Electron. Commun. Probab. **20** (2015), no. 8, 1–11.
DOI: 10.1214/ECP.v20-3636
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Concentration and exact convergence rates for expected Brownian signatures

Hao Ni* Weijun Xu†

Abstract

The signature of a d-dimensional Brownian motion is a sequence of iterated Stratonovich integrals along the Brownian paths, an object taking values in the tensor algebra over $\mathbb{R}^d$. In this article, we derive the exact rate of convergence for the expected signatures of piecewise linear approximations to Brownian motion. The computation is based on the identification of the set of words whose coefficients are of the leading order, and the convergence is concentrated on this subset of words. Moreover, under the choice of l^1 tensor norm, we give the explicit value of the leading term constant.

Keywords: expected signature; Brownian motion.
AMS MSC 2010: 60G15.
Submitted to ECP on July 1, 2014, final version accepted on January 24, 2015.

1 Introduction

Let $(e_1, \cdots, e_d)$ be the standard basis of $\mathbb{R}^d$, $d \geq 2$, and let

$$B_t = \sum_{j=1}^{d} B_t^j e_j,$$

where B_t^j's are independent standard one dimensional Brownian motions. The signature of B is a sequence of Stratonovich iterated integrals along the sample paths ([7], [8]). We give the precise definition below.

Definition 1.1. For every $n \geq 1$ and every word $w = e_{i_1} \cdots e_{i_n}$ with length n, define

$$C_{s,t}^{w} = \int_{s<u_1<\cdots<u_n<t} \circ dB_{u_1}^{i_1} \cdots \circ dB_{u_n}^{i_n} \tag{1.1}$$

in the sense of Stratonovich integral. For each $n \geq 0$, let

$$X_{s,t}^{n}(B) = \sum_{|w|=n} C_{s,t}^{w} w,$$

where the sum is taken over all words of length n. We use the convention $C_{s,t}^{w} \equiv 1$ if w is the empty word. Then, the infinite sequence

$$X_{s,t}(B) = \big(1, X_{s,t}^{1}(B), \cdots, X_{s,t}^{n}(B), \cdots\big)$$

is the (Stratonovich) signature of B over time interval $[s,t]$.

*University of Oxford, UK. E-mail: hao.ni@maths.ox.ac.uk
†University of Warwick, UK. E-mail: weijun.xu@warwick.ac.uk

Remark 1.2. It is sometimes more convenient to write the signatures in terms of tensors, i.e.,

$$X^n_{s,t} = \int_{s<u_1<\cdots u_n<t} \circ dB_{u_1} \otimes \cdots \otimes \circ dB_{u_n},$$

and $C^w_{s,t}$ defined in (1.1) is the coefficient of w in X. This is equivalent to Definition 1.1.

The study of the signature of a path dates back to K.T.-Chen in 1950's. In a series of papers ([1], [2], [3]), he developed algebraic properties of these multiple iterated integrals, and showed that piecewise smooth paths are characterized by their iterated path integrals over a fixed time interval. Hambly and Lyons ([8]) gave a quantitative version of this result, and extended it to all paths of bounded variation. They showed that, paths of bounded variation in $\mathbb{R}^d$ are uniquely determined by their signatures up to tree-like equivalence.

These iterated integrals also play a fundamental role in rough paths, where Lyons ([9]) used them to develop an integration theory along paths of any regularity.

As for random paths, the expected signature is an important object to study as it determines the law of compactly supported measure on path space, and this is anticipated to be true for more general stochastic processes (see [4] for a recent proof for processes under certain integrability conditions), the foremost example being Brownian motion. The computation of the expected signature of Brownian motion also leads to cubature on Wiener space ([11]).

The expected signature for Brownian motion was first derived by Fawcett ([5]), and then independently by Lyons and Victoir ([11]). In this note, we show that the expected signature of piecewise linear approximation to Brownian motion with mesh size $\frac{1}{M}$ converges to that of Brownian motion with rate $\frac{1}{M}$. Moreover, under the choice of l^1 tensor norm, we give the explicit value of the leading term constant.

More precisely, let $B^{(M)}$ denote the piecewise linear approximation to Brownian motion with mesh size $\frac{1}{M}$. Let

$$\phi(T) = \mathbb{E}X_{0,T}(B), \qquad \phi^M(T) = \mathbb{E}X_{0,T}(B^{(M)}),$$

where the expectation is taken for each component, and our main theorem is then the following.

