Area anomaly in the rough path Brownian scaling limit of hidden Markov walks

Olga Lopusanschi, Damien Simon

Area anomaly in the rough path Brownian scaling limit of hidden Markov walks

OLGA LOPUSANSCHI and DAMIEN SIMON

Laboratoire de Probabilités, Statistique et Modélisation (UMR 8001),
Sorbonne Université and CNRS,
4 place Jussieu, 75005 Paris, France.
E-mail: olga.lopusanschi@gmail.com,
damien.simon@sorbonne-universite.fr

AbstractWe study the convergence in rough path topology of a certain class of discrete processes, the hidden Markov walks, to a Brownian motion with an area anomaly. This area anomaly, which is a new object, keeps track of the time-correlation of the discrete models and brings into light the question of embeddings of discrete processes into continuous time. We also identify an underlying combinatorial structure in the hidden Markov walks, which turns out to be a generalization of the occupation time from the classical ergodic theorem in the spirit of rough paths.

Contents

1. Introduction

1.1. The context

Rough paths theory was introduced by T. Lyons in 1998 (see, for example, [18, 17]) in order to provide a deterministic setting to stochastic differential equations (SDEs) of the type

$$dy_t = f(y_t)[dx_t]$$

where $(y_t)_t$ is a path in a finite-dimensional vector space V', $(x_t)_t$ is a path in another finite-dimensional vector space V with Hölder regularity $\alpha < 1$ (which is often the case for stochastic processes) and $f : V' \to Hom(V, V')$ is a smooth map. Whenever the classical Young integration [19] fails (which is the case for $\alpha < 1/2$), paths may be lifted (in a non-unique way) to a larger more abstract space, the *rough path space*, for which existence, uniqueness and continuity of the solution map hold and become easier to prove.

The rough path space of level $k \geq 1$ corresponds to paths with values in

$$T_1^{(k)}(V) = V \oplus V^{\otimes 2} \oplus \dots V^{\otimes k}$$

and a finite-variation path $x : [0, T] \to V$ is lifted to a path $S(\gamma) : [0, T] \to T_1^{(k)}(V)$, called the step-$k$ signature of the path, through the formulae:

$$S(x)(t) = (S_1(x)(t), \dots, S_k(x)(t))$$
$$S_j(x)(t) = \int_{0<s_1<\dots<s_j<t} dx_{s_1} \otimes dx_{s_2} \otimes \dots \otimes dx_{s_j}$$

The coefficients of the signature satisfy a set of algebraic relations (shuffle product and concatenation product) and of analytic bounds (depending on the Hölder regularity $1/2 < \alpha$ of x), which will be detailed below. For a regularity $\alpha < 1/2$, some of the integrals above are not well-defined any more but the algebraic relations and the analytic bounds are kept as a new *definition* of the signature, and therefore also defines formally the iterated integrals.

Some examples of rough paths. The space of rough paths of level 2, for example, contains the enhanced Brownian motion [9, 4]

$$t \mapsto (B_t, (1/2)B_t \otimes B_t + A_t^{\text{Lévy}}) \tag{1}$$

where $(B_t)_t$ is a V-valued Brownian motion and $A_t^{\text{Lévy}}$ its Lévy area (with values in $V \wedge V$). The Brownian motion has regularity $1/2^-$ and its Lévy area is defined through the usual stochastic calculus by

$$A_t^{\text{Lévy}} = \frac{1}{2} \int_{0<s_1<s_2<t} (dB_{s_1} \otimes dB_{s_2} - dB_{s_2} \otimes dB_{s_1})$$

where the stochastic integration may be either in the Itô or Stratonovich sense.

The space of rough paths also contains less trivial objects such as two-dimensional area bubbles [14], defined as the limit signature of the sequence of paths

$$x_n(t) = \frac{1}{\sqrt{n}} (\cos(nt), \sin(nt))$$

which turns around $(0,0)$ faster as n increases. The signature is given by $((0,0), (t/2)e_1 \wedge e_2)$ (where (e_1, e_2) is the canonical basis of $\mathbb{R}^2$): the first level is constant at $(0,0)$, which corresponds to a constant path at $(0,0)$, whereas the second level corresponds to an accumulation of area at constant speed. This shows that the rough path space contain more information than the usual path space about the microscopic structure of approximations.

A quick look at the rough path space also shows that a combination of the two previous examples, i.e. the signature

$$S(B)(t) = \left(B_t, (1/2) B_t \otimes B_t + A_t^{\text{Lévy}} + t\Gamma \right), \tag{2}$$

where (B_t) is a two-dimensional Brownian motion and Γ is a fixed element of $V \wedge V$, is a valid rough path. From the point of view of stochastic integration however, this object is new since the construction of the Lévy area does not depend on the choice of Itô or Stratonovich integration and thus the Γ term is not an artefact of the choice of the stochastic integral. Thus, one may wonder whether this type of objects is relevant for applications and continuous time limits of discrete models.

Continuous time limit of discrete models in the rough path space: existing results. A first result is the generalization of Donsker theorem for i.i.d. random variables by Breuillard et al. [4], which, without surprise, gives the convergence in rough path topology to the enhanced Brownian motion (1) under suitable finite moment hypotheses.

A first result of convergence to (2) with non-zero Γ was obtained in [16]. It relies mostly on a geometric construction of periodic graphs and a non-reversible random walk dynamics on the graph. The idea is the following: at the scale of the period of the graph, the dynamics looks like the one of a random walk but, at the smallest scale of the periodic pattern, windings may occur and contribute to the area anomaly Γ.

The generalization to hidden Markov walks. The present paper presents a much more general construction of discrete time models converging to (2) in the rough path topology. In particular, the construction does not rely anymore on geometric properties, is purely probabilistic and the emergence of a non-zero Γ relies only on probabilistic short time correlations of the discrete time model.

Many discrete models may enter the present settings and exhibit a non-zero area anomaly in the limit. It was indeed a surprise to us that no such result have been previously considered in the literature, despite the vast literature about stochastic integration, continuous-time limits and rough paths.

The motivation for considering hidden Markov walks as defined below is multiple. When considering limits such as (2), it is very tempting to consider ad hoc models

obtained by a discretization of the limit but it does not shed any new light on the model and one of our purposes was to avoid such an approach. The central limit theorem and its process extension, Donsker theorem, show that obtaining anomalous behaviour around a normal law/Brownian motion from Markov models requires a fine tuning of the correlations at the discrete level. This is precisely what the present paper describes through hidden Markov walk.

Moreover, hidden Markov walks are simple objects from the point of view of numerical simulations and thus the present construction may lead to interesting discrete approximations, as for example in [7, 15].

Finally, the proofs of the results presented below introduce new objects in the study of discrete time Markov chains: iterated occupation times, which are a discrete analogue of iterated integrals and signature of the rough path theory. These iterated occupation times satisfy algebraic relations, such as quasi-shuffle and concatenation, and are of independent interest in the general theory of Markov chains.

1.2. Formulation of the results

1.2.1. Hidden Markov walks and the first theorem.

We first define the discrete time model we will consider throughout the paper and we choose to call it *hidden Markov walk*, which is a particular case of hidden Markov chain as introduced in [2].

Definition 1.1 (hidden Markov walk)**.** *Let E be a countable set and V a finite dimensional real vector space. A* hidden Markov walk *is a process $(R_n, X_n)_{n\in\mathbb{N}}$ on $E \times V$ such that:*

1. *the process $(R_n)_{n\in\mathbb{N}}$ is a Markov chain on E;*
2. *conditionally on the process (R_n), the increments $F_n = X_{n+1} - X_n$ are independent and have marginal laws such that, for any Borel set A of V, any $n \in \mathbb{N}$ and any $r \in E$, it holds:*

$$\mathbb{P}\left(F_n \in A | \sigma(R)\right) = \mathbb{P}\left(F_n \in A | R_n\right), \tag{3}$$

$$\mathbb{P}\left(F_n \in A | R_n = r\right) = \mathbb{P}\left(F_1 \in A | R_1 = r\right). \tag{4}$$

The process $(R_n, F_n)_{n\in\mathbb{N}}$ corresponds to the usual definition of a hidden Markov chain. The additional vector space structure of V allows one to consider the (F_n) as increments and to add them to obtain the process (X_n).

A classical way of embedding the discrete process $(X_n)_{n\in\mathbb{N}}$ in continuous time is the so-called Donsker embedding.

Definition 1.2 (Donsker embedding)**.** *Let $N \in \mathbb{N}$ and V be a vector space. Let $(x_n)_{0\leq n\leq N}$ be a V-valued sequence. The Donsker embedding $\iota_N(x)$ is the path $\iota_N(x) : [0, 1] \to V$ such*

that, for any $0 \leq k < N$ *and any* $s \in [0,1]$,

$$\iota_N(x)\left((1-s)\frac{k}{N} + s\frac{k+1}{N}\right) = (1-s)x_k + sx_{k+1},$$

which interpolates linearly between the (x_n) *and accelerate time by a factor* N.

We also define, for any real number $s > 0$, the dilation operators:

$$\begin{aligned} \delta_s : T_1^{(n)}(V) &\to T_1^{(n)}(V) \\ u_1 \otimes \ldots \otimes u_l &\mapsto s^l(u_1 \otimes \ldots \otimes u_l) \end{aligned}$$

where the $(u_i)_{1\leq i\leq l}$ are vectors of V and l an integer smaller than n.

We may now state our first theorem of convergence to an anomalous enhanced Brownian motion. The precise definition of the topological rough paths space $\mathcal{C}^\alpha([0,1], G^2(V))$ is given in section 2.1.

Theorem 1.1. *Let* $(R_n, X_n)_{n\in\mathbb{N}}$ *be a hidden Markov walk on* $E \times V$ *such that, for a fixed* $r_0 \in E$,

(i) $X_0 = 0$ *a.s. and* $R_0 = r_0$ *a.s.*

(ii) the Markov chain (R_n) *is irreducible and positive recurrent, with invariant probability* ν, *and the first return time*

$$T_1 = \inf\{n \geq 1; R_n = R_0\} \tag{5}$$

has finite moments of all orders.

(iii) the increments $F_n = X_{n+1} - X_n$ *satisfy, for all* $p \in \mathbb{N}$,

$$\sup_{r\in E} \mathbb{E}\left[||F_1||^p | R_1 = r\right] < \infty \tag{6}$$

(iv) the walk (X_n) *is centred, i.e. satisfies*

$$\sum_{r\in E} \nu(r)\mathbb{E}\left[F_1 | R_1 = r\right] = 0_V \tag{7}$$

Then, for any $1/3 < \alpha < 1/2$, *the sequence of* $G^2(V)$*-valued continuous time processes* $(\delta_{N^{-1/2}} \circ S \circ \iota_N(X)(t))_{t\in[0,1]}$ *(where* S *is the signature of the path) converges in law in the rough path topology of* $C^{0,\alpha\text{-}H\ddot{o}l}([0,1], G^2(V))$ *to the enhanced Brownian motion (as defined in* (2)*) with covariance matrix* C *and anomalous area drift* Γ. *Moreover, the limit law does not depend on the choice of the initial law of* R_0.

Explicit formulae for C *and* Γ *are presented in* (29), (30), (45a) *and* (45b).

