and t_d is

$$t_c \circ t_d = \sum_{\eta\in X^*}\sum_{\xi\in W^*} (\xi\circ c, \eta)\, \eta\circ d\otimes\xi.$$

Lemma 6: For transductions t_c and t_d it follows that $t_c \circ t_d = t_{c\circ d}$.

Proof: Using the associativity of the composition product, observe that

$$\begin{aligned} t_c\circ t_d &= \sum_{\eta\in X^*}\sum_{\xi\in W^*} (\xi\circ c,\eta)\,\eta\circ d\otimes\xi \\ &= \sum_{\xi\in W^*}\sum_{\eta\in X^*} (\xi\circ c,\eta)(\eta\circ d)\otimes\xi \\ &= \sum_{\xi\in W^*} ((\xi\circ c)\circ d)\otimes\xi \\ &= \sum_{\xi\in W^*} (\xi\circ(c\circ d))\otimes\xi = t_{c\circ d}. \end{aligned}$$

■

The relationship between the composition of transductions and the cascade of two Fliess signature operators is established next.

Theorem 6: The cascade of two Fliess signature operators S_c and S_d is described by $S_c \circ S_d = S_{c\circ d}$.

Proof: Note that $S_c \circ S_d = (t_c, (t_d, P[U]))$. Therefore, it is only necessary to show that $(t_c, (t_d, P[U])) = (t_{c\circ d}, P[U])$ holds. That is,

$$\begin{aligned} (t_c,(t_d,P[U])) &= \sum_{\eta\in X^*} (t_c,\eta)E_\eta[F_d[U]] \\ &= \sum_{\xi\in W^*,\,\eta\in X^*} \eta\,(\xi\circ c,\eta)\, E_\eta[F_d[U]] \\ &= \sum_{\xi\in W^*} \xi\, F_{\xi\circ c}[F_d[U]] \\ &= \sum_{\xi\in W^*} \xi\, E_\xi[F_{c\circ d}[U]] \\ &= (t_{c\circ d}, P[U]). \end{aligned}$$

The claim then follows trivially since

$$\begin{aligned} S_c\circ S_d(P[U]) &= (t_c,(t_d,P[U])) \\ &= (t_{c\circ d}, P[U]) \\ &= S_{c\circ d}(P[U]). \end{aligned}$$

■

V. Conclusions

A treatment of Fliess operators driven by geometric p-rough paths was presented. It was shown that the signature of the input path is mapped to the signature of the output path via a transduction constructed in terms of the composition product. Introducing rough paths solved the open problem of how the cascade of these operators can be well-posed when driven by input signals with p-variation larger than 1. In essence, the stochastic setting was dropped and stochastic processes were characterized as p-rough paths. Transductions also facilitated the computation of the signature for the cascade of two Fliess signature operators.

References

[1] K. T. Chen, Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula, *Ann. Math.* **65** (1957) 163-178.

[2] L. A. Duffaut Espinosa, *Interconnections of Nonlinear Systems Driven by L_2-Itô Stochastic Processes*, Ph.D. Dissertation, Old Dominion University, 2009.

[3] L. A. Duffaut Espinosa, W. S. Gray, and O. R. González, On the local convergence of Fliess operators driven by L_2-Itô processes, *Proc. 49th IEEE Conf. on Decision and Control*, Atlanta, Georgia, 2010, pp. 2631-2637.

[4] M. Fliess, Transductions de séries formelles, *Discrete Math.* **10** (1974) 57-74.

[5] M. Fliess, Fonctionnelles causales non linéaires et indéterminées non commutatives, *Bull. Soc. Math. France* **109** (1981) 3-40.

[6] P. Friz and N. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, Cambridge University Press, London, 2010.

[7] W. S. Gray and Y. Li, Generating series for interconnected analytic nonlinear systems, *SIAM J. Control Optim.* **44** (2005) 646-672.

[8] W. S. Gray and M. Thitsa, On the radius of convergence of self-excited feedback connected analytic nonlinear systems, *Proc. 49th IEEE Conf. on Decision and Control*, Atlanta, Georgia, 2010, pp. 7092-7098.

[9] W. S. Gray and Y. Wang, Fliess operators on L_p spaces: convergence and continuity, *Systems Control Lett.* **46** (2002) 67-74.

[10] K. Hara and M. Hino, Fractional order Taylor's series and the neo-classical inequality, *Bull. Lond. Math. Soc.* **42** (2010) 467-477.

[11] G. Jacob, Sur un théorème de Shamir, *Inform. and Control* **27** (1975) 218-261.

[12] T. Lyons, M. Caruana and T. Lévy, Differential equations driven by rough paths, in *École d'Eté de Probabilités de Saint-Flour XXXIV*, ser. Lecture Notes in Mathematics, J. Morel, F. Takens, and B. Teisser, Eds. Springer-Verlag, Berlin, 2004, vol. 1908, 1-93.

[13] T. Lyons and Z. Qian, *System Control and Rough Paths*, Oxford University Press, Oxford, 2002.

[14] Y. Wang, *Algebraic Differential Equations and Nonlinear Control Systems*, Ph.D. Dissertation, Rutgers University, 1990.