Theorem 1.3. For each $n \geq 0$, let π_n denote the projection from the tensor algebra to $(\mathbb{R}^d)^{\otimes n}$. Then,

(i) $\pi_2(\phi(T)) = \pi_2(\phi^M(T))$, and $\pi_{2n-1}(\phi(T)) = \pi_{2n-1}(\phi^M(T)) = 0$ for all $n \geq 1$.

(ii) For each $n \geq 2$, if $\mathbb{R}^d$ is endowed with the l^1 norm, and $(\mathbb{R}^d)^{\otimes 2n}$ is given the projective tensor norm (to be defined in the next section), then

$$\lim_{M\to+\infty} \frac{M}{T} \left\| \pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T)) \right\| = \frac{d-1}{3\cdot(n-2)!} \left(\frac{dT}{2}\right)^{n-1}. \tag{1.2}$$

The first part of the theorem is an immediate consequence of the basic properties of $\phi(T)$ and $\phi^M(T)$, which we will establish in section 2 below. The proof of the second claim is more involved. The core part of the proof is to identify for each n the words whose coefficients are of order $\frac{1}{M}$, which turns out to be a rather small subset of words of length $2n$. The coefficients of all other words are of order $\mathcal{O}(\frac{1}{M^2})$. That is to say,

$\left\|\pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T))\right\|$ is concentrated on this small subset. We will give precise meaning in section 4 below.

It should be noted that the exact value of the right hand side of (1.2) depends on the choice of tensor norm and the equally spaced piecewise linear approximation. However, the concentration described above is due to the intrinsic nature of Brownian signatures, and remains unchanged under different tensor norms.

Following the same line of argument, one can generalize the above statement to the form below.

Theorem 1.4. If $\|\cdot\|$ is endowed with l^1 norm on $(\mathbb{R}^d)^{\otimes 2n}$, then we have

$$\left\|\pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T))\right\| = \sum_{k=1}^{n-1} C_{k,n}\left(\frac{T}{M}\right)^k, \tag{1.3}$$

where the constants $C_{k,n}$ depends on k, n and the dimension, but are independent of M and T. In particular, $C_{1,n}$ is given by Theorem 1.3 above.

Remark 1.5. Note that the sum in the right hand side of (1.3) stops at $n-1$. This fact depends crucially on the choice of l^1 norm as in that case, the norm is simply given by the sum of all components, all of which are polynomials in $\frac{T}{M}$ up to order $n-1$. This in general is not true for other norms, where one necessarily gets an infinite sequence.

We will mainly focus on the proof of Theorem 1.3, and Theorem 1.4 will become an easy consequence of that. Before we proceed, we first give a brief introduction to tensor norms.

A note on tensor norms

For each $n \geq 1$, the n-tensor space $(\mathbb{R}^d)^{\otimes n}$ is a real vector space with basis

$$\{e_{i_1} \cdots e_{i_n} : 1 \leq i_1, \cdots, i_n \leq d\},$$

where the e_j's are standard basis of $\mathbb{R}^d$. The tensor algebra over $\mathbb{R}^d$ is then the direct sum

$$T(\mathbb{R}^d) := \mathbb{R} \oplus \mathbb{R}^d \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes n} \oplus \cdots.$$

Although it is common to identify $(\mathbb{R}^d)^{\otimes n}$ with $\mathbb{R}^{d^n}$, which gives the Hilbert Schmidt norm, in many cases, some other norms appear more useful. Throughtout this paper, we will use the l^1 norm on each $(\mathbb{R}^d)^{\otimes n}$, defined by

$$\|v\| = \sum_i |a_i|$$

if $v \in (\mathbb{R}^d)^{\otimes n}$ can be written as a linear combination of basis elements $\{v_i\}$ by $v = \sum_i a_i v_i$.

It should be noted that when $\mathbb{R}^d$ is endowed with the l^1 norm, then the l^1 and projective norms on $(\mathbb{R}^d)^{\otimes n}$ coincide for every n. But in general, the projective norm is more complicated if $\mathbb{R}^d$ is endowed with other norms. For more details of projective tensor norms and some of their significance, we refer to the paper [8].

Our paper is organized as follows. In section 2, we give some formulae and basic properties of the expected signatures of Brownian motion and its piecewise linear approximations. Section 3 is devoted to the proof of Theorem 1.3. In section 4, we briefly explain how our arguments lead to Theorem 1.4.