The enhanced Brownian motion obtained in the limit is not the standard one, with identity covariance matrix and zero area drift, but has a covariance $\mathbb{E}\left[B_t^{(i)} B_t^{(j)}\right] = C_{ij}t$

and an area $\mathbb{E}\left[A_t^{(ij)}\right] = t\Gamma_{ij}$ (the standard Lévy area is centred). As detailed during the proof of the theorem and in section 2.3.2, the main ingredient for the existence of a non-zero Γ_{ij} is the fact that the area covered during one excursion of the Markov chain (R_n) may be *not* centered and we have:

$$\mathbb{E}_{r_0}\left[A_{T_1}^{(ij)}\right] = \Gamma_{ij}\mathbb{E}_{r_0}\left[T_1\right] \tag{8}$$

where T_1 is the length of the excursion ans $A_{T_1}^{(ij)}$ is the area covered by the discrete walk X_n between 0 and T_1.

The first hypothesis is not restrictive at all since there is global translation invariance. The irreducibility of (R_n) is not restrictive since one may always restrict E to one of its irreducible component. The hypotheses of positive recurrence and of finite moments of T_1 are important for the proofs but are trivially satisfied whenever E is finite. The hypothesis on the moments of the increments is already required in [4] to obtain rough path convergence of random walks. The assumption on centring is not restrictive: the centring is important to describe TCL-like fluctuations around the law-of-large-number asymptotics of X_n.

1.2.2. Iterated occupation times.

Definition and basic algebraic properties. The proof of theorem (1.1) uses extensively conditional expectations with respect to the process $\sigma(R)$ and the hidden Markov structure reduces iterated integrals with values in tensor products of the space V to combinatorial quantities related to the single process $(R_n)_{n\in\mathbb{N}}$. It appears that these combinatorial quantities also have a nice algebraic structure and interesting asymptotics.

Definition 1.3 (iterated occupation time of a sequence)**.** *Let E be a set and let $(x_n)_{0\leq n<N}$ be a sequence of elements of E. For any $k \geq 1$ and any sequence $u = (u_1, \ldots, u_k) \in E^k$, the iterated occupation time of (x_n) at u is defined as:*

$$\begin{aligned} L_u(x) = \mathrm{card}\{&(i_1, \ldots, i_k) \in \{0, ..., N-1\}^k; \\ &i_1 < i_2 < \ldots < i_k \quad \textit{and} \quad x_{i_1} = u_1, x_{i_2} = u_2, \ldots, x_{i_k} = u_k\} \end{aligned} \tag{9}$$

and, if $k = 0$, the empty sequence is noted ϵ and by convention $L_\epsilon(x) = 1$.

For $k = 1$, $L_u(x)$ counts the number of times the sequence x visits u, hence the name of iterated occupation times for larger k. Another writing of (9) makes the relation with iterated integrals clearer:

$$L_u(x) = \sum_{0\leq i_1<\ldots<i_k<N} \mathbf{1}_{x_{i_1}=u_1} \ldots \mathbf{1}_{x_{i_k}=u_k} \tag{10}$$

The iterated sum structure endows the iterated occupation times with both a concatenation structure and a quasi-shuffle structure.

Property 1.1 (concatenation structure)**.** *Let $(x_n)_{0\leq n<N}$ be a finite sequence of elements of E of length N. Let $u=(u_1,\dots,u_k)\in E^k$ for some k. Let M be an integer smaller than N. Then it holds*

$$L_{u_1\dots u_k}(x)=\sum_{p=0}^{k}L_{u_1\dots u_p}((x_n)_{0\leq n<M})L_{u_{p+1}\dots u_k}((x_n)_{M\leq n<N}) \tag{11}$$

where, by convention, $L_\epsilon(x)=1$ for ϵ the zero length sequence.

The quasi-shuffle property requires some additional combinatorial definitions. There are various approaches to quasi-shuffles and the interested reader may refer to [11, 12] for more results on this notion.

Definition 1.4 (quasi-shuffle product)**.** *Let E be a set and let $A(E)$ be the algebraic direct sum*

$$A(E)=\oplus_{k\geq 0}\oplus_{u\in E^k}\mathbb{R}u$$

(for $k=0$, the empty sequence is written ϵ) endowed with the shift operators $a\cdot u$ defined for $a\in E$ and $u=(u_1,\dots,u_k)\in E^k$ by

$$a\cdot u=(a,u_1,\dots,u_k)$$

The quasi-shuffle product $\widehat{\sqcup\!\sqcup}$ is defined recursively on the canonical basis by:

$$\begin{aligned}\epsilon\widehat{\sqcup\!\sqcup}\epsilon&=\epsilon\\(a\cdot u)\widehat{\sqcup\!\sqcup}\epsilon&=a\cdot u\\\epsilon\widehat{\sqcup\!\sqcup}(b\cdot v)&=b\cdot v\\(a\cdot u)\widehat{\sqcup\!\sqcup}(b\cdot v)&=a\cdot(u\widehat{\sqcup\!\sqcup}(b\cdot v))+b\cdot((a\cdot u)\widehat{\sqcup\!\sqcup}v)\\&\quad+\mathbf{1}_{a=b}a\cdot(u\widehat{\sqcup\!\sqcup}v)\end{aligned}$$

and extended to $A(E)$ by linearity.

To any finite sum $a=\sum_p c_p u_p$ with elements $u_p\in\cup_k E^k$, the definition of the iterated occupation time $L_u(x)$ of a sequence is extended by:

$$L_a(x)=\sum_p c_p L_{u_p}(x) \tag{12}$$

and is thus a linear application $A(E)\to\mathbb{R}$.

When the sequences u_1 and u_2 have no element in common, the set of quasi-shuffles $u_1\widehat{\sqcup\!\sqcup}u_2$ is equal to the classical shuffle of the two sequences. We now express the fact that the linear map $L_\bullet:A(E)\to\mathbb{R}$ is a morphism of algebra for the quasi-shuffle product.

Proposition 1.1. *Let $(x_n)_{0\leq n<N}$ be a finite sequence of elements of E. For any $k,l\geq 1$ and any sequences $u=(u_1,\dots,u_k)\in E^k$ and $v=(v_1,\dots,v_l)\in E^l$, it holds:*

$$L_u(x)L_v(x)=L_{u\widehat{\sqcup\!\sqcup}v}(x) \tag{13}$$

where the r.h.s has to interpreted in $A(E)$ through the extension (12).

Checking the following identity for $k=l=1$ is left as a warm-up exercise for the complete proof in section 2.3.1

$$L_{(u_1)}(x)L_{(v_1)}(x)=L_{(u_1v_1)}(x)+L_{(v_1u_1)}(x)+\begin{cases}L_{(u_1)(x)} & \text{if } u_1=v_1\\ 0 & \text{else}\end{cases}$$

The replacement of shuffles of iterated integrals by quasi-shuffles of iterated sums is related to the fact that the Lebesgue measure puts zero mass on singlets nor on $d-1$-dimensional subspaces of $\mathbb{R}^d$.

From hidden Markov paths to iterated occupation times. As announced, the definition 1.1 of hidden Markov walks fits nicely with the framework of iterated occupation times.

Property 1.2. *Let (R_n,X_n) be a hidden Markov walk on $E\times V$. Let $N\in\mathbb{N}$ and $\mathbf{X}_N:[0,1]\to V$ be the piecewise linear path $\mathbf{X}_N=\iota_N(X)$. Then, under suitable integrability hypothesis for the existence of the conditional expectation, there exist deterministic coefficients $(f_u)_{u\in E}$ in V, $(c_u)_{u\in E}$ in $V\otimes V$ and $(b_{u,v})_{u,v\in E}$ in $V\otimes V$ such that*

$$\mathbb{E}\left[\int_0^1 d\mathbf{X}_N(s)\middle|\sigma(R)\right]=\sum_{u\in E}f_uL_u((R_n)_{0\leq n<N})$$

$$\mathbb{E}\left[\int_{0<s_1<s_2<1}d\mathbf{X}_N(s_1)\otimes d\mathbf{X}_N(s_2)\middle|\sigma(R)\right]=\sum_{u\in E}c_uL_u((R_n)_{0\leq n<N})+\sum_{(u,v)\in E}b_{u,v}L_{uv}((R_n)_{0\leq n<N})$$

and more generally such decompositions on iterated occupation time hold for higher iterated integrals.

One checks that iterated integrals of level two are mapped to iterated occupation times of both level one and level two: the absence of coherent grading is related to the fact that iterated integrals satisfy a shuffle property whereas iterated sums satisfy only a quasi-shuffle property.

Description of C and Γ. There are explicit expressions for the deterministic coefficients C and Γ of theorems 1.1 and 1.2 that are easy to derive but that we choose to skip here. Section 2.3.2 is dedicated to such computations giving in particular explicit expressions relating iterated occupation times with C and Γ.

Generalized ergodic theorem for iterated occupation times. Following the same type of proof as in section 2.2 based on excursion theory of Markov chain, one can prove the following asymptotic property.

Proposition 1.2. *Let $(R_n)_{n\in\mathbb{N}}$ be a Markov chain on a countable space E such that it is irreducible and positive recurrent, with invariant probability ν. Then, for any sequence $(u_1,\dots,u_k)\in E_k$,*

$$\frac{L_{u_1\dots u_k}((R_n)_{0\le n<N})}{N^k} \xrightarrow[N\to\infty]{a.s.,L^1} \frac{\nu(u_1)\nu(u_2)\dots\nu(u_k)}{k!} \tag{14}$$

This result looks uninteresting at first sight and corresponds to the discrete equivalent of the convergence of $\delta_{N^{-1}}\circ S\circ\iota_N(X)$ to a zero limit due to the centring of the increments and the law of large number. A more interesting asymptotic result for iterated occupation times consists in considering fluctuations around the a.s. limit (14). This is done partially in section 2.3.3, even if it would be interesting to have a more general study of iterated occupation times.

1.2.3. The general question of embeddings.

Theorem 1.1 already encompasses a wide variety of discrete models but we formulate below a generalization of it by observing the two following facts.

First, the definition 1.1 of hidden Markov walk requires only that V is a semi-group in order to build $X_n = F_0\cdot F_1\cdot\ldots\cdot F_{n-1}$ out of its increments F_n and that it embeds nicely in $G^2(V)$ in order to formulate the theorem.

Besides, the choice of Donsker embedding is particular since it is one of the simplest embeddings one may consider on V but one may choose more useful embeddings if the discrete model already has its own geometrical embedding (see for example the round-about model described in section 1.3).

We thus generalize the definition 1.1 to the case of general embeddings.

Definition 1.5. *Let E be a countable set and V a finite-dimensional vector space. A hidden Markov path $((R_n,\mathbf{F}_n)_{n\in\mathbb{N}},(X_t)_{t\in\mathbb{R}_+})$ with regularity α is a process such that:*

(i) the process $(R_n)_{n\in\mathbb{N}}$ is a Markov chain on E;
(ii) the r.v. $\mathbf{F}_n$ have values in $C^{0,\alpha\text{-Höl}}([0,1],G_2(V))$
(iii) conditionally on the process (R_n), the increments $\mathbf{F}_n$ are independent *and have marginal laws such that, for any Borel set A, any $n\in\mathbb{N}$ and any $r\in E$,*

$$\mathbb{P}\left(\mathbf{F}_n\in A|\sigma(R)\right)=\mathbb{P}\left(\mathbf{F}_n\in A|R_n\right) \tag{15}$$
$$\mathbb{P}\left(\mathbf{F}_n\in A|R_n=r\right)=\mathbb{P}\left(\mathbf{F}_1\in A|R_1=r\right) \tag{16}$$

and moreover $\mathbf{F}_n(0)=0_{G^2(V)}$ a.s.;
(iv) the process (X_t) is obtained by concatenating the increments $\mathbf{F}_n$, i.e., for any $n\in\mathbb{N}$, for any $t\in[n,n+1[$,

$$X_t=\mathbf{F}_0(1)\cdot\ldots\mathbf{F}_{n-1}(1)\ldots\cdot\mathbf{F}_n(t-n) \tag{17}$$

where $\cdot$ is the product in $G_2(V)$.