Notations. In the rest of the paper, $\|\cdot\|_n$ will denote the projective tensor norm on $(\mathbb{R}^d)^{\otimes n}$. We will omit the subscript n and simply write $\|\cdot\|$ if no confusion may arise. We use π_n to denote the projection from $T(\mathbb{R}^d)$ onto $(\mathbb{R}^d)^{\otimes n}$. Also, if $x \in T(\mathbb{R}^d)$, and w is a word, then $C^w(x)$ will denote the coefficient of w in x. Finally, for fixed T and M, we write $\Delta t = \frac{T}{M}$.

2 The expected signatures of Brownian motion and its piecewise linear approximations

In this part, we give some formulae and propositions of $\phi(T)$ and $\phi^M(T)$. We first introduce some notations. For any word w, let $N_i(w)$ denote the number of occurrences of the letter e_i in w. For each $n \geq 0$, let

$$\mathcal{S}_{2n} = \{w : w = e_{i_1}^2 \cdots e_{i_n}^2, 1 \leq i_1, \cdots, i_n \leq d\},$$

and

$$\mathcal{K}_{2n} = \{w : |w| = 2n, \;\; N_i(w) \;\text{ is even for all } i\}.$$

The following formula for $\phi(T)$ was proven by Fawcett in [5] as well as by Lyons and Victoir in [11].

Proposition 2.1. Let B be a d-dimensional Brownian motion. Then,

$$\phi(T) = \mathbb{E}[X_{0,T}(B)] = \exp\left[\frac{T}{2}\sum_{j=1}^{d} e_j \otimes e_j\right].$$

It is immediate from the proposition that if $w \in \mathcal{S}_{2n}$ for some n, then

$$C^w(\phi(T)) = \frac{1}{n!}\left(\frac{T}{2}\right)^n, \tag{2.1}$$

and $C^w(\phi(T)) = 0$ for all other w's.

Lemma 2.2. Fix an arbitrary $n \in \mathbb{N}$. If $w \in \mathcal{K}_{2n}$ such that $N_k(w) = 2m_k$ for $k = 1, \cdots, d$, then for each $t \geq 0$, we have

$$C^w(\phi^1(t)) = \frac{\lambda_w}{n!}\left(\frac{t}{2}\right)^n,$$

where $\lambda_w = \binom{n}{m_1, \cdots, m_d} \Big/ \binom{2n}{2m_1, \cdots, 2m_d} \leq 1$. On the other hand, $C^w(\phi^1(t)) = 0$ for all $t \geq 0$ and all words w that do not belong to any of the $\mathcal{K}_{2n}$'s.

Proof. If $\gamma = (\gamma^1, \cdots, \gamma^d)$ is a straight line, and $w = e_{j_1} \cdots e_{j_k}$, then

$$C^w(X_{0,t}(\gamma)) = \frac{1}{k!}\gamma^{j_1}(t) \cdots \gamma^{j_k}(t).$$

Taking expectation of both sides gives

$$C^w(\phi^1(t)) = \frac{1}{k!}(\mathbb{E}(B_t^1)^{p_1})\cdots(\mathbb{E}(B_t^d)^{p_d}),$$

where p_l is the number of occurrences of the letter e_l in w. It is then clear that $C^w(\phi^1(t)) = 0$ if any of the p_l's is odd. For $w \in \mathcal{K}_{2n}$, let $2m_k$ be the number of occurrences of e_k, then

$$C^w(\phi^1(t)) = \frac{1}{(2n)!}(\mathbb{E}(B_t^1)^{2m_1})\cdots(\mathbb{E}(B_t^d)^{2m_d}),$$

and the conclusion of the lemma follows from Gaussian moments.

□

Corollary 2.3. For any $w \in \mathcal{S}_{2n}$, we have

$$C^w(\phi^M(T)) \leq C^w(\phi(T)).$$

Proof. By the expression, (2.1), it suffices to show $C^w(\phi^M(T)) \leq \frac{1}{n!}\left(\frac{T}{2}\right)^n$. Since the increments of Brownian motion are independent, we have $\phi^M(T) = \phi^1(\Delta t)^{\otimes M}$. This implies

$$C^w(\phi^M(T)) = \sum C^{v_1}(\phi^1(\Delta t))\cdots C^{v_M}(\phi^1(\Delta t)),$$

where $\Delta t = \frac{T}{M}$, and the sum is taken over all $v_1 * \cdots * v_M$ such that each v_j is in $\mathcal{S}_{2k}$ for some k. By Lemma 2.2, we have

$$\begin{aligned} C^w(\phi^M(T)) &\leq \left(\frac{\Delta t}{2}\right)^n \frac{1}{n!} \sum_{k_1+\cdots+k_M=n} \binom{n}{k_1,\cdots,k_M} \\ &= \frac{1}{n!}\left(\frac{T}{2}\right)^n, \end{aligned}$$

where we have used the fact that $\lambda_{v_j} \leq 1$, and each v_j has even length. □

Lemma 2.4. For each $n, M \in \mathbb{N}$ and $T \geq 0$, we have

$$\|\pi_{2n}(\phi(T))\| = \left\|\pi_{2n}(\phi^M(T))\right\| = \frac{1}{n!}\cdot\left(\frac{dT}{2}\right)^n.$$

Proof. That $\|\pi_{2n}(\phi(T))\| = \frac{1}{n!}\cdot\left(\frac{dT}{2}\right)^n$ is immediate from Proposition 2.1. In order the prove the second one, we note that

$$\left\|\pi_{2n}(\phi^1(t))\right\| = \frac{1}{n!}\left(\frac{dt}{2}\right)^n \tag{2.2}$$

for all n and t. By independent increments of Brownian motion, we have

$$\pi_{2n}(\phi^M(T)) = \sum_{k_1+\cdots+k_M=n} \pi_{2k_1}(\phi^1(\Delta t))\otimes\cdots\otimes\pi_{2k_M}(\phi^1(\Delta t)).$$

By properties of the projective norm and the positivity of all entries, we can change the sum with the norm $\|\cdot\|$, and get

$$\left\|\pi_{2n}(\phi^M(T))\right\| = \sum_{k_1+\cdots+k_M=n} \left\|\pi_{2k_1}(\phi^1(\Delta t))\right\| \cdots \left\|\pi_{2k_M}(\phi^1(\Delta t))\right\|.$$

By (2.2) and the multinomial theorem, we get

$$\left\|\pi_{2n}(\phi^M(T))\right\| = \frac{1}{n!}\left(\frac{dT}{2}\right)^n,$$

thus proving the lemma. □

Note that the above lemma is true only for the l^1 norm. For Hilbert Schmidt norm, we have $\|\pi_{2n}(\phi(T))\| > \left\|\pi_{2n}(\phi^M(T))\right\|$. The next proposition will be very useful for proving the main theorem. It is an immediate consequence of the previous lemma.

Proposition 2.5. $\left\|\pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T))\right\| = 2\sum_{w\in\mathcal{K}_{2n}\setminus\mathcal{S}_{2n}} C^w(\phi^M(T))$.

Proof. By Corollary 2.3, we have

$$\begin{aligned}&\left\|\pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T))\right\| \\ =& \sum_{w\in\mathcal{K}_{2n}\setminus\mathcal{S}_{2n}} C^w(\phi^M(T)) + \sum_{w\in\mathcal{S}_{2n}} [C^w(\phi(T)) - C^w(\phi^M(T))],\end{aligned}$$

where we have used the fact that all the $C^w(\phi^M(T))$'s are non-negative. Also, Lemma 2.4 implies that the two terms on the right hand side are equal. Thus, we arrive at the conclusion of the proposition. □

3 Proof of Theorems 1.3 and 1.4

This section is mainly devoted to the proof of Theorem 1.3, and we will explain how one can get Theorem 1.4 as a corollary. The first part of Theorem 1.3 is an immediate consequence of Proposition 2.1 and Lemma 2.2. To prove the second part, we need a more detailed study of the coefficients of words in $\mathcal{K}_{2n}$. By Proposition 2.5, it suffices to consider the words in $\mathcal{K}_{2n}\setminus\mathcal{S}_{2n}$. Let

$$\mathcal{E} = \{e_ie_je_ie_j, e_ie_je_je_i : 1 \le i,j \le d, i \ne j\}.$$

For each $k = 0, 1, \cdots, n-2$, define

$$\mathcal{W}_{2n}^k = \{v * v' * v'' : v \in \mathcal{S}_{2k}, v' \in \mathcal{E}, v'' \in \mathcal{S}_{2n-4-2k}\},$$

and let

$$\mathcal{W}_{2n} := \bigcup_{k=1}^{n-2} \mathcal{W}_{2n}^k.$$

Then $\mathcal{W}_{2n} \subset \mathcal{K}_{2n}\setminus\mathcal{S}_{2n}$. We will show that for each n, the set of words whose coefficients are of order $\frac{1}{M}$ is precisely $\mathcal{W}_{2n}\cup\mathcal{S}_{2n}$. We then compute the sum of the coefficients (with absolute values) in $\mathcal{W}_{2n}$, and those in $\mathcal{S}_{2n}$ will be obtained by symmetry. We now study the coefficients of words in $\mathcal{K}_{2n}\setminus(\mathcal{S}_{2n}\cup\mathcal{W}_{2n})$ and in $\mathcal{W}_{2n}$, respectively.