Given this generalized definition, theorem 1.1 admits the following generalization.

Theorem 1.2. *Let E be a countable set and V a finite-dimensional vector space. Let $r_0 \in E$. Let $((R_n, \mathbf{F}_n)_{n\in\mathbb{N}}, (X_t)_{t\in\mathbb{R}_+})$ be a hidden Markov path such that:*

(i) $X_0 = 0_{G^2(V)}$ *a.s. and* $R_0 = r_0$ *a.s.*

(ii) the Markov chain (R_n) is irreducible and positive recurrent, with invariant probability ν, and the first return time $T_1 = \inf\{n \geq 1; R_n = R_0\}$ has finite moments for all integer p:

$$\mathbb{E}_{r_0}\left[T_1^p\right] < \infty \tag{18}$$

(iii) the increments $\mathbf{F}_n$ take values in $\bigcap_{\frac{1}{3}<\beta<\frac{1}{2}} C^{0,\beta\text{-Höl}}([0,1], G^2(V))$ and satisfy, for all $p \in \mathbb{N}$, the bound

$$\sup_{r\in E}\sup_{\frac{1}{3}<\beta<\frac{1}{2}} \mathbb{E}\left[d_\beta(\mathbf{F}_1, 0_{G^2(V)})^p \middle| R_1 = r\right] < \infty \tag{19}$$

(iv) the walk (X_t) is centred, i.e. satisfies

$$\sum_{r\in E} \nu(r)\mathbb{E}\left[\pi_1(\mathbf{F}_1(1))|R_1 = r\right] = 0_V \tag{20}$$

where $\pi_1(u)$ is the component in V of $u \in G^2(V) \subset V \oplus (V \otimes V)$.

Then, for any $1/3 < \alpha < 1/2$, the sequence of processes $(\delta_{N^{-1/2}}(X_{Nt}))_{t\in[0,1]}$ converges in law in the rough path topology of $C^{0,\alpha\text{-Höl}}([0,1], G^2(V))$ to the enhanced Brownian motion (as defined in (2)) with covariance matrix C and area anomaly $\Gamma \in V \wedge V$ given in (29) and (30), for any $1/3 < \alpha < 1/2$. Moreover, the limit law does not depend on the choice of r_0.

In particular, one verifies easily that theorem 1.1 is a consequence of theorem 1.2 by choosing the linear interpolation:

$$\mathbf{F}_n(t) = \exp(tF_n) = (tF_n, 0) \in V \oplus (V \wedge V) \simeq G^2(V) \tag{21}$$

(see section 2.1 for the exact definition of the Lie group $G^2(V)$). More generally, hypothesis (19) is satisfied as soon as the embeddings are smooth or Lipschitz and the increments have finite moments.

As for theorem 1.1, the area anomaly Γ obtained in (45b) is related to the area of the walk (X_n) covered during one excursion of the process (R_n). There may be now two contributions: one of them is related to the fact that the $\mathcal{F}_n$ may contain an area drift (this is the new part due to the nonlinear embeddings) and one of them is related to the area produced by correlations during one excursion.

1.3. Examples and heuristics

We present in this section various models to illustrate the previous theorems. The explicit formulae for C and Γ are given below in equations (45a) and (45b).

A random walk in $\mathbb{C}$ with rotating increments. Let $\omega = e^{2i\pi/L}$ be a root of unity with $L \geq 3$. Let $(U_k)_{k\in\mathbb{N}}$ be a sequence of i.i.d. real random variables with finite moments of all order. Identifying canonically $\mathbb{C}$ with $\mathbb{R}^2$, the process $(X_n)_{n\in\mathbb{N}}$ defined by $X_0 = 0$ and, for any $n \geq 1$

$$X_n = \sum_{k=0}^{n-1} \omega^k U_k$$

is a hidden Markov walk. Indeed, one may choose $E = \mathbb{Z}/L\mathbb{Z}$ and the deterministic dynamics $R_n = n \mod L$. The increment $F_n = \omega^{R_n} U_n$ depends only on R_n and of r.v. independent of the process (R_n). The first return time is constant $T_1 = L$ and thus $\mathbb{E}_0[T_1^p] = L^p < \infty$. Moreover one has the following computations.

$$\mathbb{E}_0[X_{T_1}] = \left(\sum_{k=0}^{L-1} \omega^k\right) \mathbb{E}[X_1] = 0$$

$$C_{ij} = \frac{1}{L}\mathbb{E}_0\left[X_{T_1}^{(i)} X_{T_1}^{(j)}\right] = \frac{\mathrm{Var}(X_1)}{2}\delta_{ij}$$

$$\Gamma = \frac{\cos(\pi/L)}{4\sin(\pi/L)}\mathbb{E}[X_1]^2 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

In the present case, the return time has an almost sure value and thus we may improve the proof of theorem 1.2 in order to relax the finite moment hypothesis on the U_k.

Spending time turning around. A case described by theorem 1.2 but not by theorem 1.1 is given by the following construction. We fix $E = \{1, 0\}$. If $R_n = 1$, then a centred random vector U_n is chosen and the path increment is the straight line:

$$\mathbf{F}_n(t) = (tU_n, 0) \tag{22}$$

If $R_n = 0$, then the path increment is a circle $c(t) = r(\cos(2\pi t) - 1, \sin(2\pi t))$. The covered area at time 1 is thus πr^2 and $\mathbf{F}_n(1) = (0, \pi r^2)$ where $\mathbb{R}^2 \wedge \mathbb{R}^2$ is identified to $\mathbb{R}$.

The process R_n is a Markov chain with transition matrix

$$Q = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

in the basis $(1, 0)$. Theorem 1.2 can be applied. An excursion corresponds to the sequence of states 1 or $100\ldots0$ where the number of 0 is a geometric law (starting at 1) and one obtains:

$$C = \frac{1}{\mathbb{E}_1[T_1]}\mathbb{E}_1[X_{T_1} \otimes X_{T_1}] = \frac{b}{a+b}\mathbb{E}[U_1 \otimes U_1]$$

$$\Gamma = \pi r^2 \frac{a}{a+b}$$

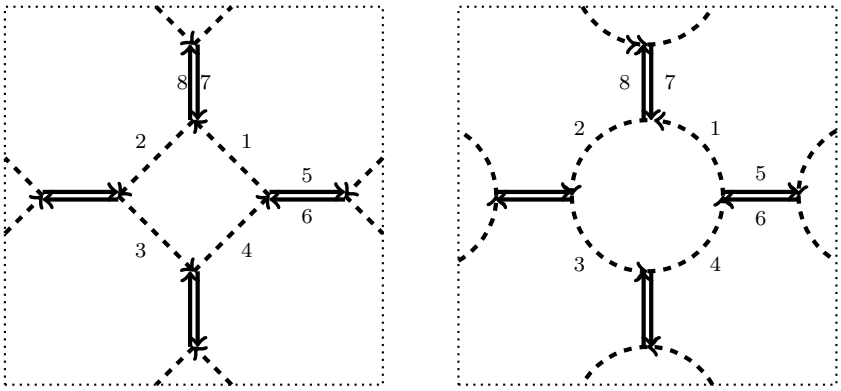

Figure 1. Diamond (left) and round-about (right) models on $\mathbb{Z}^2$-periodic patterns. At each vertex, there is exactly two incoming and two out-coming edges, one plain and one dashed. The dashed out-coming edge is chosen with probability p and the plain one with probability $1-p$. The space state E corresponds to the types of arrow. The difference between the two cases corresponds to the curved dashed arrow: in the roundabout model, the increment is the same but with an additional covered area.

If b goes to zero, one recovers the area bubbles mentioned in the introduction. If a goes to 0, there are no circles and one recovers the classical random walk.

Diamond and round-about models: the question of correlations and embeddings. We introduce two other models that illustrate theorem 1.2 in figure 1. Both models have the same space $E = \{1, 2, \dots, 8\}$. A value $r \in E$ corresponds to a unique type of edges on a $\mathbb{Z}^2$-periodic graph. In both models, the edges have the same increments in the plane $\mathbb{R}^2$ but differ by their embeddings: in the diamond model, all the embeddings are straight lines whereas, in the round-about model, part of the edges are circle arcs, which cover a non-zero area.

At the end of an arrow, there is exactly two out-coming edges, one plain and one dashed. The dashed out-coming edge is chosen with probability p and the plain one with probability $1 - p$.

If $p \geq 1/2$, the walk tends to be trapped in the diamonds/round-abouts and thus accumulates a covered area, which in the continuous limit, contributes to the area anomaly Γ. In the diamond model, all the contributions to Γ are of these type. In the round-about model, there is an additional contribution to Γ corresponding to the area covered by the circle arcs.

2. Mathematical tools and proofs.

2.1. The rough path space and its topology

All our notations and definitions follow closely the ones introduced in [4] and [9] and thus we sketch only the tools needed for the proofs. The space $G^2(V)$ is the subset of $T_1^{(2)}(V) = V \oplus (V \otimes V)$ of elements (v, M) such that there exists a smooth path $x : [0,1] \to V$ whose signature $S(x)(1)$ is equal to (v, M). It is easy to see that $G^2(V)$ is a subgroup isomorphic to $V \oplus V \wedge V$ since the symmetric part of M has to be equal to $(1/2)v \otimes v$ and thus can be skipped from the description. The group law is defined as

$$(a, A) \cdot (b, B) = (a + b, A + B + (1/2)(a \otimes b - b \otimes a))$$

and the inverse is given by $(a, A)^{-1} = (-a, -A)$. We define the two canonical projections:

$$\begin{array}{ll} \pi_1 : G^2(V) \to V & \qquad \pi_2 : G^2(V) \to V \wedge V \\ (a, A) \mapsto a & \qquad (a, A) \mapsto A \end{array}$$

The Carnot-Caratheodory norm $||u||$ of an element u is the infimum of the lengths[1] of smooth paths x such that $S(x)(1) = u$ and it induces a distance on $G^2(V)$ through $d(u_1, u_2) = \left|\left|u_1 \cdot u_2^{-1}\right|\right|$, making $G^2(V)$ a geodesic space.

Given two smooth paths $x_1, x_2 : [0,1] \to V$ such that $x_1(0) = x_2(0) = 0_V$, we introduce, for any $\alpha \in (1/3, 1/2)$, the distance:

$$d_\alpha(x_1, x_2) = \sup_{(s,t)\in[0,1]^2} \frac{\left|\left|S_2(x_1)(s,t) \cdot S_2(x_2)(s,t)^{-1}\right|\right|}{|t - s|^\alpha} \tag{23}$$

where $S_2(x)(s,t) = S_2(x)(t) \cdot S_2(x)(s)^{-1}$ is the $G^2(V)$-valued increment of the signature.