3.1 Words with negligible coefficients

The purpose of this part is to show that for each n, there exists a constant $C = C(d,n)$ such that

$$\sum_{w\in\mathcal{K}_{2n}\setminus(\mathcal{S}_{2n}\cup\mathcal{W}_{2n})} C^w(\phi^M(T)) < \frac{CT^n}{M^2} \tag{3.1}$$

for all large M. For $w \in \mathcal{K}_{2n}$ with $w = e_{i_1}e_{i_2}\cdots e_{i_{2n-1}}e_{i_{2n}}$, let

$$p(w) = |\{k : i_{2k-1} \neq i_{2k}\}|.$$

In other words, $p(w)$ counts the number of non-square pairs in the word w. For each $k = 0, \cdots, n$, define

$$\mathcal{P}^k_{2n} = \{w \in \mathcal{K}_{2n} : p(w) = k\}.$$

It is clear that $\mathcal{P}^0_{2n} = \mathcal{S}_{2n}$, $\mathcal{P}^1_{2n}$ is empty, $\mathcal{W}_{2n} \subset \mathcal{P}^2_{2n}$, and

$$\mathcal{K}_{2n} = \bigcup_{k=0}^{n} \mathcal{P}^k_{2n}$$

as a disjoint union. We will now show that for any $w \in \mathcal{P}^k_{2n}$, we have

$$C^w(\phi^M(T)) < \frac{CT^n}{M^{\lfloor (k+1)/2\rfloor}}. \tag{3.2}$$

We first consider the case $k = 2$. If $w \in \mathcal{P}^2_{2n}$, then it can be expressed as

$$w = \cdots e_ie_j\cdots e_ie_j\cdots, \qquad \text{or} \qquad w = \cdots e_ie_j\cdots e_je_i\cdots,$$

where $i \neq j$, and all other pairs are squares. Without loss of generality, we can assume w has the form

$$w = e^2_{i_1}\cdots e^2_{i_a}\underbrace{e_ie_j * u' * e_ie_j}_{u}e^2_{j_1}\cdots e^2_{j_b},$$

where $u' \in \mathcal{S}_{2r}, r \geq 0$, and $a+b+r = n-2$. Let $u = e_ie_j*u'*e_ie_j$. Since $\phi^M(T) = \phi^1(\Delta t)^{\otimes M}$, we have

$$C^w(\phi^M(T)) = \sum C^{v_1}(\phi^1(\Delta t))\cdots C^{v_M}(\phi^1(\Delta t)), \tag{3.3}$$

where the sum is taken over the collection of words $(v_1, \cdots, v_M)$ such that (i) $v_1 * \cdots * v_M = w$, and (ii) for each j, either $v_j \in \mathcal{S}_{2l}$ for some $l \geq 0$, or $v_j = v' * u * v''$, where $v' \in \mathcal{S}_{2a'}, v'' \in \mathcal{S}_{2b'}$ for some $a', b' \geq 0$[1]. The idea is that the two non-square terms must be grouped together (along with any squares between these two pairs, if they exist) in order for the product on the right hand side of (3.3) not being zero. This will give at most $n-1$ 'atoms' in the decomposition, and the total number of the elements in the sum will be $\mathcal{O}(M^{n-1})$.

More precisely, by Lemma 2.2, for each decomposition $(v_1, \cdots, v_M)$ in the sum, we have

$$C^{v_1}(\phi^1(\Delta t))\cdots C^{v_M}(\phi^1(\Delta t)) \leq \left(\frac{\Delta t}{2}\right)^{a+b+r+2} < (\Delta t)^n, \tag{3.4}$$

[1] Condition (ii) guarantees that every term in the sum is positive. In fact, by Lemma 2.2, if $(v_1, \cdots, v_M)$ satisfies condition (i) but not (ii), then we will have

$$C^{v_1}(\phi^1(\Delta t))\cdots C^{v_M}(\phi^1(\Delta t)) = 0.$$

and we can bound $C^w(\phi^M(T))$ by counting the number of elements in the sum on the right hand side of (3.3). This is exactly the number of nonnegative integer solutions to

$$x_1 + \cdots + x_M = a + b + 1,$$

which equals

$$\binom{M+a+b}{M-1} = \binom{M+n-2-r}{n-1-r} < (M+n)^{n-1-r}. \tag{3.5}$$

Combining the above bound with (3.4), we have

$$C^w(\phi^M(T)) < [(M+n)\Delta t]^{n-1-r}(\Delta t)^{r+1} < \left(\frac{T+n\Delta t}{2}\right)^n \cdot \frac{1}{M^{r+1}},$$

and this is true for all $w \in \mathcal{P}_{2n}^2$. Now, if $w \in \mathcal{P}_{2n}^2 \setminus \mathcal{W}_{2n}$, then $r \geq 1$, and

$$C^w(\phi^M(T)) < \frac{CT^n}{M^2}.$$

The argument for $k \geq 3$ is similar. In order to produce more 'atoms', the best possible choice is to group the consecutive two non-square pairs together, and in the case of odd k, one atom should contain three non-square pairs[2]. Below are two figures for even and odd k's, respectively.

$$k \text{ even}: \qquad \cdots \underbrace{e_{i_1}e_{i_2}\cdots e_{i_3}e_{i_4}}_{u_1}\cdots\cdots \underbrace{e_{i_{k-3}}e_{i_{k-2}}\cdots e_{i_{k-1}}e_{i_k}}_{u_{\frac{k}{2}}}\cdots$$

$$k \text{ odd}: \qquad \cdots \underbrace{e_{i_1}e_{i_2}\cdots e_{i_3}e_{i_4}\cdots e_{i_5}e_{i_6}}_{u_1}\cdots\cdots \underbrace{e_{i_{k-3}}e_{i_{k-2}}\cdots e_{i_{k-1}}e_{i_k}}_{u_{\frac{k-1}{2}}}\cdots$$

As we can see, this will give at most $n - \left\lfloor \frac{k+1}{2} \right\rfloor$ 'atoms' in the decompositions. Thus, by the same computation of the number of elements for such decompositions, we can show that

$$C^w(\phi^M(T)) < \frac{CT^n}{M^{\lfloor (k+1)/2 \rfloor}}$$

for all $w \in \mathcal{P}_{2n}^k$ with $k \geq 3$, where C depends on n only. Since

$$\mathcal{K}_{2n} \setminus (\mathcal{S}_{2n} \cup \mathcal{W}_{2n})^c = (\mathcal{P}_{2n}^2 \setminus \mathcal{W}_{2n}) \cup \mathcal{P}_{2n}^3 \cup \cdots \cup \mathcal{P}_{2n}^n,$$

and note that the number of elements in $\mathcal{K}_{2n} \setminus (\mathcal{S}_{2n} \cup \mathcal{W}_{2n})^c$ depends on d and n only, we conclude (3.1) with a constant $C = C(d, n)$.

3.2 Words in $\mathcal{W}_{2n}$

Fix $0 \leq k \leq n-2$ and $w_k \in \mathcal{W}_{2n}^k$, then

$$w_k = e_{i_1}^2 \cdots e_{i_k}^2 * u * e_{j_1}^2 \cdots e_{j_{n-2-k}}^2,$$

where $u \in \mathcal{E}$ as defined at the beginning of this section. Similar as before, we have

$$C^{w_k}(\phi^M(T)) = \sum_{\mathcal{X}(w_k)} C^{v_k^1}(\phi^1(\Delta t)) \cdots C^{v_k^M}(\phi^1(\Delta t)),$$

[2] For example, the three pairs are e_1e_2, e_2e_3 and e_3e_1.

where $\mathcal{X}(w_k)$ is the set of words $(v_k^1, \cdots, v_k^M)$ such that (i) $v_k^1 * \cdots * v_k^M = w$, and (ii) for each j, either $v_j \in \mathcal{S}_{2l}$ for some $l \geq 0$, or $v_j = u' * u * u''$, where $u' \in \mathcal{S}_{2a}, u'' \in \mathcal{S}_{2b}$ for some $a, b \geq 0$.