The d_α-closure of the set of signatures $(S(x))$ of smooth paths x at finite distance from the signature of the constant zero path is the Polish space (for d_α)

$$C^{0,\alpha\text{-Höl}}([0,1], G^2(V)) \tag{24}$$

In practice, the study of convergence in law of processes in the topology of this rough path space is made easier by two useful tools: equivalence of norms and Kolmogorov-Centsov tightness criterion. First, there is equivalence of norms on $G^2(V)$ between the Carnot-Caratheodory norm $||(a, A)||$ defined through geodesics and the norm $||(a, A)||'$ defined by:

$$||(a, A)||' = \sum_{i=1}^{d} |\pi_1^{(i)}(a)| + \sum_{1\le i<j\le d} |\pi_2^{(ij)}(A)|^{1/2} \tag{25}$$

where $\pi_1^{(i)}(a)$ and $\pi_2^{(ij)}(A)$ are the components of a and A in a given basis. Bounding $d_\alpha(x_1, x_2)$ thus only requires suitable bounds on every coefficient of the components of the paths.

[1] V is assumed to be Euclidean.

All the tightness criteria required by our theorem deal with $1/3 < \alpha < 1/2$ and thus, following [4], tightness of a sequence of processes $((X^{(N)}(t))_{t\in[0,1]})_{N\geq 1}$ in a fixed $C^{0,\alpha\text{-Höl}}([0,1], G^2(V))$ requires only that there exists $p \geq 1$ such that $\alpha \leq (2p-1)/(4p)$ and there exists C such that for any $s, t \in [0,1]$, it holds:

$$\sup_{N\geq 1} \mathbb{E}\left[\left|\left|X^{(N)}(t)\cdot X^{(N)}(s)^{-1}\right|\right|^{4p}\right] \leq C|t-s|^{2p} \tag{26}$$

2.2. Proof of the convergence theorem and formulae for the anomalous area drift

In this section, we fix a hidden Markov path $((R_n, \mathbf{F}_n)_{n\in\mathbb{N}}, (X_t)_{t\in\mathbb{R}_+})$ in $G^2(V)$ satisfying the hypothesis of theorem 1.2. The proof of theorem 1.2 relies on the following steps:

- we cut the trajectories of (R_n) into excursions;
- we then study the convergence of the accelerated geodesic interpolation $(\widehat{X}_t)$ of the process (X_t) between two successive return times of (R_n);
- we compare the finite-dimensional marginals of the two processes $\widehat{X}$ and X in the limit;
- we prove the tightness of the sequence of processes $(X^{(N)})_N$.

Cutting into independent excursions. The proof of theorem 1.2 relies on the division of the process $((R_n, \mathbf{F}_n)_{n\in\mathbb{N}}, (X_t)_{t\in\mathbb{R}_+})$ into time windows $[T_k, T_{k+1})$ corresponding to excursions of the Markov process (R_n) (see [13] for a general theory of excursions of Markov processes).

Proposition 2.1. *Let $(T_k)_{k\in\mathbb{N}}$ be the sequence of excursion times defined by:*

$$\begin{cases} T_0 & = 0 \\ T_{k+1} & = \inf\{n > T_k; R_n = R_{T_k}\}, \qquad k \in \mathbb{N} \end{cases} \tag{27}$$

Then, for any $r_0 \in E$, conditionally on $\{R_0 = r_0\}$, the r.v. $(\widehat{F}_k)_{k\in\mathbb{N}_1}$ defined by

$$\widehat{F}_k = \mathbf{F}_{T_{k-1}}(1)\cdot \mathbf{F}_{T_{k-1}+1}(1)\cdot \ldots \cdot \mathbf{F}_{T_k-1}(1) \tag{28}$$

is a sequence of independent and identically distributed $G^2(V)$-valued random variables.

Proof. The recurrence of (R_n) implies that all the T_k are finite a.s. and, from the general theory of discrete Markov processes, the excursions of the process (R_n) are independent and identically distributed. Each $\widehat{F}_k$ is a product of r.v. indexed by times belonging to the same excursion of the process (R_n) and the hidden Markov structure implies the result. □

Property 2.1 (moments of $\widehat{F}_1$)**.** *Under the hypotheses of theorem (1.2), the r.v. $(\widehat{F}_k)_{k\in\mathbb{N}}$ are independent and satisfy:*

(i) for all $p \in \mathbb{N}$ and all $r_0 \in E$, $\mathbb{E}_{r_0}\left[\left|\left|\widehat{F}_1\right|\right|^p\right] < \infty$

(ii) the projection on V are centred and have a finite covariance

$$\mathbb{E}_{r_0}\left[\pi_1(\widehat{F}_1)\right] = 0_V \qquad \mathbb{E}_{r_0}\left[\pi_1^{(i)}(\widehat{F}_1)\pi_1^{(j)}(\widehat{F}_1)\right] = C_{ij}\mathbb{E}_{r_0}\left[T_1\right] < \infty \tag{29}$$

where C_{ij} is symmetric and does not depend on the choice of r_0.

(iii) the expectation of the second level is given by

$$\mathbb{E}_{r_0}\left[\pi_2^{(ij)}(\widehat{F}_1)\right] = \Gamma_{ij}\mathbb{E}_{r_0}\left[T_1\right] \tag{30}$$

where Γ_{ij} is antisymmetric and does not depend on r_0.

Proof. Let $p \in \mathbb{N}$ be a fixed integer. The first point uses sub-additivity of the norm and the bound $(a_1 + \ldots + a_n)^p \leq n^p(a_1^p + \ldots + a_n^p)$ for positive numbers:

$$\left|\left|\widehat{F}_1\right|\right|^p \leq T_1^p \sum_{0\leq k<T_1} ||\mathbf{F}_k(1)||^p$$

$$\mathbb{E}_{r_0}\left[\left|\left|\widehat{F}_1\right|\right|^p \middle| \sigma(R)\right] \leq T_1^p \sum_{0\leq k<T_1} \mathbb{E}_{r_0}\left[||\mathbf{F}_k(1)||^p|\sigma(R)\right] \leq T_1^{p+1}C$$

using hypotheses (19) and (18).

The second point uses the classical representation property of invariant measure as marginal of the excursion measure of additive functionals: for any ν-integrable function $f : E \to \mathbb{R}$, it holds

$$\sum_{r\in E} f(r)\nu(r) = \frac{1}{\mathbb{E}_{r_0}\left[T_1\right]}\mathbb{E}_{r_0}\left[\sum_{0\leq k<T_1} f(R_k)\right] \tag{31}$$

for any $r_0 \in E$. In the present case, we apply this formula to

$$\mathbb{E}_{r_0}\left[\pi_1(\widehat{F}_1)\right] = \mathbb{E}_{r_0}\left[\sum_{0\leq k<T_1} \mathbb{E}_{r_0}\left[\pi_1(\mathbf{F}_k(1))|R_k\right]\right]$$

and use the centring hypothesis (20). C and Γ do not involve additive functionals (see below section 2.3.2 for more information and explicit formulae), however the independence with respect with r_0 can be proved in the same way as the previous property. □

Convergence of the excursion-geodesic extracted process. Out of the independent $G^2(V)$-valued r.v. $(\widehat{F}_k)$, we follow [4] and build the geodesic-interpolated processes $(\widehat{X}_t))_{t\in\mathbb{R}_+}$ defined by

$$\widehat{X}_t = \widehat{F}_1 \cdot \ldots \cdot \widehat{F}_{\lfloor t\rfloor} \cdot g(\widehat{F}_{\lfloor t\rfloor+1}, \{t\}) \tag{32}$$

where $g : G^2(V) \times [0,1] \to G^2(V)$ is defined such that $g(u,\cdot)$ is the geodesic in $G^2(V)$ joining $0_{G^2(V)}$ and u. For any real number t, $\lfloor t \rfloor$ and $\{t\}$ are respectively its integer and fractional parts.

The relation with the initial process X_t is that, for any $k \in \mathbb{N}$,

$$\widehat{X}_k = X_{T_k} \tag{33}$$

However, the sequence of processes $(\delta_{N^{-1/2}}(\widehat{X}_{Nt}))$ is such that a corrected version of the results of [4] can be applied and provide the following theorem and the following lemma. By the word *corrected*, we mean that theorem 3 of [4] should either require also the centering of the $\pi_2(\xi_i)$ in order to have no drift area or, if not, include an area drift Γ in the enhanced Brownian motion at the limit. One may be convinced for example by considering increments (ξ_i, a) where ξ_i is an random increment in V and a is a constant in $V \wedge V$. Then $\mathbf{W}^{(n)}_{k/n}$ is given by $\delta_{n^{-1/2}}\left(e^{\xi_1} \otimes \ldots \otimes e^{\xi_k}\right) \otimes (0, ka/n)$ since the last term is central in $G^2(V)$ and converges a.s. to the limit process $(0, ta)$ (using Slutsky's lemma, one has convergence to the anomalous enhanced Brownian motion). The proof of theorem 3 of [4] remains unchanged: the use of Stroock-Varadhan theorem is still valid but identifies a non-zero additional drift term in $V \otimes V$. The tightness criterion remains the same up to recentering of the area, which is costless. An alternative way of identifying this drift term is present in [6] in a more general context and with a precise description of all the terms.

Theorem 2.1 (from [4]). *Let $((R_n, \mathbf{F}_n)_{n\in\mathbb{N}}, (X_t)_{t\in\mathbb{R}_+})$ be a hidden Markov path in $G^2(V)$ satisfying the hypothesis of theorem 1.2 and let $1/3 < \alpha < 1/2$. Let $\beta = 1/\mathbb{E}_{r_0}[T_1]$. The sequence of processes $((\delta_{N^{-1/2}}(\widehat{X}_{N\beta t})_{t\in[0,1]})_{N\in\mathbb{N}^*}$ converges in distribution in the space $C^{0,\alpha\text{-Höl}}([0,1], G^2(V))$ to the enhanced Brownian motion with covariance matrix C and area anomaly Γ given in property 2.1. Moreover, the limit law does not depend on r_0.*

We emphasize that the dependence on r_0 in the construction of $\widehat{X}_k$ is due to the construction of the excursions of the process (R_n) and of their length T_1: the independence of C and Γ in property 2.1 requires a normalization by $\mathbb{E}_{r_0}[T_1]$, which is included in the previous theorem as a slow-down of the time scale. This can be also seen in eq. (33), using the a.s. asymptotic equivalent $T_k \simeq \mathbb{E}[T_1]\, k$.

Useful bounds on products of increments $\widehat{F}_n$.

Lemma 2.1 (from [4]). *For all integer $p \geq 1$, there exists $C'_p > 0$ such that, for any integers $n < m$, the following bound holds:*

$$\mathbb{E}\left[\left|\left|\widehat{F}_n \cdot \widehat{F}_{n+1} \cdot \ldots \cdot \widehat{F}_{m-1}\right|\right|^{4p}\right] \leq C_p (m-n)^{2p} \tag{34}$$

This lemma inherited directly from [4] corresponds to equation $(*)$ of [4] and its proof relies on the centering of the r.v. $\pi_1^{(i)}(\widehat{F}_n)$ and a equivalence of $||\cdot||$ with a norm which is a polynomial in the components of $\widehat{F}_n$.

However, this lemma is not precise enough for our purposes because of the fluctuations of the time scale between $\widehat{X}_{N\beta t}$ and the process X_{Nt} (more precisely $T_k/k \to \mathbb{E}[T_1]$ only asymptotically) and we need the following improved *maximal* version of the previous lemma.

Lemma 2.2. *For all integer $p \geq 1$, there exist $C'_p > 0$ such that, for any integers $n < m$, the following bounds hold:*

$$\mathbb{E}\left[\sup_{n\leq k\leq m-1}\left|\left|\widehat{F}_n\cdot\widehat{F}_{n+1}\cdot\ldots\cdot\widehat{F}_k\right|\right|^{4p}\right] \leq C'_p(m-n)^{2p} \tag{35}$$

$$\mathbb{E}\left[\sup_{n\leq k\leq l\leq m-1}\left|\left|\widehat{F}_k\cdot\widehat{F}_{k+1}\cdot\ldots\cdot\widehat{F}_l\right|\right|^{4p}\right] \leq 2^{4p}C'_p(m-n)^{2p} \tag{36}$$

Proof. We start with the proof of (35). Using the distance (25), it is enough to prove the existence of constants A_i and B_{ij} such that:

$$\mathbb{E}\left[\sup_{n\leq k\leq m-1}\left|\pi_1^{(i)}(\widehat{F}_n\cdot\widehat{F}_{n+1}\cdot\ldots\cdot\widehat{F}_k)\right|^{4p}\right] \leq A_i(m-n)^{2p} \tag{37a}$$

$$\mathbb{E}\left[\sup_{n\leq k\leq m-1}\left|\pi_2^{(ij)}(\widehat{F}_n\cdot\widehat{F}_{n+1}\cdot\ldots\cdot\widehat{F}_k)\right|^{2p}\right] \leq B_{ij}(m-n)^{2p} \tag{37b}$$

for any $1 \leq i,j \leq d$. By definition of $G^2(V)$, we have:

$$\begin{aligned}
\pi_1^{(i)}(\widehat{F}_n\cdot\widehat{F}_{n+1}\cdot\ldots\cdot\widehat{F}_k) &= \sum_{l=n}^{k}\pi_1^{(i)}(\widehat{F}_l)\\
\pi_2^{(ij)}(\widehat{F}_n\cdot\widehat{F}_{n+1}\cdot\ldots\cdot\widehat{F}_k) &= \sum_{l=n}^{k}\pi_2^{(ij)}(\widehat{F}_l)\\
&\quad+\sum_{n\leq l_1<l_2\leq k}\left(\pi_1^{(i)}(\widehat{F}_{l_1})\pi_1^{(j)}(\widehat{F}_{l_2})-\pi_1^{(j)}(\widehat{F}_{l_1})\pi_1^{(i)}(\widehat{F}_{l_2})\right)
\end{aligned}$$

The r.v. $\pi_1^{(i)}(\widehat{F}_l)$ are i.i.d. centred random variable and thus the sequence $(M_k^{(1)})_{k\geq n}$ of r.v. $M_k^{(1)} = \sum_{l=n}^{k}\pi_1^{(i)}(\widehat{F}_l)$ is a martingale and Doob's maximal inequality gives the first bound (37a) since $\mathbb{E}\left[(M_{m-1}^{(1)})^{4p}\right] \leq A'_i(m-n)^{2p}$.

We introduce now $a^{(ij)} = \mathbb{E}\left[\pi_2^{(ij)}(\widehat{F}_1)\right]$ and observe that, if $M_k^{(2)} = \pi_2^{(ij)}(\widehat{F}_n\cdot\ldots\cdot\widehat{F}_k) - (k-n)a^{(ij)}$, $(M_k^{(2)})_{k\geq n}$ is also a martingale since both $\left(\sum_{l=n}^{k}(\pi_2^{(ij)}(\widehat{F}_l)-a^{(ij)})\right)_{k\geq n}$ and

$$\left(\sum_{n\leq l_1<l_2\leq k}\left(\pi_1^{(i)}(\widehat{F}_{l_1})\pi_1^{(j)}(\widehat{F}_{l_2})-\pi_1^{(j)}(\widehat{F}_{l_1})\pi_1^{(i)}(\widehat{F}_{l_2})\right)\right)_{k\geq n}$$

are martingales. Moreover it holds:

$$\left|\pi_2^{(ij)}(\widehat{F}_\cdot \ldots \cdot \widehat{F}_k)\right|^{2p} \leq 2^{2p}\left(|a^{(ij)}|^{2p}(k-n)^{2p} + M_k^{2p}\right) \leq C_1(m-n)^{2p} + M_k^{2p}$$

Using again Doob's maximal inequality, we are left to prove

$$\mathbb{E}\left[(M_{m-1}^{(2)})^{2p}\right] \leq C_2(m-n)^{2p} \tag{38}$$

which can be done by direct computation. Hence we obtain eq. (37b).

The proof of (36) is obtained by the norm subadditivity through the following inequality for $n \leq k \leq l \leq m$

$$\begin{aligned}\left|\left|\widehat{F}_k \cdot \ldots \cdot \widehat{F}_l\right|\right|^{4p} &= \left|\left|(\widehat{F}_n \cdot \ldots \cdot \widehat{F}_{k-1})^{-1} \cdot (\widehat{F}_n \cdot \ldots \cdot \widehat{F}_l)\right|\right|^{4p} \\ &\leq 2^{4p}\left(\left|\left|\widehat{F}_n \cdot \ldots \cdot \widehat{F}_{k-1}\right|\right|^{4p} + \left|\left|\widehat{F}_n \cdot \ldots \cdot \widehat{F}_l\right|\right|^{4p}\right)\end{aligned}$$

Using then (35) gives directly (36). □

Useful bounds on the renewal process $(T_k)_{k\in\mathbb{N}}$. The sequence of excursion times $(T_k)_{k\in\mathbb{N}}$ is a renewal process on $\mathbb{N}$ (see [1] for a good introduction). We introduce, for any $x \in \mathbb{R}_+$, the integer-valued random variable $K(x)$ defined by

$$K(x) = \max\{n \in \mathbb{N}; T_n \leq x\}$$

This is the unique integer $K(x)$ such that $T_{K(x)} \leq x < T_{K(x)+1}$. The proofs below require deviation estimates of $K(x) - \beta x$ with $\beta = 1/\mathbb{E}[T_1]$ as well as moment estimates on increments $T_{K(x)+1} - T_{K(x)}$. We emphasize on the *inspection paradox*: although $T_{n+1} - T_n$ has the same law as T_1, $T_{K(x)+1} - T_{K(x)}$ is stochastically larger than T_1. However, in our case, the following property still holds:

Proposition 2.2. *Let (T_k) be a renewal process such that the $(T_{n+1} - T_n)_{n\in\mathbb{N}}$ are i.i.d. and there exists $P \in \mathbb{N}$ such that, for all $p \leq P$, $\mathbb{E}[T_1^p] < \infty$. Then, for all $p \leq P-1$, there exists $C_p > 0$ such that, for all $x \in \mathbb{R}_+$,*

$$\mathbb{E}\left[\left(T_{K(x)+1} - T_{K(x)}\right)^p\right] \leq C_p < \infty$$

Proof. It is sufficient to prove the result for x integer since $K(\lfloor x \rfloor) = K(x)$. We introduce the delay $D_j = T_j - T_{j-1}$ for $j \geq 1$ and we now decompose the event $\{D_{K(x)+1} = n\}$ along the values of $K(x)$:

$$\begin{aligned}\{D_{K(x)+1} = n\} &= \bigcup_{j\in\mathbb{N}} \{K(x) = j \text{ and } D_{j+1} = n\} \\ &= \bigsqcup_{j\in\mathbb{N}} \{T_j \leq x < T_j + n \text{ and } D_{j+1} = n\} \\ &= \bigsqcup_{x-n<k\leq x} \bigsqcup_{j\in\mathbb{N}} \{T_j = k \text{ and } D_{j+1} = n\}\end{aligned}$$

We now obtain

$$\mathbb{P}\left(D_{K(x)+1}=n\right)=\sum_{x-n<k\leq x}\sum_{j\in\mathbb{N}}\mathbb{P}\left(T_j=k \text{ and } D_{j+1}=n\right)$$

For fixed j, the r.v. T_j and D_{j+1} are independent by construction and D_{j+1} has the same law as T_1. We thus obtain

$$\begin{aligned}\mathbb{P}\left(D_{K(x)+1}=n\right)&=\mathbb{P}\left(T_1=n\right)\sum_{x-n<k\leq x}\sum_{j\in\mathbb{N}}\mathbb{P}\left(T_j=k\right)\\&\leq\mathbb{P}\left(T_1=n\right)\sum_{x-n<k\leq x}1\leq n\mathbb{P}\left(T_1=n\right)\end{aligned}$$

The expected uniform finite bounds on the moments of $T_{K(x)+1}-T_{K(x)}$ is then an easy consequence. □

We will also need the following deviation estimation on the r.v. $K(x)$ for the proof of lemma 2.3.

Proposition 2.3. *Let (T_k) be a renewal process such that the $(T_{n+1}-T_n)_{n\in\mathbb{N}}$ are i.i.d. and, there exists $p\in\mathbb{N}$, $\mathbb{E}\left[T_1^p\right]<\infty$. Let $\beta=1/\mathbb{E}\left[T_1\right]$. Then, for all real $a>1$ and $x>0$, it holds:*

$$\mathbb{P}\left(K(x)-\beta x\geq a\right)\leq C_p\frac{(\beta x+a)^{p/2}}{(a-1)^p} \tag{39a}$$

$$\mathbb{P}\left(K(x)-\beta x\leq -a\right)\leq\begin{cases}C_p\frac{(\beta x-a+1)^{p/2}}{(a-1)^p} & \text{for } a\leq\beta x\\0 & \text{for } a>\beta x\end{cases} \tag{39b}$$

Proof. We first relate the event $\{|K(x)-\beta x|\geq a\}$ to events related to the (T_n) using the equality $\{K(x)\geq n\}=\{T_n\leq x\}$:

$$\begin{aligned}\mathbb{P}\left(K(x)-\beta x\geq a\right)&=\mathbb{P}\left(K(x)\geq\lfloor\beta x+a\rfloor\right)\leq\mathbb{P}\left(T_{\lfloor\beta x+a\rfloor}\leq x\right)\\&\leq\mathbb{P}\left(T_{\lfloor\beta x+a\rfloor}-\mathbb{E}\left[T_1\right]\lfloor\beta x+a\rfloor\leq-(a-1)\mathbb{E}\left[T_1\right]\right)\end{aligned}$$

As a sum of independent centered r.v. with finite moments, it holds $\mathbb{E}\left[|T_n-n\mathbb{E}\left[T_1\right]|^p\right]\leq A_p n^{p/2}$. Using now Markov inequality produces (39a). The second inequality is obtained in the same way by considering complementary events:

$$\mathbb{P}\left(K(x)-\beta x\leq -a\right)\leq\mathbb{P}\left(K(x)\leq\lfloor\beta x-a\rfloor\right)=\mathbb{P}\left(T_{\lfloor\beta x-a\rfloor+1}>x\right)$$

□

Comparison of the finite-dimensional marginals of the two processes.