Intuitively, when M is large, most contributions to the sum come from the decompositions (or more precisely, allocations) with the further restriction that u and each single square are located in different v_j's. More precisely, let

$$\mathcal{X}'(w_k) := \left\{v_k^1 * \cdots * v_k^M = w : \text{ for each } j \leq M, \; v_k^j = u \text{ or } e_l^2 \text{ for some } l\right\}.$$

Then, $\mathcal{X}'(w_k) \subset \mathcal{X}(w_k)$, and

$$|\mathcal{X}'(w_k)| = \binom{M}{n-1}.$$

Their difference is

$$|\mathcal{X}(w_k) \setminus \mathcal{X}'(w_k)| = \binom{M+n-2}{n-1} - \binom{M}{n-1} = \mathcal{O}(M^{n-2}).$$

Also, for each $(v_k^1, \cdots, v_k^M) \in \mathcal{X}(w_k) \setminus \mathcal{X}'(w_k)$, we have

$$C^{v_k^1}(\phi^1(\Delta t)) \cdots C^{v_k^M}(\phi^1(\Delta t)) \leq \left(\frac{\Delta t}{2}\right)^n, \tag{3.6}$$

and thus

$$\sum_{\mathcal{X}(w_k) \setminus \mathcal{X}'(w_k)} C^{v_k^1}(\phi^1(\Delta t)) \cdots C^{v_k^M}(\phi^1(\Delta t)) = \mathcal{O}\left(\frac{1}{M^2}\right).$$

On the other hand, for every $(v_k^1, \cdots, v_k^M) \in \mathcal{X}'(w_k)$, Lemma 2.2 implies that

$$C^{v_k^1}(\phi^1(\Delta t)) \cdots C^{v_k^M}(\phi^1(\Delta t)) = \frac{1}{6}\left(\frac{\Delta t}{2}\right)^n.$$

Since $|\mathcal{X}'(w_k)| = \binom{M}{n-1}$, combining the above equality with (3.6), we get

$$C^{w_k}(\phi^M(T)) = \frac{1}{12 \cdot (n-1)!}\left(\frac{T}{2}\right)^{n-1} \Delta t + \mathcal{O}\big(\frac{1}{M^2}\big),$$

which holds for each $w_k \in \mathcal{W}_{2n}^k$. Note that there are $4d^{n-2}\binom{d}{2}$ words in $\mathcal{W}_{2n}^k$ for each k, summing over k from 0 to $n-2$, we get

$$\sum_{w \in \mathcal{W}_{2n}} C^w(\phi^M(T)) = \frac{(d-1)T}{6M \cdot (n-2)!}\left(\frac{dT}{2}\right)^{n-1} + \mathcal{O}\big(\frac{1}{M^2}\big). \tag{3.7}$$

3.3 Putting all together

We are now in a position to prove the main claim. By Proposition 2.5, we have

$$\left\|\pi_{2n}(\phi^M(T)) - \pi_{2n}(\phi(T))\right\| = 2\sum_{w \in \mathcal{K}_{2n} \setminus \mathcal{S}_{2n}} C^w(\phi^M(T)). \tag{3.8}$$

Also by (3.1), we know that the coefficients of the words in $\mathcal{K}_{2n} \setminus (\mathcal{S}_{2n} \cup \mathcal{W}_{2n})$ are of order $O(\frac{1}{M^2})$, and thus

$$\sum_{w \in \mathcal{K}_{2n} \setminus \mathcal{S}_{2n}} C^w(\phi^M(T)) = \sum_{w \in \mathcal{W}_{2n}} C^w(\phi^M(T)) + \mathcal{O}(\frac{1}{M^2}).$$

Substituting (3.7) into the right hand side, and combining it with (3.8), we get

$$\left\| \pi_{2n}(\phi^M(T)) - \pi_{2n}(\phi(T)) \right\| = \frac{(d-1)T}{3M \cdot (n-2)!} \left(\frac{dT}{2} \right)^{n-1} + \mathcal{O}(\frac{1}{M^2}).$$

Multiplying $\frac{M}{T}$ on both sides, and letting $M \to +\infty$, we get

$$\lim_{M \to +\infty} \frac{M}{T} \left\| \pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T)) \right\| = \frac{d-1}{3 \cdot (n-2)!} \left(\frac{dT}{2} \right)^{n-1}.$$

Thus we have completed the proof of Theorem 1.3.