Lemma 2.3. *Let $\beta = 1/\mathbb{E}[T_1] \in (0, +\infty)$. Let $t \in [0,1]$. For any $\epsilon > 0$, it holds:*

$$\mathbb{P}\left(\left|\left|\delta_{N^{-1/2}}(\widehat{X}_{N\beta t})^{-1} \cdot \delta_{N^{-1/2}}(X_{Nt})\right|\right| > \epsilon\right) \xrightarrow{N\to\infty} 0 \tag{40}$$

Using the same argument repeatedly, the same convergence in probability to 0 can be extended to a family $(t_1, \dots, t_m)$ and thus, by Slutsky's lemma, it shows that both processes have the same finite-dimensional marginal laws.

Proof. The idea behind the lemma comes from eq. (33): the stopping times T_k are asymptotically equivalent to $k\mathbb{E}[T_1]$. The differences between the two processes have two origins: first, the difference between T_k and $k\mathbb{E}[T_1]$ and second, the difference between the geodesic interpolation and the stochastic path X_t.

The two r.v. $\widehat{X}_{Nt\beta}$ and X_{Nt} are given by:

$$\begin{aligned}
\widehat{X}_{N\beta t} &= \widehat{F}_1 \cdot \ldots \cdot \widehat{F}_{\lfloor N\beta t\rfloor} \cdot g(\widehat{F}_{\lfloor Nt\beta\rfloor+1}, \{Nt\beta\}) \\
X_{Nt} &= \widehat{F}_1 \cdot \ldots \cdot \widehat{F}_{K(\lfloor Nt\rfloor)} \cdot \mathbf{F}_{T_{K(Nt)}}(1) \cdot \ldots \cdot \mathbf{F}_{\lfloor Nt\rfloor-1}(1) \cdot \mathbf{F}_{\lfloor Nt\rfloor}(\{Nt\})
\end{aligned}$$

where $K(u)$ is the unique integer such that

$$T_{K(u)} \le u < T_{K(u)+1}.$$

We then obtain the bound:

$$\begin{aligned}
\left|\left|\widehat{X}_{N\beta t}^{-1} \cdot X_{Nt}\right|\right|^q \le 3^q \Big(& \left|\left|g(\widehat{F}_{\lfloor N\beta t\rfloor+1}, \{Nt\beta\})\right|\right|^q \\
& + \left|\left|\widehat{F}_{\min(K(Nt),\lfloor Nt\beta\rfloor)+1} \cdot \ldots \cdot \widehat{F}_{\max(K(Nt),\lfloor Nt\beta\rfloor)}\right|\right|^q \\
& + \left|\left|\mathbf{F}_{T_{K(Nt)}}(1) \cdot \ldots \cdot \mathbf{F}_{\lfloor Nt\rfloor-1}(1) \cdot \mathbf{F}_{\lfloor Nt\rfloor}(\{Nt\})\right|\right|^q \Big)
\end{aligned}$$

The first norm is trivially bounded by $\left|\left|\widehat{F}_{\lfloor N\beta t\rfloor+1}\right|\right|^q$, which has a finite expectation from property 2.1. The third norm is also easy to bound in expectation:

$$\begin{aligned}
& \mathbb{E}\left[\left|\left|\mathbf{F}_{T_{K(Nt)}}(1) \cdot \ldots \cdot \mathbf{F}_{\lfloor Nt\rfloor-1}(1) \cdot \mathbf{F}_{\lfloor Nt\rfloor}(\{Nt\})\right|\right|^q \middle| \sigma(R)\right] \\
& \le (\lfloor Nt\rfloor - T_{K(Nt)})^q \sum_{k=T_{K(Nt)}}^{\lfloor Nt\rfloor} \mathbb{E}\left[d_\alpha(\mathbf{F}_k, 0_{G^2(V)})^q \middle| R_k\right] \\
& \le A(\lfloor Nt\rfloor - T_{K(Nt)})^{q+1} \le A(T_{K(Nt)+1} - T_{K(Nt)})^{q+1}
\end{aligned}$$

and thus the expectation of the q-th power of the norm is finite from proposition 2.2 and hypothesis 18.

In order to bound from above the expectation of the second norm, we introduce the sequence $u_N = \lfloor N^\gamma \rfloor$ for some real $\gamma > 0$ chosen below. We then obtain, using $q = 4p$ and lemma 2.2:

$$\mathbb{E}\left[\left|\left|\widehat{F}_{\min(K(Nt),\lfloor Nt\beta\rfloor)-1}\cdot\ldots\cdot\widehat{F}_{\max(K(Nt),\lfloor Nt\beta\rfloor)}\right|\right|^{4p}\mathbf{1}_{|K(Nt)-\lfloor Nt\beta\rfloor|\leq u_N}\right]$$
$$\leq \mathbb{E}\left[\sup_{\lfloor N\beta t\rfloor-u_N\leq k\leq l\leq\lfloor N\beta t\rfloor+u_N}\left|\left|\widehat{F}_k\cdot\ldots\cdot\widehat{F}_l\right|\right|^{4p}\right]\leq K_1u_N^{2p}$$

The event $|K(Nt) - \lfloor N\beta t\rfloor| > u_N$ has a small probability for large N. Moreover the previous norm contains at most $\lfloor Nt\rfloor$ terms since $K(Nt) \leq \lfloor Nt\rfloor$ and $\beta > 1$. We use lemma 2.2 and Cauchy-Schwarz inequality to obtain

$$\begin{aligned}J(u_N) =&\mathbb{E}\left[\left|\left|\widehat{F}_{\min(K(Nt),\lfloor Nt\beta\rfloor)}\cdot\ldots\cdot\widehat{F}_{\max(K(Nt),\lfloor Nt\beta\rfloor)-1}\right|\right|^{4p}\mathbf{1}_{|K(Nt)-\lfloor Nt\beta\rfloor|>u_N}\right]\\ \leq&\mathbb{E}\left[\sup_{1\leq k\leq l\leq\lfloor Nt\rfloor}\left|\left|\widehat{F}_k\cdot\ldots\cdot\widehat{F}_l\right|\right|^{4p}\mathbf{1}_{|K(Nt)-\lfloor Nt\beta\rfloor|>u_N}\right]\\ \leq&K_2N^{2p}\,\mathbb{P}\left(|K(Nt)-\lfloor Nt\beta\rfloor|>u_N\right)^{1/2}\end{aligned}$$

Proposition 2.3 for $p = 2$ immediately gives the following bound for N large enough:

$$J(u_N)\leq K_4\frac{N^{2p+1/2}}{u_N}$$

Collecting all the results with the dilation $\delta_{N^{-1/2}}$ gives for $4p + 3 \leq r$

$$\mathbb{E}\left[\left|\left|\delta_{N^{-1/2}}(\widehat{X}_{N\beta t})^{-1}\cdot\delta_{N^{-1/2}}(X_{Nt})\right|\right|^{4p}\right]\leq\frac{A_1}{N^{2p}}+\frac{K_1u_N^{2p}}{N^{2p}}+K_4\frac{N^{1/2}}{u_N}$$

Any choice $1/2 < \gamma < 1$ implies that the expectation tends to zero, hence the convergence in probability. □

Tightness of the initial process.

Lemma 2.4 (tightness)**.** *Under the hypotheses of theorem 1.2, the tightness criterion* (26) *holds for the sequence of processes* $(\delta_{N^{-1/2}}(X_{Nt}))_{t\in[0,1]}$.

Proof. The proof is similar to the one of the previous lemma. We fix $s < t$. If $\lfloor Ns\rfloor = \lfloor Nt\rfloor = j$, then

$$\left|\left|X_{Ns}(s)^{-1}\cdot X_{Nt}(t)\right|\right|^{4p}=||\mathbf{F}_j(Nt-Ns)||^{4p}\leq d_\beta(\mathbf{F}_j,0_{G^2(V)})^{4p}(Nt-Ns)^{4p\beta}$$

for any $1/3 < \beta < 1/2$. Hypothesis (19) gives

$$\mathbb{E}\left[\left|\left|X_{Ns}(s)^{-1}\cdot X_{Nt}(t)\right|\right|^{4p}\right]\leq C(Nt-Ns)^{4p\beta}$$

for any $\beta < 1/2$, hence the bound $N^{2p}(t-s)^{2p}$ by taking $\beta \to 1/2$.

If $\lfloor Ns \rfloor < \lfloor Nt \rfloor$, we introduce the event

$$E = \left\{ T_{K(Ns)+1} < T_{K(Nt)} \right\}$$

which corresponds to at least one complete excursion separating Ns and Nt. On E^c, we use the decomposition

$$X_{Ns}^{-1} \cdot X_{Nt} = \left(\mathbf{F}_{\lfloor Ns \rfloor}(\{Ns\})^{-1} \cdot \mathbf{F}_{\lfloor Ns \rfloor}(1) \right) \cdot \mathbf{F}_{\lfloor Ns \rfloor + 1}(1) \cdot \ldots \cdot \mathbf{F}_{\lfloor Nt \rfloor - 1}(1) \cdot \mathbf{F}_{\lfloor Nt \rfloor}(\{Nt\})$$

and, on E, we factorize the excursions using the $\widehat{F}_k$:

$$\begin{aligned} X_{Ns}^{-1} \cdot X_{Nt} = & \mathbf{F}_{\lfloor Ns \rfloor}(\{Ns\})^{-1} \cdot \mathbf{F}_{\lfloor Ns \rfloor}(1) && \text{(41a)} \\ & \cdot \mathbf{F}_{\lfloor Ns \rfloor + 1}(1) \cdot \ldots \cdot \mathbf{F}_{T_{K(Ns)+1} - 1}(1) && \text{(41b)} \\ & \cdot \left(\widehat{F}_{K(Ns)+2} \cdot \ldots \cdot \widehat{F}_{K(Nt))} \right) && \text{(41c)} \\ & \cdot \mathbf{F}_{T_{K(Nt)}}(1) \cdot \ldots \cdot \mathbf{F}_{\lfloor Nt \rfloor - 1}(1) && \text{(41d)} \\ & \cdot \mathbf{F}_{\lfloor Nt \rfloor}(\{Nt\}) && \text{(41e)} \end{aligned}$$

First case: on E. We will use repeatedly the following ordering valid on E:

$$T_{K(Ns)} \leq Ns < \lfloor Ns \rfloor + 1 \leq T_{K(Ns)+1} < T_{K(Nt)} \leq \lfloor Nt \rfloor \leq Nt < T_{K(Nt)+1}$$

We call I_1, I_2, I_3, I_4 and I_5 the respective norms of the terms (41a), (41b),(41c), (41d) and (41e). We then have

$$\left|\left| \delta_{N^{-1/2}}(X_{Ns})^{-1} \cdot \delta_{N^{-1/2}}(X_{Nt}) \right|\right|^{4p} \leq \frac{5^{4p}}{N^{2p}} \sum_{k=1}^{5} I_k^{4p}$$

Bounding I_1^{4p} and I_5^{4p} from above uses the distance (23)

$$\begin{aligned} ||I_1||^{4p} &\leq d_\beta(\mathbf{F}_{\lfloor Ns \rfloor}, 0_{G^2(V)})^{4p} (1 - \{Ns\})^{4p\beta} \\ ||I_5||^{4p} &\leq d_\beta(\mathbf{F}_{\lfloor Nt \rfloor}, 0_{G^2(V)})^{4p} (\{Nt\})^{4p\beta} \end{aligned}$$

Taking $\sigma(R)$-conditional expectation with the bound (19) gives a bound valid for all $\beta < 1/2$:

$$\begin{aligned} \mathbb{E}\left[||I_1||^{4p} \middle| \sigma(R) \right] &\leq C(1 - \{Ns\})^{4p\beta} \leq C(Nt - Ns)^{4p\beta} \\ \mathbb{E}\left[||I_5||^{4p} \middle| \sigma(R) \right] &\leq C(\{Nt\})^{4p\beta} \leq C(Nt - Ns)^{4p\beta} \end{aligned}$$

and thus for $\beta \uparrow 1/2$ we obtain the expected bound $N^{2p}(t-s)^{2p}$ for $\mathbb{E}\left[||I_1||^{4p} \mathbf{1}_E \right]$ and $\mathbb{E}\left[||I_5||^{4p} \mathbf{1}_E \right]$.

The bounds on I_2 and I_4 are similar and we write down only the one for I_2 using again hypothesis (19):

$$\begin{aligned}
||I_2||^{4p} &\leq \left(\sum_{k=\lfloor Ns \rfloor + 1}^{T_{K(Ns)+1}-1} d_\beta(\mathbf{F}_k, 0_{G^2(V)}) \right)^{4p} \\
&\leq \left(T_{K(Ns)+1} - \lfloor Ns \rfloor - 1 \right)^{4p} \sum_{k=\lfloor Ns \rfloor + 1}^{T_{K(Ns)+1}-1} d_\beta(\mathbf{F}_k, 0_{G^2(V)})^{4p} \\
&\leq C \left(T_{K(Ns)+1} - \lfloor Ns \rfloor - 1 \right)^{4p+1} \\
&\leq C \left(T_{K(Ns)+1} - T_{K(Ns)} \right)^{2p+1} (Nt - Ns)^{2p}
\end{aligned}$$

where the last inequality is true on E only. The excursion times have finite moments from proposition 2.2 and hypothesis 18 and thus:

$$\mathbb{E}\left[||I_2||^{4p} \mathbf{1}_E \right] \leq C'(t-s)^{2p} \tag{42}$$

The bound on I_3 can be obtained using lemma 2.2. Since the number $K(Nt) - K(Ns)$ between Ns and Nt is necessarily smaller than $N(t-s)$, we have the bound

$$\mathbb{E}\left[I_3(Ns, Nt)^{4p} \mathbf{1}_E \right] \leq \mathbb{E}\left[\sup_{1 \leq k \leq N(t-s)} \left|\left| \widehat{F}_{K(Ns)+2} \cdot \ldots \cdot \widehat{F}_{K(Ns)+1+k} \right|\right|^{4p} \right]$$

We call $Z(T_{K(Ns)+1})$ the positive r.v. in the r.h.s. since it is is a product of r.v. $\mathbf{F}_j(1)$ with $j \geq T_{K(Ns)+1}$. The filtration $(\mathcal{F}_n)$ is defined as $\mathcal{F}_n = \sigma((R_k, F_k); k \leq n)$. We now have:

$$\mathbb{E}\left[Z(T_{K(Ns)+1}) \right] = \sum_{p \in \mathbb{N}} \mathbb{E}\left[Z(T_{p+1}) \mathbf{1}_{K(Ns)=p} \right] = \sum_{p \in \mathbb{N}} \mathbb{E}\left[\mathbb{E}\left[Z(T_{p+1}) \middle| \mathcal{F}_{T_{p+1}} \right] \mathbf{1}_{K(Ns)=p} \right]$$

since the event $\{K(Ns) = p\} = \{T_p \leq Ns < T_{p+1}\}$ is $\mathcal{F}_{T_{p+1}}$-measurable. Using the strong Markov property for the hidden Markov chain $(R_n, \mathbf{F}_n)$ and the fact that $R_{T_{p+1}} = r_0$, we obtain

$$\mathbb{E}\left[Z(T_{K(Ns)+1}) \right] = \sum_{p \in \mathbb{N}} \mathbb{E}\left[Z(0) \mathbf{1}_{K(Ns)=p} \right] = \mathbb{E}\left[Z(0) \right]$$

We got rid of the dependency on $K(Ns)$ and lemma 2.2 gives the final inequality

$$\mathbb{E}\left[I_3(Ns, Nt)^{4p} \mathbf{1}_E \right] \leq B N^{2p} (t-s)^{2p}$$

Second case: on E^c. On E^c, we have the following ordering:

$$T_{K(Ns)} \leq \lfloor Ns \rfloor \leq Ns < \lfloor Ns \rfloor + 1 \leq \lfloor Nt \rfloor \leq Nt < T_{K(Ns)+2} \tag{43}$$

The difference is that the term I_3 is absent and I_2 and I_4 may be combined such that, on E^c:

$$\left|\left|X_{Ns}^{-1}\cdot X_{Nt}\right|\right|^{4p}\leq 3^{4p}\left(||I_1||^{4p}+||I_{24}||^{4p}+||I_5||^{4p}\right)$$

where $I_{24}=\mathbf{F}_{\lfloor Ns\rfloor+1}(1)\cdot\ldots\cdot\mathbf{F}_{\lfloor Nt\rfloor-1}(1)$. The upper bounds on I_1 and I_5 are the same as in the previous case. For I_{24}, we have on E^c

$$||I_{24}||^{4p}\,\mathbf{1}_{E^c}\leq(\lfloor Nt\rfloor-\lfloor Ns\rfloor-1)^{4p}\sum_{k=\lfloor Ns\rfloor+1}^{\lfloor Nt\rfloor-1}d_\beta(\mathbf{F}_k,0_{G^2(V)})^{4p}$$

and thus now using (19) and the ordering (43) on E^c

$$\begin{aligned}\mathbb{E}\left[||I_{24}||^{4p}\middle|\sigma(R)\right]\mathbf{1}_{E^c}&\leq C(\lfloor Nt\rfloor-\lfloor Ns\rfloor-1)^{4p+1}\mathbf{1}_{E^c}\\&\leq C(Nt-Ns)^{2p}(T_{K(Ns)+2}-T_{K(Ns)})^{2p+1}\mathbf{1}_{E^c}\end{aligned}$$

Hypothesis (18) then gives the desired bound $C\mathbb{E}\left[T_2^{2p+1}\right]$ on $\mathbb{E}\left[||I_{24}||^{4p}\,\mathbf{1}_{E^c}\right]$.

Collecting all the previous bounds gives the expected tightness criterion since all the bounds are of the form $AN^{2p}(t-s)^{2p}$. □

Remarks on the hypotheses (18) and (19). Some of the hypotheses of theorem 1.2 could be slightly relaxed by improving the previous proof or by considering only a fixed given $\alpha<1/2$. In the case where only one value $\alpha<1/2$ is targeted, one could use the same approach as [4] and require only a finite set of finite moments. However, in practice, one is often interested to the case $\alpha\to 1/2$. In this case, the finite moments hypothesis on T_1 cannot be relaxed since the variable $\widehat{F}_k$ are required to have moments of all orders in order to apply the results of [4]. Only efforts may be made on the requirement (19) using for example correlations between the length T_k of an excursion and the corresponding increment $\widehat{F}_k$. The bound (19) is not restrictive in practice since it encompasses already the case where the embeddings $\mathbf{F}_k$ are smooth or Lipschitz.

2.3. Iterated occupation times, the quasi-shuffle property and asymptotics

2.3.1. Proof of the quasi-shuffle property (1.1)

Proof. We first prove the following recursive decomposition. If $u=u_1\cdot u'$ with $u'=(u_2,\ldots,u_k)$ is the concatenation of the length one sequence u_1 and the sequence u', then

$$\begin{aligned}L_{u_1\cdot u'}(x)&=\sum_{0\leq i_1}\mathbf{1}_{x_{i_1}=u_1}\left(\sum_{i_1<i_2<\ldots<i_k\leq N}\mathbf{1}_{x_{i_2}=u_2}\ldots\mathbf{1}_{x_{i_k}=u_k}\right)\\&=\sum_{0\leq i_1<N}\mathbf{1}_{x_{i_1}=u_1}L_{u'}((x_n)_{i_1+1\leq n<N})\end{aligned}$$

For two sequences $u = u_1 u'$ and $v = v_1 v'$ of length larger than 1, the previous equation gives:

$$\begin{aligned}
L_{u_1\cdot u'}(x)L_{v_1\cdot v'}(x) &= \sum_{0\le i_1<j_1} \mathbf{1}_{x_{i_1}=u_1}\mathbf{1}_{x_{j_1}=v_1} L_{u'}((x_n)_{i_1+1\le n<N})L_{v'}((x_n)_{j_1+1\le n<N}) \\
&\quad + \sum_{0\le j_1<i_1} \mathbf{1}_{x_{i_1}=u_1}\mathbf{1}_{x_{j_1}=v_1} L_{u'}((x_n)_{i_1+1\le n<N})L_{v'}((x_n)_{j_1+1\le n<N}) \\
&\quad + \sum_{0\le i} \mathbf{1}_{x_i=u_1}\mathbf{1}_{u_1=j_1} L_{u'}((x_n)_{i+1\le n<N})L_{v'}((x_n)_{i+1\le n<N}) \\
&= \sum_{0\le i_1} \mathbf{1}_{x_{i_1}=u_1} L_{u'}((x_n)_{i_1+1\le n<N})L_{v_1v'}((x_n)_{i_1+1\le n<N}) \\
&\quad + \sum_{0\le j_1} \mathbf{1}_{x_{j_1}=v_1} L_{u_1u'}((x_n)_{j_1+1\le n<N})L_{v'}((x_n)_{j_1+1\le n<N}) \\
&\quad + \mathbf{1}_{u_1=v_1}\sum_{0\le i} \mathbf{1}_{x_i=u_1} L_{u'}((x_n)_{i+1\le n<N})L_{v'}((x_n)_{i+1\le n<N}) \\
&= L_{u_1\cdot u' \widehat{\sqcup\!\sqcup} v_1\cdot v'}(x)
\end{aligned}$$

from the definition of the quasi-shuffle product. The expected result is then obtained from the previous equation by recursion on the sum of the lengths of u and v by setting for convenience $L_\epsilon(x) = 0$. □

2.3.2. Relation between C, Γ and the first iterated occupation times

The covariance matrix C and the anomalous area drift Γ are obtained in property 2.1 in terms of the moments of the law of the i.i.d.r.v. $\widehat{F}_k$. It may be interesting to have more explicit formulae for C and Γ, since they describe completely the limit law.

To this purpose, we introduce, for all $u \in E$, the following expectation values, which are related to the conditional law of the $\mathbf{F}_k$ and do not depend on the law of the process $(R_n)_{n\in\mathbb{N}}$,

$$\begin{aligned}
f_u &= \mathbb{E}_{r_0}\left[\pi_1(\mathbf{F}_1(1))|R_1 = u\right] \\
c_u &= \mathbb{E}_{r_0}\left[\pi_1(\mathbf{F}_1(1)) \otimes \pi_1(\mathbf{F}_1(1))|R_1 = u\right] \in V \odot V \\
\gamma_u &= \mathbb{E}_{r_0}\left[\pi_2(\mathbf{F}_1(1))|R_1 = u\right] \in V \wedge V
\end{aligned}$$

where $V \odot V$ is the symmetric subspace of $V \otimes V$. The proof of proposition 2.1 uses the fact that the invariant probability of the Markov chain (R_n) satisfies, for any $u \in E$,

$$\nu(u) = \frac{\mathbb{E}_{r_0}\left[L_u((R_n)_{0\le n<T_1})\right]}{\mathbb{E}_{r_0}\left[T_1\right]}$$

Using the iterated occupation times, we now define, for any $k \in \mathbb{N}$, the iterated measure on E^k:

$$\nu_{k,r_0}(u_1,\ldots,u_k) = \frac{\mathbb{E}_{r_0}\left[L_{u_1\ldots u_k}((R_n)_{0\le n<T_1})\right]}{\mathbb{E}_{r_0}\left[T_1\right]}, \tag{44}$$

which coincide, for $k = 1$, with ν. For $k \neq 1$, it depends on the initial point r_0: this can be seen for example in the total mass related to the moments of T_1, which depends on the initial point r_0. However, the expression below for C and Γ do not depend on r_0. It would be interesting to understand in more details the dependence on the initial point, but it is more a question of general theory of Markov processes than a rough path question.

Property 2.2. *The covariance matrix C and the anomalous area drift Γ, defined in property 2.1 and appearing in theorems 1.1 and 1.2 are given by:*

$$C_{ij} = \sum_{u \in E} c_u^{(ij)} \nu(u) + \sum_{(u,v) \in E^2} (f_u^{(i)} f_v^{(j)} + f_u^{(j)} f_v^{(i)}) \nu_2(u,v) \tag{45a}$$

$$\Gamma_{ij} = \sum_{u \in E} \gamma_u^{(ij)} \nu(u) + \frac{1}{2} \sum_{(u,v) \in E^2} (f_u^{(i)} f_v^{(j)} - f_u^{(j)} f_v^{(i)}) \nu_2(u,v) \tag{45b}$$

Proof. The proof is left to the reader and uses only the definition of a hidden Markov chain and the definition of f_u, c_u, γ_u, ν and ν_2. □

The previous formula (45b) for Γ shows that there may be two ways of creating a non-zero Γ. The first way — which is a bit trivial — uses non-zero contributions γ_u and corresponds to path increments $\mathbf{F}_k$ which already have a non-zero area $\pi_2(\mathbf{F}_k(1))$ in average: this is the case for example in the round-about model of section 1.3. However, this is impossible in the case of theorem 1.1 since geodesics in V are straight lines with zero area.

The second way is much more interesting since it may create a non-zero anomalous area drift Γ even in the context of theorem 1.1: it is based on the ν_2 contribution to Γ in (45b). In particular, it is absent from [4]: random walks are a particular case of hidden Markov walk for which E can be chosen to have cardinal 1 and $T_1 = 1$ a.s. and thus $\nu_k = 0$ for $k \geq 2$. One also checks easily that this term also vanishes for reversible Markov chain (R_n), for which $\nu_2(u,v) = \nu_2(v,u)$ for all $(u,v) \in E^2$.

2.3.3. Asymptotic fluctuations of the iterated occupation times and a remark on non-geometric rough paths

Property 1.2 gives an almost sure limit of the rescaled quantities $L_u((R_n))_{n \in \mathbb{N}}$; however, their fluctuations are more difficult to describe due to the quasi-shuffle property. We show here how to handle the question using our previous theorem 1.2. To make things simpler, we focus on the case where u is of length at most 2 and E is finite.

We now introduce the following quantities:

$$L_{u_1}^{\star}((R_n)_{0 \leq n < N}) = \sum_{0 \leq k < N} (\mathbf{1}_{R_k = u_1} - \nu(u_1)) = L_{u_1}((R_n)_{0 \leq n < N}) - N\nu(u_1)$$

If E is finite, we introduce the finite-dimensional vector space $\mathbb{R}^E$ with its canonical basis $(e_u)_{u\in E}$. We define the increments $F_k^\star = \sum_{u\in E}(\mathbf{1}_{R_k=u} - \nu(u))e_u$ and the walk:

$$X_N^\star = \sum_{u\in E} L_u^\star((R_n)_{0\le n<N}))e_u \tag{46}$$

Property 2.3. *The process $(R_n, X_n^\star)_{n\in\mathbb{N}}$ is a hidden Markov walk in $E\times\mathbb{R}^E$ with centred increments $(F_n^\star)$. If (R_n) is irreducible, positive recurrent with finite moments of the first return time T_1, then theorem 1.1 may be applied to describe the scaling limit of the process.*

This implies in particular, without any surprise, that $L_u((R_n)_{0\le n<N})/N$ has (joint) Gaussian fluctuations of order $N^{-1/2}$ around its a.s. limit $\nu(u)$. But this also gives result about the second iterated occupation times $L_{u_1u_2}((R_n)_{0\le n<N})$. To this purpose, we compute explicitly the iterated integral of the process $\iota_N(X^\star)$:

$$\begin{aligned}\frac{1}{N}\int_{0<s_1<s_2<1} d\iota_N(X^\star)(s_1)\otimes d\iota_N(X^\star)(s_2) &= \frac{1}{2N}\sum_{k=0}^{N-1} F_k^\star\otimes F_k^\star \\ &\quad + \frac{1}{N}\sum_{0\le k<l<N} F_k^\star\otimes F_l^\star\end{aligned}$$

The l.h.s. converges in law to $B_1\otimes B_1/2 + A_1^{\text{Lévy}} + \Gamma$. The first term of the r.h.s. converges a.s. and thus in law to a deterministic constant by the ergodic theorem. The second term of the r.h.s. is related to the modified iterated occupation time through:

$$\begin{aligned}\frac{1}{N}\sum_{0\le k<l<N} F_k^{(u_1)}F_l^{(u_2)} &= \frac{1}{N}\sum_{0\le k<l<N}(\mathbf{1}_{R_k=u_1}-\nu(u_1))(\mathbf{1}_{R_l=u_2}-\nu(u_2)) \\ &=: \frac{1}{N}L_{u_1u_2}^\star((R_n)_{0\le n<N})\end{aligned} \tag{47}$$

Relating $L_{u_1u_2}^\star$ to $L_{u_1u_2}$ gives the information about joint fluctuations of the collection of r.v. $L_{u_1u_2}((R_n)_{0\le n<N})$ and additive functionals of the $L_u((R_n)_{0\le n<N})$.

Remark on non-geometric rough paths. The construction of the present section is purely combinatorial and involves only integer numbers for the indices: one may wonder why integrals of the Donsker embedding $\iota_N(X^\star)$ should be preferred to the choice of iterated sums for the signature in $T^2(V)$ (and not $G^2(V)$ any more) such as:

$$\left(\sum_{0\le k<N} F_k^\star, \sum_{0\le k<l<N} F_k^\star\otimes F_l^\star\right)$$

There is indeed no reason to prefer one to another and it may depend on the context. The positive result is that there is no reason since both constructions differ by a term

$(1/2)\sum_{0\leq k<N} F_k^\star \otimes F_k^\star$, which belongs to the center of $T^2(V)$, is symmetric and whose limit is governed in the present case by the law of large numbers and is given by $(0, Kt)$ where K is a deterministic symmetric matrix. This additional term breaks the geometric rough path property but, as emphasized in exercise 2.14 (page 23) of [8] (and remarks disseminated in the corresponding chapter) or [10], this subtlety does not make a big difference from an analytic perspective.

3. Some open questions and extensions

We have seen how to build non-trivial rough paths above Brownian motion from very simple and intuitive processes such as hidden Markov chains. This raises various questions.

A key role in the emergence of a non-zero area anomaly Γ is played by the short-time correlations of the underlying Markov chain $(R_n)_{n\in\mathbb{N}}$. Exact renormalization on the time scale is due to the excursion decomposition: it would be interesting to generalize it to more general processes, such as α-mixing processes as described in [3].

On one hand, we have put restrictive hypothesis on the return times T_1 and the moments of the increments $\mathbf{F}_n$, so that the limit belongs to the Brownian universality class. On the other hand, rough paths structure may also describe Lévy processes as described in [6]. It would be interesting to build discrete time models that converge to such Lévy processes and contain all types of admissible anomalies such as Γ.

The generalization of theorem 1.1 to theorem 1.2 uses embeddings. This question of a discrete structure on top of piecewise paths is similar to the theory of piecewise-deterministic Markov processes (PDMP) in continuous time and it may be interesting to study space-time renormalized PDMP using the present rough path approach.

Rough paths are a particular case of much more general regularity structures as introduced by [8]. One may expect that such regularity structures may contain a wide class of anomalies (both in the sense of our area anomaly and in the sense of anomalies in field theory, i.e. a broken symmetry in the discretization or regularization restored by counter-terms in the continuous limit) and it may interesting to understand them from correlations in discrete models, as done in the present paper. In particular, anomalous enhanced Brownian motion is a particular case of translation of rough paths as described in [5]: the renormalization scheme described in this reference corresponds to our excursion-based renormalization. It would also be interesting in this context to introduce branched version of our iterated occupation times.

The other novelty is the introduction of iterated occupation time and the emergence of shuffle or quasi-shuffle products already at the discrete level. Such products also appear in other domains of algebra or combinatorics, for example in the theory of multiple zeta values or periods: it would be interesting to examine whether relevant Markov chains could be related to such theories.

Acknowledgements

D. S. is partially funded by the Grant ANR-14CE25-0014 (ANR GRAAL). D.S. and O.L. thank Lorenzo Zambotti for stimulating discussions and useful suggestions.

References

[1] S. Asmussen. *Applied Probability and Queues*. Springer-Verlag New York, 2003.
[2] L. E. Baum and T. Petrie. Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *Ann. Math. Statist.*, 37(6):1554–1563, 1966.
[3] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1999.
[4] E. Breuillard, P.K. Friz, and M. Huesmann. From random walks to rough paths. *Proc. Amer. Math. Soc.*, 137(10):3487–3496, 2009.
[5] Y. Bruned, I. Chevyrev, P.K. Friz and R. Preiß. A rough path perspective on renormalization. *J. Funct. Ana.*, 2019.
[6] I. Chevyrev. Random walks and Lévy processes as rough paths. *Probab. Theory Relat. Fields*, 170:891–932, 2018.
[7] A. Davie. Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express*, pages 009–40, 2007.
[8] P.K. Friz and M. Hairer. *A Course on Rough Paths - With an introduction to regularity structures*. Springer, 2014.
[9] P.K. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge University Press, 2010.
[10] M. Hairer and D. Kelly. Geometric versus non-geometric rough paths. *Annals of Institute Henri Poincaré*, 51(1):207–251, 2015.
[11] M. Hoffman. Quasi-shuffle products. *Journal of Algebraic Combinatorics*, 11:49–68, 2000.
[12] M. Hoffman and K. Ihara. Quasi-shuffle products revisited. *Journal of Algebra*, 481:293–326, 2017.
[13] J. Norris. *Markov Chains*. Cambridge University Press, 1997.
[14] A. Lejay and T. Lyons. On the importance of the Lévy area for studying the limits of functions of converging stochastic processes. Application to homogenization. *Current Trends in Potential Theory*, page 63–84, 2003.
[15] Y. Liu and S. Tindel. Discrete rough paths and limit theorems. arXiv:1707.01783, 2017.
[16] O. Lopusanschi and D. Simon. Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs. *Stochastic Processes and Applications*, 128(7):2404–2426, 2017.
[17] T. Lyons, M. Caruana, and Th. Lévy. *Differential Equations Driven by Rough Paths (École d'été de Probabilités de Saint-Flour XXXIV)*. Springer, 2004.
[18] T. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.
[19] L. Young. *The Theory of Integration*. The University Press, 1927.