3.4 Proof of Theorem 1.4

It remains to explain why the expansion on the right hand side of (1.3) stops at $n-1$, and why all the coefficients are independent of M. Again, in light of Proposition 2.5, we only need to consider words in $\mathcal{K}_{2n} \setminus \mathcal{S}_{2n}$. Note that the contribution of each word w to the norm $\|\pi_{2n}(\cdot)\|$ is the sum of all values of its decomposition and 'allocation'[3].

More precisely, for each $w \in \mathcal{K}_{2n} \setminus \mathcal{S}_{2n}$, there is a unique maximal decomposition into 'even' words

$$w = w_1 * \cdots * w_k, \qquad k < n$$

in the sense that each $w_j \in \mathcal{S}_{2l_j}$ for some l_j, and that no further even decomposition is possible. We necessarily have $k < n$ since w is not in $\mathcal{S}_{2n}$.

Similar as before, the total number of 'allocations' of these k subwords into M slots is the number of nonnegative integer solutions to

$$x_1 + \cdots + x_M = k,$$

which equals

$$\binom{M+k-1}{M-1} = \frac{1}{k!} M(M+1) \cdots (M+k-1), \tag{3.9}$$

easily seen to be a polynomial in M with degree at most $n-1$, and the lowest degree term being M. On the other hand, by Lemma 2.2, the contribution to the norm of each allocation is $\lambda (\Delta t)^n$, where λ depends on the decomposition and allocation only. Thus, for each $w \in \mathcal{K}_{2n} \setminus \mathcal{S}_{2n}$, since the right hand side of (3.9) has no constant term (the lowest degree term is M), the coefficient C^w is a polynomial in $\frac{T}{M}$ with degree at most $n-1$ and all coefficients independent of M. Since $\left\| \pi_{2n}(\phi(T)) - \pi_{2n}(\phi^M(T)) \right\|$ is twice the sum of all such C^w's, it is also a polynomial in $\frac{T}{M}$ with degree at most $n-1$. We have thus concluded the proof of Theorem 1.4.

[3] Since all the terms we deal with here are nonnegative, we can neglect the absolute values, and simply sum all of them to get the l^1 norm.

References

[1] K.-T. Chen, Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula, *Annals of Mathematics*, Vol.65, No.1 (1957), pp.163-178. MR-0085251

[2] K.-T. Chen, Integration of paths - a faithful representation of paths by noncommutative formal power series, *Transactions of the A.M.S.*, Vol.89, No.2 (1958), pp.395-407. MR-0106258

[3] K.-T. Chen, Iterated path integrals, *Bulletin of the A.M.S.*, Vol.83, No.5 (1977), pp.831-879. MR-0454968

[4] I.Chevyrev, A characteristic function on the space of signatures of geometric rough paths, arxiv preprint, 2013.

[5] T.Fawcett, Problems in stochastic analysis: connections between rough paths and non-commutative harmonic analysis, PhD Thesis, University of Oxford, 2003.

[6] P.K.Friz, N.Victoir, Multidimensional Stochastic Processes as Rough Paths, *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, 2010. MR-2604669

[7] Y.LeJan, Z.Qian, Stratonovich's signatures of Brownian motion determine Brownian sample paths, to appear in *Probability Theory and Related Fields*, 2014. MR-3101845

[8] B.M.Hambly, T.J.Lyons, Uniqueness for the signature of a path of bounded variation and the reduced path group, *Annals of Mathematics*, Vol.171, No.1 (2010), pp.109-167. MR-2630037

[9] T.J.Lyons, Differential equations driven by rough signals, Rev.Mat.Iber, Vol.14, No.2 (1998), pp.215-310. MR-1654527

[10] T.J.Lyons, H.Ni, Expected signature of two dimensional Brownian Motion up to the first exit time of the domain, arXiv:1101.5902, 2011.

[11] T.J.Lyons, N.Victoir, Cubature on Wiener space, *Proceedings: Mathematical, Physical and Engineering Sciences*, Vol. 460, No. 2041, Stochastic Analysis with Applications to Mathematical Finance, pp. 169-198, 2004. MR-2052260

[12] H.Ni, The expected signature of a stochastic process, PhD Thesis, University of Oxford, 2012. MR-3271737

Acknowledgments. We thank our supervisor Terry Lyons for his support and helpful discussions. We also thank the referee for carefully reading the manuscript and giving valuable suggestions.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])
- Non profit, sponsored by IMS[2], BS[3], PKP[4]
- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1] OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3] BS: Bernoulli Society http://www.bernoulli-society.org/
[4] PK: Public Knowledge Project http://pkp.sfu.ca/
[5] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm