JEMS

R.F. Bass · B.M. Hambly · T.J. Lyons

Extending the Wong-Zakai theorem to reversible Markov processes

Abstract We show how to construct a canonical choice of stochastic area for paths of reversible Markov processes satisfying a weak Hölder condition, and hence demonstrate that the sample paths of such processes are rough paths in the sense of Lyons. We further prove that certain polygonal approximations to these paths and their areas converge in p-variation norm. As a corollary of this result and standard properties of rough paths, we are able to provide a significant generalization of the classical result of Wong-Zakai on the approximation of solutions to stochastic differential equations. Our results allow us to construct solutions to differential equations driven by reversible Markov processes of finite p-variation with $p < 4$.

1. Introduction

1.1. Overview

Differential equations of the form

$$dy_t = f(y_t) \, dx_t \tag{1}$$

arise in many areas of mathematics. Cartan development [13], the path ordered exponential, the stochastic differential equations of Itô [12], control theory and even the standard linear filter of engineering provide examples of this formalism. These are all dynamical systems subject to an exogenous control. As the control is not necessarily smooth, it is important to give meaning to (1) for "rough paths".

R.F. Bass: Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA

B.M. Hambly: Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford, OX1 3LB, UK

T.J. Lyons: Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford, OX1 3LB, UK

Classical analysis gives (1) meaning if x has bounded variation. Itô calculus gives it meaning if x is a typical Brownian path in $\mathbb{R}^d$ and, more generally, for almost all semi-martingale paths. In this paper we use the theory of rough paths [17] and a generalization of a theorem of Lévy to give meaning to the equation for a much wider class of Markov processes with values in $\mathbb{R}^d$.

Suppose x_t is a smooth path in a vector space V and f is a linear map from this vector space to the Lipschitz vector fields on some state space M. Observe that $\frac{dx_t}{dt}$ is an element of V, and if $v \in V$, then $f(.)v$ is a vector field, $f(y)v$ is a tangent vector at y and (1) can be interpreted as the differential equation

$$\frac{dy_t}{dt} = f(y_t)\frac{dx_t}{dt}. \tag{2}$$

If we reparameterise time for the path x, then this produces a solution y that is simply a reparameterisation of the original solution. This shows immediately that in some sense, the parameter t and its differential "dt" are irrelevant. In addition it demonstrates how it is unnatural to expect that the path x_t should always be (piecewise) differentiable and that therefore one should be able to interpret (1) for more general paths.

The first serious extension of (1) to a class of non-differentiable paths x came in 1942-44 when Itô published [10], [11], papers that were to have a profound long term impact. He observed, using an integral constructed via an L^2 isometry, that (1) could be given meaning for *almost every* (multi-dimensional) Brownian path. That is, if W is Wiener measure on the continuous functions, then he constructed a map $I_f : C[0, 1] \to C[0, 1]$ so that for W-almost every path x the identity (1) holds. He used his result to construct new stochastic processes, and in so doing immediately demonstrated existence of solutions to wide classes of elliptic and parabolic partial differential equations. For over 50 years, his work, and the work of others in this area have profoundly affected engineering and more recently the world of finance.

Brownian paths are nowhere differentiable and it has been difficult to build a bridge between the classical interpretation of (2) and the meaning given to (1) by Itô. The first step is to build a slightly different "Stratonovich" integral and differential equation. Any reasonably smooth Itô differential equation can be recast as a Stratonovich equation and vice versa. The Stratonovich differential equation, like its smooth counterpart is invariant under changes in co-ordinates. The second important bridge between the Itô-Stratonovich stochastic differential equations and classical differential equations came in work of Wong and Zakai [23], [24]. They considered the one dimensional version of a stochastic differential equation of the form (1) and proved that if $x^{(n)}$ is a sequence of bounded variation paths converging uniformly to a Brownian path, x, then the solutions to the corresponding differential equations converge to the Stratonovich solution to the original stochastic differential equation. This work has been extended by Clark [4], for the more difficult multi-dimensional Brownian case, by Stroock and Varadhan [22], in a support theorem for SDE's and, much later, Chap. 6 of Ikeda and Watanabe [9] focused on the same issues, but in greater depth.

Despite this important bridge with classical systems there is still a significant limitation in the interpretation of (1) provided by Itô. The function I_f, introduced above is only defined up to null sets of paths. It does not give an analytic interpretation of (1) for any given path x. Moreover, these methods only apply to a class of stochastic processes known as semi-martingales.

One significant attempt to provide a pathwise definition of (1) for a class of paths rich enough to encompass the SDEs of Itô can be found in [17]. In the first instance one considers the map I_f as a function defined on the space $S(V)$ of paths of bounded variation in a Banach space V. Then [17] introduces a class of metrics on $S(V)$ of p-variation type (a precise definition is in Sect. 2) and proves that I_f is a uniformly continuous function on bounded sets in this metric. This approach allows x to be any "rough path" in the completion of $S(V)$. We denote this completion by $\Omega_{g,p}(V)$.[1]

Classical work of Lévy explains how one can regard almost every Brownian path as an element of $\Omega_{g,p}(\mathbb{R}^d)$ for $p > 2$, and so gives a pathwise interpretation for stochastic differential equations driven by Brownian motion and more generally by semi-martingales. In this paper we prove that a much wider class of diffusion processes in $\mathbb{R}^d$ can be regarded as having sample paths in $\Omega_{g,p}(\mathbb{R}^d)$; in particular we have a good interpretation of (1) for these processes. This approach provides an extension of the Wong-Zakai theorem as we prove that solutions to (1) can be viewed as the limit of a Cauchy sequence of polygonal approximations. The result is striking because processes such as diffusions with paths on connected fractal subsets of $\mathbb{R}^d$ are included and these are rather far from the semi-martingale paradigm considered by Itô. Our conclusions cannot (to our knowledge) be recovered using the usual Itô calculus methods.

Lévy also observed the convergence of the sequence obtained by taking the integral $\int \omega \circ dx^{(n)}$ of the area form $\omega = x^i dx^j - x^j dx^i$ against the piecewise linear dyadic approximations $x^{(n)}$ to a Brownian path x. The resulting limit is known as the Lévy area process and can be regarded as a solution to a differential equation of the form (1). In fact it is the antisymmetric 2 tensor component of the differential equation describing the noncommutative exponential function (4). Our main technical achievement is a generalisation of this result. We produce a Lévy area and show that it has finite p-variation for a class of reversible Markov processes satisfying a certain uniform Hölder condition.

1.2. Outline

The structure of the paper is as follows. Firstly we give an introduction to the theory of rough paths to provide the algebraic and analytic framework for our results. Then we show how to carry through the following program for constructing a canonical version of the Lévy area and higher order iterated integrals for a class of paths in $\mathbb{R}^d$ that are rougher than Brownian motion.

1. We begin by proving the existence of a canonical area between two *fixed* times for paths of a reversible Markov process with suitable α-variation, where the

---

[1] The notation $\Omega_p(V)$ is reserved for a larger space.

process is started from its stationary measure. Without loss of generality, the interval can be taken to be $[0, 2^{-n}]$. Our approach will be to take the dyadic polygonal approximations to the path of our process over this fixed time interval, compute the area for these approximations and then show the existence of the limit. This argument is essentially probabilistic, and exploits the reversible structure of the process and our assumptions about its sample path properties. At the same time as we construct this area we estimate its moments. A crucial part of our argument is the forward-backward martingale argument of Lyons and Zheng [20].

2. The next stage is to construct the area *process*. In other words we need to define the area simultaneously for all pairs of times $s < t \in [0, 1]$ and estimate the p-variation of the resulting two time parameter process. This is an essentially deterministic argument, and can be deduced from arguments that follow from Lemma 3 of Sect. 2 of [19]. The area and its p-variation are deterministically controlled by areas and increments defined over dyadic intervals alone. These form a countable collection and, with the moment conditions, give control over the whole path segment.

3. The previous steps show that (at least when $p < 3$) the Markov process, when enhanced by this new stochastic area process, is a rough path in the sense of [17]. It can be approximated by smooth paths in such a way that the paths and areas converge in the appropriate p-variation sense. However, to prove the Wong-Zakai theorem in the desired form we will need to prove that the polygonal approximations we used to construct the area over the fixed interval do actually converge in the p-variation sense along the whole interval. This follows quite easily from the work in [19] but is not stated there. The Wong-Zakai theorem can be deduced from this result using the main theorem in [17].

4. For $3 \leq p < 4$ the construction of the area, and the estimates on its variation still hold true, however, as they stand they are not sufficient to construct a rough path, or deduce the analogue of the Wong-Zakai theorem. As the roughness of the path increases, we require an increasing but finite number of higher order iterated integrals in order to solve differential equations of the form (1). For $3 \leq p < 4$ one requires a second "area like" integral to converge[2] as well. In Sect. 5 we will show that for $p < 4$ the same Hölder condition that we used previously, implies the existence of the required third order iterated integral. The proof of the analogue of the Wong-Zakai theorem then follows in the same way as before.

Finally we show that the conditions for existence of area hold for Dirichlet processes, that is diffusions associated with uniformly elliptic operators in divergence form; diffusion processes on nested fractals [16] and on Sierpinski carpets [1], [2]. We will also give an example of a non-reversible diffusion on the Sierpinski gasket for which the canonical choice of area does not exist. Finally we conclude with some open problems.

---

[2] For $2 \leq p < 3$ it is automatic from the theory of rough paths that the third order iterated integral will converge if the path and area converge.

We will sometimes need arbitrary constants in the paper and use the convention that c_i are constants that change though they remain fixed within a proof. For fixed constants appearing in Section n we will write $c_{n.i}$

2. Rough paths

2.1. The algebra of iterated integrals

The solution of differential equations driven by rough paths requires the construction of iterated integrals of the driving noise. The first iterated integral is just the increment of the path. Our aim is to construct the higher order iterated integrals of a path and before we do this it will be helpful to set up some notation. Iterated integrals have a strong algebraic content and it will be important to understand this a little.

Let S be the space of finite piecewise smooth path segments $\{x(u) \mid u \in [0, t], \ x(0) = 0\}$ in our vector space V, where we associate any path segment

$$\{x'(u) \mid u \in [s, t], \ x'(s) = a\} \quad \text{with} \quad \{x'(u + s) - a \mid u \in [0, t - s]\}.$$

There is a natural associative multiplication on S given by concatenation. Let $x, y \in S$ be paths defined on $[0, t]$, $[0, s]$, respectively; then

$$\begin{aligned} x * y &= z, \\ z(u) &= x(u), & u \in [0, t], \\ z(u) &= x(t) + y(u - t), & u \in [t, t + s], \end{aligned}$$

defines a path in S defined on the interval $[0, t + s]$. We now show that there exists a natural projection of S into the free nilpotent group of step ν, a projection that respects the multiplication $*$.

Let $T^{(\nu)}$ denote the truncated tensor algebra given by

$$T^{(\nu)} = \mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus V^{\otimes 3} \ldots \oplus V^{\otimes \nu},$$

with multiplication given by $\otimes$ with the convention that products producing tensors of degree greater than ν are zero. $T^{(\nu)}$ is an associative algebra. We give it a Lie bracket in the canonical way by defining $[\mathbf{x}, \mathbf{y}] = \mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x}$. This algebra is clearly nilpotent of degree ν (nested brackets of degree greater than ν always produce zero). The sub-Lie algebra $\mathcal{A}^{(\nu)}$ generated by V is of particular importance. It is a graded space

$$\mathcal{A}^{(\nu)} := V \oplus [V, V] \oplus [V, [V, V]] \ldots \oplus \overbrace{[V, [V, \ldots [V, V]]]}^{\nu \text{ times}},$$

where for example $[V, [V, V]]$ is the linear space spanned by all Lie brackets of triples from V. Because we are in the truncated tensor algebra the exponential series always converges, and it makes sense to consider $G^{(\nu)} = \exp \mathcal{A}^{(\nu)}$. Recall that for $\mathbf{a}, \mathbf{b} \in T^{(\nu)}$

$$e^{\mathbf{a}} e^{\mathbf{b}} = e^{m[\mathbf{a}, \mathbf{b}]},$$

where m[**a**, **b**] is the usual Campbell-Baker-Hausdorff formula:

$$m\,[\mathbf{a},\,\mathbf{b}] = \mathbf{a} + \mathbf{b} + \frac{1}{2}\,[\mathbf{a},\,\mathbf{b}] + \frac{1}{12}[\mathbf{a},\,[\mathbf{a},\,\mathbf{b}]] + \frac{1}{12}[\mathbf{b},\,[\mathbf{b},\,\mathbf{a}]] + \dots$$

From this it is clear that $e^{\mathbf{a}}e^{-\mathbf{a}} = 1$ and that $G^{(v)}$ is closed under tensor multiplication and hence a group. It is the free v-step nilpotent group over V, $\mathcal{A}^{(v)}$ is its Lie algebra and $T^{(v)}$ is the enveloping algebra. Moreover, there is an inverse map. Let $\mathbf{x} = 1 + \mathbf{y} \in G^{(v)}$, where $\mathbf{y}$ has 0 for a constant term. Define $\log \mathbf{x} = \log{[1 + \mathbf{y}]}$ via the series

$$\log{[1 + \mathbf{y}]} = \mathbf{y} - \frac{\mathbf{y}^2}{2} + \frac{\mathbf{y}^3}{3} - \frac{\mathbf{y}^4}{4} \dots$$

which is a finite series. Thus this series converges for any $\mathbf{x} \in G^{(v)}$ whose scalar component is one. The function log is then inverse to exp when restricted to $G^{(v)}$. From this we see that the correspondence between elements of $G^{(v)}$ and $\mathcal{A}^{(v)}$ is one to one and that we may regard them as two different co-ordinate charts for the group. In the one chart multiplication is given by tensor multiplication, in the other by the Campbell-Baker-Hausdorff formula.

The connection between the groups we defined above and the space of paths S comes via the iterated integrals. There is a natural map of S into $G^{(v)}$. Let $x \in S$ and define

$$\mathbf{x}_{0t}^r = \int \dots \int_{0 \le u_1 \le \dots \le u_r \le t} dx_{u_1} \dots dx_{u_r},$$

where of course $\mathbf{x}_{0t}^r \in V^{\otimes r}$. We may take the $\mathbf{x}_{0t}^r, r \le v$ together as a sequence to construct our map,

$$\phi\,(x) = \mathbf{x}_{0t}^{(v)} := \left(1, \mathbf{x}_{0t}^1, \mathbf{x}_{0t}^2, \dots, \mathbf{x}_{0t}^v\right).$$

Obviously $\phi(x) \in T^{(v)}$; in fact it takes its values in $G^{(v)}$ and the map ϕ is a homomorphism in the sense that

$$\phi\,(x) \otimes \phi\,(y) = \phi(x * y).$$

Both of these facts can easily be proved using the two elementary but key identities which together explain the significance of the iterated integrals

$$\mathbf{x}_{0t}^{(v)} = \mathbf{x}_{0u}^{(v)} \otimes \mathbf{x}_{ut}^{(v)}, \tag{3}$$

and the non-commutative exponential characterization:

$$d\mathbf{x}_{0t}^{(v)} = \mathbf{x}_{0t}^{(v)} \otimes dx_t, \tag{4}$$

$$\mathbf{x}_{00}^{(v)} = (1, \mathbf{0}, \mathbf{0}, \dots, \mathbf{0}).$$

In analogy with the argument above, it is obvious from (3) that the range of ϕ is closed under multiplication, and from (4) that the image of a path run backwards is the inverse of the image of the path run forwards.

2.2. The Lévy area

Initially we define the area for a path $x \in S$.

Definition 2.1. *The 2-parameter process* $A_{st} = (A_{st}^{ij})_{ij} \in V^{\otimes 2}$ *given by the antisymmetric part of the second iterated integral,*

$$A_{st}^{ij}(x) = \frac{1}{2} \int \int_{s < u_1 < u_2 < t} dx_{u_1}^i dx_{u_2}^j - dx_{u_1}^j dx_{u_2}^i, \quad s, t \in [0, L]$$

will be called the canonical area process associated to the piecewise smooth path x.

Fix a pair of co-ordinates i, j and so project the path into $\mathbb{R}^2$, then the closed directed planar curve defined by (x_u^i, x_u^j), $u \in [s, t]$ and the projection of the chord $\overrightarrow{x_t x_s}$ encloses an area $A_{st}^{ij}(x)$, providing orientation and multiplicity are taken into account. Note that $A(x)$ is dependent on the trajectory of the path x and not on the speed of passage, so that if $\tau(t)$ is a continuous time change

$$A_{st}(x \circ \tau) = A_{\tau(s)\tau(t)}(x).$$

It also depends on the inner product on the underlying space - but only in a simple fashion.

If the path x is nowhere differentiable, perhaps generated by a stochastic process, then, if the piecewise polygonal approximations to the path along the dyadic partition converge, it will be called the Lévy area. The dyadic partitions are not crucial to this result and any other geometrically increasing sequence of partitions that is deterministic and has intervals of equal length would produce the same analysis. We anticipate that the resulting area does not depend upon the particular partition but do not have a proof of this.

We can now interpret the ideas on rough paths in relation to this Lévy area. We have seen that there is a map from S into $G^{(v)}$ for any v. If we are to solve differential equations driven by the paths of a stochastic process it is clear from (4) that one must be able to construct the iterated integrals $\mathbf{x}_{0,t}^{(v)}$ associated to the path. The map into $G^{(1)}$ simply takes x to the increment $x_{0t} = x_t - x_0 = \int_{0 \le u \le t} dx_u$ and is clearly defined for all paths (and is obviously a homomorphism between concatenation of paths and addition in V). The map into $G^{(2)}$ is more easily seen when we change variables to Lie co-ordinates

$$\begin{aligned}
\mathbf{x}_{0t}^{(2)} &= \exp\left[\int_{0 \le u \le t} dx_u + \frac{1}{2} \int \int_{0 \le u_1 \le u_2 \le t} [dx_{u_1}, dx_{u_2}]\right] \\
&= \exp[x_{0t} + A_{0t}],
\end{aligned}$$

where $A_{0t} = \frac{1}{2} \int \int_{0 \le u_1 \le u_2 \le t} [dx_{u_1}, dx_{u_2}]$ is the area determined by the path. If our path is Brownian motion, this will be the Lévy area of the path.

Thus constructing the Lévy area is equivalent to constructing, for every subsegment of the path of our stochastic process, a lift into the group $G^{(2)}$ and solving the exponential differential equation at least in step 2. In addition, if we can establish

analytic conditions on these incremental lifts we will, by results in [17], ensure the existence of canonical choices of lift to all $G^{(\nu)}$ where $\nu > 2$ providing $p < 3$.

The analytic conditions are expressed in terms of p-variation and we discuss here the p-variation norm for a rough path. Suppose $\{x_t, \ t \in [0, T]\}$ is a piecewise smooth path in a Banach space V. It is said to have finite p-variation if

$$\omega(s, t) = \sup_{D \in \mathcal{D}} \sum_D |x_{u_{i+1}} - x_{u_i}|^p < \infty, \quad [s, t] \subset [0, T],$$

where $\mathcal{D}$ denotes the set of all partitions of $[s, t]$ and if $D \in \mathcal{D}$, then $D = \{s \leq \cdots < u_i < u_{i+1} < \cdots \leq t\}$ is a dissection of the time interval $[s, t]$. We can define p-variation for a multiplicative functional (rough path) $\mathbf{X}_{st} = (X_{st}, A_{st})$ in Lie coordinates (we write $X_{st} = X_t - X_s$ for the path increment).

Definition 2.2. *The p-variation distance between (X_{st}, A_{st}), $(\tilde{X}_{st}, \tilde{A}_{st})$ is given by $\omega_{\mathbf{X}\tilde{\mathbf{X}}}(s, t)$, where we define*

$$\omega_{\mathbf{X}\tilde{\mathbf{X}}}(s, t) = \sup_{D \in \mathcal{D}} \sum_D (|X_{u_{j-1}u_j} - \tilde{X}_{u_{j-1}u_j}|^p + |A_{u_{j-1}u_j} - \tilde{A}_{u_{j-1}u_j}|^{p/2}).$$

As the area is the antisymmetric part of the second iterated integral, controlling $\omega_{\mathbf{X}0}$ using area is equivalent to controlling it using the second iterated integral X^2. We can extend our definition of p-variation to multiplicative functionals in $G^{(\nu)}$, up to any degree ν, as

$$\omega'_{\mathbf{X}, \tilde{\mathbf{X}}}(s, t) = \sup_{D \in \mathcal{D}} \sum_D \sum_{m=1}^{\nu} \left| X^m_{u_{j-1}u_j} - \tilde{X}^m_{u_{j-1}u_j} \right|^{p/m}.$$

We can also write our definition of p-variation distance in Lie coordinates and state the following fact relating the p-variation in ordinary and Lie coordinates. If we define, for $\mathbf{X}, \tilde{\mathbf{X}} \in G^{(\nu)}$ with constant term 1,

$$\omega_{\mathbf{X}\tilde{\mathbf{X}}}(s, t) = \sup_{D \in \mathcal{D}} \sum_D \sum_{m=1}^{\nu} \left| (\log \mathbf{X}_{u_{j-1}u_j})^m - (\log \tilde{\mathbf{X}}_{u_{j-1}u_j})^m \right|^{p/m},$$

then there is a constant $c_{2.1}$ such that

$$c_{2.1}^{-1} \omega_{\mathbf{X}\tilde{\mathbf{X}}}(s, t) \leq \omega'_{\mathbf{X}\tilde{\mathbf{X}}}(s, t) \leq c_{2.1} \omega_{\mathbf{X}\tilde{\mathbf{X}}}(s, t).$$

We say that $\mathbf{X}$ has finite p-variation if $\omega(s, t) = \omega_{\mathbf{X}0}(s, t)$ or $\omega'(s, t)$ is finite.

Once we have these ideas we can formulate a key result of [17], contained in Theorem 2.2.1.

Theorem 2.3 (Exponential extension). *Let $\mathbf{x}_{0t}^{(\nu)}$ be a continuous two time function satisfying the algebraic condition (3) and the analytic condition that it has finite p-variation for all $p > \nu$. Then, for any N, there is a unique extension of $\mathbf{x}_{0t}^{(\nu)}$ to $\mathbf{x}_{0t}^{(N)}$ satisfying the algebraic and analytic conditions with the same p. Moreover the map is continuous.*

Indeed it can be shown that the extension map is uniformly continuous on bounded sets. In particular if we can show that the iterated integrals for our polygonal path approximations are Cauchy in this p-variation metric, the result provides a construction of all iterated integrals of our stochastic path.

Finally we can give a meaning to the solution of the stochastic differential equation (1). Let X be a path which has finite p-variation for all $p > \nu$. If there is a sequence of piecewise smooth approximations x_n, to the path X, which are Cauchy in the p-variation metric involving the first $[\nu]$ iterated integrals of X, and the vector fields are Lipschitz of smoothness $\gamma > \nu$, then the sequence of solutions y_n to (1) is also Cauchy in the p-variation metric for the same $p > \nu$ and hence we can define the Stratonovich solution Y to the stochastic differential equation.

3. Existence of area

In this section and for the rest of the paper, we will assume that we have a continuous path generated by a reversible Markov process and that it lies in $V = \mathbb{R}^d$. We note that this path may be a function of the underlying Markov process. Now fix a time interval and consider a sequence of polygonal approximations to the path. Our aim is to show that the sequence of areas associated to these polygonal approximations converge.

Now consider a general continuous path $\{x_u; u \in [0, T]\}$. Let $t_i^n = i2^{-n}T$, then $t^n = \{t_i^n : 0 \le i \le 2^n\}$ is the n-th dyadic partition of the interval $[0, T]$. Let $x_i^n = x_{t_i^n}$ be the path sampled at these dyadic times. We may extend the path x^n to a piecewise linear path by interpolation:

$$x_{u2^{-n}T}^n = (i + 1 - u)x_i^n + (u - i)x_{i+1}^n, \ i < u < i + 1, \quad 0 \le i < 2^n.$$

The paths x^{n+1} and x^n agree at the times in the partition t^n. As one goes from x^n to x^{n+1} one can associate with each linear segment of x^n, two edges from x^{n+1} making an oriented triangle, and by considering all the intervals in t^n we obtain a sequence of 2^n oriented triangles (see Fig. 1). If x is Markov and we define the filtration generated by the n-th approximation,

$$\mathcal{F}_n = \sigma\left(x_{t_i^n} : i \le 2^n\right),$$

then we note that the triangles are independent random variables given $\mathcal{F}_n$.

Now fix attention on the path $\{x_t; t \in [0, T]\}$ and let $A^n = A_{0T}(x^n)$ be the area generated by the n-th successive approximation to the path. We study the convergence of A^n. Now $A^0 = 0$, and $A^n = \sum_{j=1}^n A^j - A^{j-1}$ where $A^{n+1} - A^n = \sum_{i=0}^{2^n-1} A_{t_i^n t_{i+1}^n}(x^{n+1})$ is the sum of the areas in the 2^n triangles mentioned above and illustrated in the schematic picture shown as Fig. 1.

Our aim is to show that the limit of

$$A^{k+1} = \sum_{n=1}^k \sum_{i=0}^{2^n-1} A_{t_i^n t_{i+1}^n}(x^{n+1}), \tag{5}$$
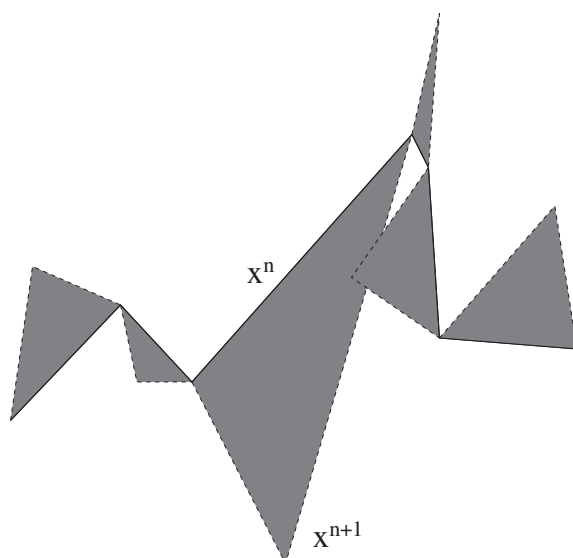
as $k \to \infty$, exists.

Fig. 1. The increase in area due to the addition of the next dyadic approximation to the path

A naive approach to the construction of the area would be to compute the area of each triangle and then add them all together. This would fail if the path were Brownian - in the sense that the sum is not absolutely convergent. However, it is obvious from reflection symmetry that in the Brownian case the elements $A_{t_i^n t_{i+1}^n}(x^{n+1})$ satisfy $E\left(A_{t_i^n t_{i+1}^n}(x^{n+1})|\mathcal{F}_n\right) = 0$ and thus A^n is an $\mathcal{F}_n$ martingale. Convergence of the area then follows easily as the bracket process is seen to converge very rapidly indeed. It seems essential to the argument that one uses some cancelation, and we observe in this paper that, in the case where x is Markov, reversibility seems to be the key.

Assumption 3.1. *Assume the path x is generated by a continuous reversible Markov process $\{X_t; \ t \in (0, T]\}$, started from a finite stationary measure μ which we take to have mass one[3], on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$.*

We now make the assumption that the β moments of the increments of the process are controlled.

Assumption 3.2. *There exist constants $\alpha < 4$ and $c_{3.1}(\beta)$ such that for all $\beta \geq 1$,*

$$\sup_x E^x|X_0 - X_t|^\beta \leq c_{3.1}(\beta)t^{\beta/\alpha}, \ \ \forall 0 < t \leq 1.$$

We denote by E^x the expectation of the process started from a point x and hence this assumption is a uniform Hölder condition on the paths. We will see that this

---

[3] As the problems we are interested in are of a local nature, we assume this measure is finite. Lebesgue measure, and Brownian motion in $\mathbb{R}^d$ could be made to fit our hypotheses by quotienting onto a torus etc.

implies that our path has finite p-variation for all $p > \alpha$. These two assumptions will be assumed until Sect. 6, where we will show that they hold for a variety of examples. Our main theorem in this section is the following.

Theorem 3.3. *For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumptions 3.1, 3.2, the area $A_{0T}(X)$ up to time T exists (in the sense of (5)) and there are constants $c_{3.2}(\beta)$ such that for all $\beta \geq 1$,*

$$EA_{0T}^{\beta}(X) \leq c_{3.2}(\beta)T^{2\beta/\alpha}.$$

We prove this result through several preliminary lemmas. Consider the following filtrations

$$\mathcal{G}_t = \sigma(X_u; 0 \leq u \leq t), \quad \tilde{\mathcal{G}}_t = \sigma(X_u; T - t \leq u \leq T), \tag{6}$$

determined by the path run forward in time and the path run backward in time respectively.

Firstly, in the spirit of [20], we show that there is a forward-backward martingale decomposition for functions of reversible Markov processes. Let $\pi(x, y)$ denote a path from x to y.

Lemma 3.4. *Let $\psi : C(\mathbb{R} : V) \to \mathbb{R}$ be a measurable function satisfying $\psi(\pi(x, y)) = -\psi(\pi(y, x))$, for all paths $\pi(x, y) \in C(\mathbb{R} : V)$ and $x, y \in V$. For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumption 3.1, and $E^x(|\psi(\pi(X_s, X_t))||\mathcal{G}_s), E^x(|\psi(\pi(X_s, X_t))||\tilde{\mathcal{G}}_{T-t}) < \infty$ for all $s, t \in [0, T]$, we have a forward-backward martingale decomposition,*

$$\sum_{i=0}^{2^n-1} \psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) = \frac{1}{2}M_{2^n}^n + \frac{1}{2}\tilde{M}_0^n - E(\psi(\pi(X_{t_{2^n}^n}, X_{t_{2^n-1}^n}))|\tilde{\mathcal{G}}_0)$$
$$+ E(\psi(\pi(X_{t_0^n}, X_{t_1^n}))|\mathcal{G}_0) \tag{7}$$

where the martingales are given by

$$M_k^n = \sum_{i=0}^{k-1} \left(\psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) - E(\psi(\pi(X_{t_i^n}, X_{t_{i+1}^n}))|\mathcal{G}_{t_i^n})\right),$$

and

$$\tilde{M}_k^n = -\sum_{i=k+1}^{2^n} \left(\psi(\pi(X_{t_i^n}, X_{t_{i-1}^n})) - E(\psi(\pi(X_{t_i^n}, X_{t_{i-1}^n}))|\tilde{\mathcal{G}}_{T-t_i^n})\right).$$

Proof. By the definition of the function ψ, we have that for any path X, and any n,

$$\sum_{i=0}^{2^n-1} \psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) = -\sum_{i=0}^{2^n-1} \psi(\pi(X_{t_{i+1}^n}, X_{t_i^n}))$$
$$= \frac{1}{2}\sum_{i=0}^{2^n-1} \psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) - \frac{1}{2}\sum_{i=0}^{2^n-1} \psi(\pi(X_{t_{i+1}^n}, X_{t_i^n})).$$
$$\tag{8}$$

The process is started from its stationary measure and hence the path $\{X_t; t \in [0, T]\}$, observed going forwards or backwards from a particular point $X_{t_i^n}$, satisfies

$$E(\psi(\pi(X_{t_i^n}, X_{t_{i+1}^n}))|\mathcal{G}_{t_i^n}) = E(\psi(\pi(X_{t_i^n}, X_{t_{i-1}^n}))|\tilde{\mathcal{G}}_{T-t_i^n}), \quad 1 \le i \le 2^n - 1.$$

By (8), using the usual Doob-Meyer decomposition for the forward and the backward processes, we obtain the forward and backward martingales M_n, $\tilde{M}_n$. We include the final two terms in (7) to cater for the discrepancies at the beginning and end of the path. $\square$

The existence of area will be established by showing the existence of the limit in (5). As the area formed by a curve depends on orientation, we see that the area generated by the path traversed backwards is just the negative of the area generated by the path traversed forwards. Thus, if we define

$$\psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) = A_{t_i^n t_{i+1}^n}(x^{n+1}),$$

then we can write

$$A^{k+1} = \sum_{n=1}^{k} \sum_{i=0}^{2^n-1} A_{t_i^n t_{i+1}^n}(x^{n+1}),$$

and apply the lemma.

Now we make the simple geometric observation that the area of a triangle, $A_{t_i^n t_{i+1}^n}(x^{n+1})$ is bounded above by the product of the lengths of the vectors which define it,

$$\left|A_{t_i^n t_{i+1}^n}(x^{n+1})\right| \le \frac{1}{2}\left|x_i^n - x_{2i+1}^{n+1}\right| \left|x_{2i+1}^{n+1} - x_{i+1}^n\right|. \tag{9}$$

Lemma 3.5. *Under Assumptions 3.1, 3.2, we have*

$$E\left(\left|A_{t_i^n t_{i+1}^n}(x^{n+1})\right| \big| \mathcal{G}_{t_i^n}\right) \le \frac{1}{2}c_{3.1}(2)T^{2/\alpha}2^{-2n/\alpha}. \tag{10}$$

Proof. By (9), we can bound the random variable $|A_{t_i^n t_{i+1}^n}(x^{n+1})|$ by the product of the lengths of the increments of the path x^{n+1}.

$$E\left(\left|A_{t_i^n t_{i+1}^n}(x^{n+1})\right| \big| \mathcal{G}_{t_i^n}\right) \le \frac{1}{2}E\left(\left|x_i^n - x_{2i+1}^{n+1}\right|\left|x_i^n - x_{i+1}^n\right| \big| \mathcal{G}_{t_i^n}\right).$$

Now apply Cauchy-Schwarz and our hypothesis on the increments,

$$E\left(\left|A_{t_i^n t_{i+1}^n}(x^{n+1})\right| \big| \mathcal{G}_{t_i^n}\right) \le \frac{1}{2}\left(E\left(\left|x_i^n - x_{2i+1}^{n+1}\right|^2 \big| \mathcal{G}_{t_i^n}\right)E\left(\left|x_i^n - x_{i+1}^n\right|^2 \big| \mathcal{G}_{t_i^n}\right)\right)^{1/2}.$$

$$\le \frac{1}{2}c_{3.1}(2)\left((T2^{-(n+1)})^{2/\alpha}(T2^{-n})^{2/\alpha}\right)^{1/2}$$

$$\le \frac{1}{2}c_{3.1}(2)T^{2/\alpha}2^{-2n/\alpha}, \tag{11}$$

as desired. $\square$

Proof of Theorem. In order to prove that, in (5), the limit A^m exists as $m \to \infty$ we will show that it has finite mean. For this we need to control the brackets of the martingales in the forward-backward martingale decomposition (7). The bracket processes are given by

$$[M^n]_k = \sum_{i=0}^{k-1} \left(A_{t_i^n t_{i+1}^n}(x^{n+1}) - E\big(A_{t_i^n t_{i+1}^n}(x^{n+1})\big|\mathcal{G}_{t_i^n}\big) \right)^2,$$

and

$$[\tilde{M}^n]_k = \sum_{i=k+1}^{2^n} \left(A_{t_i^n t_{i-1}^n}(x^{n+1}) - E\big(A_{t_i^n t_{i-1}^n}(x^{n+1})\big|\tilde{\mathcal{G}}_{T-t_i^n}\big) \right)^2.$$

Using the fact that

$$|x_i^n - x_{2i+1}^{n+1}|^2 |x_{2i+1}^{n+1} - x_{i+1}^n|^2 \le \frac{1}{2}\big(|x_i^n - x_{2i+1}^{n+1}|^4 + |x_{2i+1}^{n+1} - x_{i+1}^n|^4\big),$$

Lemma 3.5 and (9), we have pathwise control on the bracket of the forward martingale, as

$$[M^n]_{2^n} \le 2 \sum_{i=0}^{2^n-1} \big|A_{t_i^n t_{i+1}^n}(x^{n+1})\big|^2 + 2 \sum_{i=0}^{2^n-1} E\big(\big|A_{t_i^n t_{i+1}^n}(x^{n+1})\big|\,\big|\mathcal{G}_{t_i^n}\big)^2$$

$$\le \frac{1}{4} \sum_{i=0}^{2^{n+1}-1} \big|x_i^{n+1} - x_{i+1}^{n+1}\big|^4 + c_1 T^{4/\alpha} 2^n 2^{-4n/\alpha}. \tag{12}$$

Taking expectations we have

$$E[M^n]_{2^n} \le \frac{1}{4} \sum_{i=0}^{2^{n+1}-1} E\big|x_i^{n+1} - x_{i+1}^{n+1}\big|^4 + c_1 T^{4/\alpha} 2^{n(1-4/\alpha)}$$

$$\le c_2 T^{4/\alpha} 2^{n(1-4/\alpha)}.$$

We now need to show that we can add up each estimate for the area and keep control on the whole sum,

$$A = \sum_n \left(\frac{1}{2} M_{2^n}^n + \frac{1}{2} \tilde{M}_{2^n}^n + E\big(A_{t_{2^n}^n t_{2^n-1}^n}(x^{n+1})\big|\tilde{\mathcal{G}}_0\big) - E\big(A_{t_0^n t_1^n}(x^{n+1})\big|\mathcal{G}_0\big) \right). \tag{13}$$

We do this by establishing that the moments of each of the four sums is finite. The $L^\beta(\Omega, \mathbb{P})$ norm of the area is estimated by

$$\|A\|_\beta \le \sum_n \big\|M_{2^n}^n\big\|_\beta + \sum_n \big\|\tilde{M}_{2^n}^n\big\|_\beta +$$

$$\sum_n \big\|E\big(A_{t_{2^n}^n t_{2^n-1}^n}(x^{n+1})\,\big|\tilde{\mathcal{G}}_0\big)\big\|_\beta + \sum_n \big\|E\big(A_{t_0^n t_1^n}(x^{n+1})\big|\mathcal{G}_0\big)\big\|_\beta. \tag{14}$$

For the forward martingale we use Burkholder's inequality and (12) to show, provided $\beta \geq 2$,

$$
\begin{aligned}
E\left(M_{2^n}^n\right)^\beta &\leq c_3 E\left([M^n]_{2^n}^{\beta/2}\right) \\
&\leq c_4 E\left(\sum_{i=0}^{2^{n+1}-1} \left|x_i^{n+1} - x_{i+1}^{n+1}\right|^4 + c_1 T^{4/\alpha} 2^n 2^{-4n/\alpha}\right)^{\beta/2} \\
&\leq c_5 E\left(\sum_{i=0}^{2^{n+1}-1} \left|x_i^{n+1} - x_{i+1}^{n+1}\right|^4\right)^{\beta/2} + c_6 T^{2\beta/\alpha} 2^{n(1-4/\alpha)\beta/2} \\
&\leq c_5 2^{(n+1)(\beta/2-1)} \sum_{i=0}^{2^{n+1}-1} E\left|x_i^{n+1} - x_{i+1}^{n+1}\right|^{2\beta} + c_6 T^{2\beta/\alpha} 2^{n(1-4/\alpha)\beta/2} \\
&\leq c_7 T^{2\beta/\alpha} 2^{n(1-4/\alpha)\beta/2}.
\end{aligned}
$$

Then, for $\alpha < 4$, we have

$$
\sum_n \left\|M_{2^n}^n\right\|_\beta \leq c_7 \sum_n T^{2/\alpha} 2^{n(1-4/\alpha)/2} = c_8 T^{2/\alpha}.
$$

This is finite and decays exponentially quickly giving the almost sure existence as well.

The same approach shows that the backward martingale part is also well behaved. We can also control the correction terms using the estimate obtained in (10). For the first backward term

$$
\sum_n \left\|E\left(\left|A_{t_{2^n}^n t_{2^n-1}^n}(x^{n+1})\right| \mid \tilde{\mathcal{G}}_0\right)\right\|_\beta \leq \sum_n \frac{1}{2} c_{3.1}(2) T^{2/\alpha} 2^{-2n/\alpha} \leq c_9 T^{2/\alpha},
$$

and similarly for the first forward term.

Thus we have that all the sums are convergent and we have the almost sure existence of A_{0T}. Replacing the terms in (14) to see that $\|A\|_\beta \leq c_{10} T^{2/\alpha}$, and using monotonicity of the norm, gives the required moment estimate. $\quad\square$

Recall that A_{0T}^n is the area of the piecewise polygonal path x^n. From the above results we can easily deduce a moment estimate on the rate of convergence of these approximating sums.

Corollary 3.6. *There exists a constant $c_{3.3}(\beta)$ such that for all $\beta \geq 1$,*

$$
E\left|A_{0T} - A_{0T}^n\right|^\beta \leq c_{3.3}(\beta) T^{2\beta/\alpha} 2^{n\beta(1-4/\alpha)/2}.
$$

4. Area between an arbitrary pair of times and finite p-variation

We have proved the existence of an area random variable for a fixed pair of times. We now wish to establish the existence and continuity of a two parameter 2-tensor valued process A_{st}, which we will call the area process for a path of our underlying continuous process $\{X_t : t \in [0, T]\}$. In order to be able to establish convergence results we will establish that $\mathbf{X}_{st} = (X_{st}, A_{st})$ has finite p-variation for suitable p. We will follow the same approach as that used in [8], Sect. 4, but remark that, in common with [19], there is less work to be done as we have the area between any pair of dyadic times.

Our aim is to construct the area process and control its p-variation using dyadic intervals. First we fix an interval $[s, t]$ and define inductively, a series of disjoint dyadic intervals as follows. Let n_0 be the first integer such that there exists a $k \in \mathbb{Z}$ such that

$$\left[k2^{-n_0}, (k+1)\,2^{-n_0}\right] \subset [s, t],$$

and define $I_0 = \left[k2^{-n_0}, (k+1)\,2^{-n_0}\right]$. For $n > 0$ define I_n to be the unique dyadic interval of length $2^{-(n_0+n)}$ to the right of $\cup_{i<n} I_i$, if it is contained in $[s, t]$, and the singleton

$$\left[\sup \cup_{i<n} I_i, \sup \cup_{i<n} I_i\right],$$

if it is not. Define I_n for $n < 0$ in a similar way to be the intervals on the left.

We use this dyadic division to define the rough path for any pair of times s, t. As the path is an element in $G^{(\nu)}$ we can write

$$\mathbf{X}_{[s,t]} = \lim_{n \to \infty} \mathbf{X}_{I_{-n}} \otimes \mathbf{X}_{I_{-n+1}} \otimes \ldots \otimes \mathbf{X}_{I_0} \otimes \ldots \otimes \mathbf{X}_{I_{n-1}} \otimes \mathbf{X}_{I_n}. \qquad (15)$$

Clearly the series is eventually constant if s and t are dyadic, and so by the multiplicative property of $\mathbf{X}$ gives a compatible definition for those times. We will state our results in some generality and for the next result regard $\mathbf{X} = (1, X^1, X^2, \ldots)$ as an element of the full tensor algebra. Now observe that for $n > 0$,

$$\mathbf{X}_{I_{-n}} \otimes \mathbf{X}_{I_{-n+1}} \otimes \ldots \otimes \mathbf{X}_{I_0} \otimes \ldots \otimes \mathbf{X}_{I_{n-1}} \otimes \mathbf{X}_{I_n}$$

$$= \left(1, \sum_{i=-n}^{n} X_{I_i}, \sum_{-n \leq i < j \leq n} X_{I_i} \otimes X_{I_j} + \sum_{i=-n}^{n} X_{I_i}^2, \ldots, \right.$$

$$\left. \sum_{\substack{k=1 \\ }}^{m} \sum_{\substack{(r_1,\ldots,r_k) \\ \sum_1^k r_j = m}} \sum_{-n \leq i_1 < \cdots < i_k \leq n} X_{I_{i_1}}^{r_1} \otimes \cdots \otimes X_{I_{i_k}}^{r_k}, \ldots \right). \qquad (16)$$

We will write

$$\eta(X; r, p) = \sum_n n^\gamma \sum_{i=0}^{2^n-1} \left| X_{t_i^n t_{i+1}^n}^r \right|^{p/r},$$

where $\gamma > p - 1$, as this will be useful for the dyadic control on the p-variation of the r-th component of $\mathbf{X}$. We will abuse this notation slightly and write, for the area (the antisymmetric part of X^2),

$$\eta(A; 2, p) = \sum_n n^\gamma \sum_{i=0}^{2^n-1} |A_{t_i^n t_{i+1}^n}|^{p/2}.$$

Lemma 4.1. *There exists a universal constant* $c_{4.1}$, *depending on* p, *such that for a* $\gamma > p - 1$,

$$\omega'_{\mathbf{X}\tilde{\mathbf{X}}}(0, 1) \leq$$

$$c_{4.1} \sum_{m=1}^{\nu} \sum_{k=1}^{m} \sum_{\substack{(r_1,\dots,r_k) \\ \sum_1^k r_j = m}} \sum_{l=1}^{k} \eta(X - \tilde{X}; r_l, p)^{r_l/m} \prod_{\substack{j=1 \\ j \neq l}}^{k} \left(\eta(X; r_j, p) + \eta(\tilde{X}; r_j, p) \right)^{r_j/m}.$$

Proof. For the case of $\omega(0, 1)$ with $\nu = 2$, this is equation (2) of [19]. We extend this result to $\omega'_{\mathbf{X}\tilde{\mathbf{X}}}(0, 1)$ for any finite ν. In order to deal with higher order terms we will make use of the non-commutative identity

$$\bigotimes_{i=1}^{n} a_i - \bigotimes_{i=1}^{n} b_i = \sum_{j=1}^{n} \bigotimes_{i=1}^{j-1} a_i (a_j - b_j) \bigotimes_{i=j+1}^{n} b_i, \tag{17}$$

for any sequences $\{a_i\}$, $\{b_i\}$, where we define $\bigotimes_{i=1}^{0} a_i = 1$. Consider our dyadic partition of the interval $[s, t]$ given by $\{I_i\}_{i=-\infty}^{\infty}$. In order to control the p-variation over this interval, we compare the m-th order term in the two multiplicative functionals using (16),

$$X_{st}^m - \tilde{X}_{st}^m = \lim_{n \to \infty} \sum_{k=1}^{m} \sum_{\substack{(r_1,\dots,r_k) \\ \sum_1^k r_j = m}} \sum_{-n \leq i_1 < \cdots < i_k \leq n} X_{I_{i_1}}^{r_1} \otimes \cdots \otimes X_{I_{i_k}}^{r_k}$$

$$- \left(\sum_{k=1}^{m} \sum_{\substack{(r_1,\dots,r_k) \\ \sum_1^k r_j = m}} \sum_{-n \leq i_1 < \cdots < i_k \leq n} \tilde{X}_{I_{i_1}}^{r_1} \otimes \cdots \otimes \tilde{X}_{I_{i_k}}^{r_k} \right). \tag{18}$$

Now apply (17) to the terms in (18),

$$\sum_{-\infty < i_1 < \cdots < i_k < \infty} \left(X_{I_{i_1}}^{r_1} \otimes \cdots \otimes X_{I_{i_k}}^{r_k} - \tilde{X}_{I_{i_1}}^{r_1} \otimes \cdots \otimes \tilde{X}_{I_{i_k}}^{r_k} \right)$$

$$= \sum_{-\infty < i < \infty} \sum_{j=1}^{k} \sum_{-\infty < i_1 < \dots i_{j-1} < i_{j+1} < \cdots < i_k < \infty} \prod_{l=1}^{j-1} X_{I_{i_l}}^{r_l} \left(X_{I_i}^{r_j} - \tilde{X}_{I_i}^{r_j} \right) \prod_{l=j+1}^{k} \tilde{X}_{I_{i_l}}^{r_l}$$

$$\leq \sum_{-\infty < i < \infty} \sum_{l=1}^{k} |X_{I_i}^{r_l} - \tilde{X}_{I_i}^{r_l}| \prod_{\substack{j=1 \\ j \neq l}}^{k} \sum_{-\infty < i < \infty} \left(|X_{I_i}^{r_j}| + |\tilde{X}_{I_i}^{r_j}| \right).$$

Hölder's inequality can be applied to give the existence of a constant $c_1(r, p, \gamma)$ such that, if $\gamma > p - 1$, then

$$(\sum_{i=-\infty}^{\infty} |X_{I_i}^r|)^p \leq c_1(r, p, \gamma) \sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma |X_{I_i}^r|^p.$$

Using this twice we have

$$\left| X_{st}^m - \tilde{X}_{st}^m \right|^{p/m} \leq \sum_{k=1}^{m} \sum_{\substack{(r_1,...,r_k) \\ \sum_1^k r_j = m}} c_2(m, p, \gamma) \sum_{l=1}^{k} \left(\sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left| X_{I_i}^{r_l} - \tilde{X}_{I_i}^{r_l} \right|^{p/m} \right.$$

$$\left. \prod_{\substack{j=1 \\ j \neq l}}^{k} \sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left(\left| X_{I_i}^{r_j} \right|^{p/m} + \left| \tilde{X}_{I_i}^{r_j} \right|^{p/m} \right) \right).$$

Thus the variation over any interval can be expressed in terms of the variation over dyadics.

Consider an arbitrary partition. Firstly we use Hölder's inequality for the product of k terms and then sum over all disjoint dyadic intervals to estimate the p-variation,

$$\sup_{0 < u_1 < ... < u_r < 1} \sum_{\mathcal{D}} \left| X_{u_h u_{h+1}}^m - \tilde{X}_{u_h u_{h+1}}^m \right|^{p/m}$$

$$\leq \sup_{0 < u_1 < ... < u_r < 1} \sum_{\mathcal{D}} \sum_{k=1}^{m} \sum_{\substack{(r_1,...,r_k) \\ \sum_1^k r_j = m}} c_2 \sum_{l=1}^{k} \left(\sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left| X_{I_i(u_h, u_{h+1})}^{r_l} - \tilde{X}_{I_i(u_h, u_{h+1})}^{r_l} \right|^{p/m} \right.$$

$$\left. \prod_{\substack{j=1 \\ j \neq l}}^{k} \sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left(\left| X_{I_i(u_l, u_{l+1})}^{r_j} \right|^{p/m} + \left| \tilde{X}_{I_i(u_h, u_{h+1})}^{r_j} \right|^{p/m} \right) \right)$$

$$\leq \sup_{\mathcal{D}} \sum_{k=1}^{m} \sum_{\substack{(r_1,...,r_k) \\ \sum_1^k r_j = m}} c_2 \left(\left(\sum_{\mathcal{D}} \left(\sum_{l=1}^{k} \sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left| X_{I_i(u_h, u_{h+1})}^{r_l} - \tilde{X}_{I_i(u_h, u_{h+1})}^{r_l} \right|^{p/m} \right)^{m/r_l} \right)^{r_l/m} \right.$$

$$\left. \prod_{\substack{j=1 \\ j \neq l}}^{k} \left(\sum_{\mathcal{D}} \left(\sum_{i=-\infty}^{\infty} (|i| + 1)^\gamma \left(\left| X_{I_i(u_h, u_{h+1})}^{r_j} \right|^{p/m} + \left| \tilde{X}_{I_i(u_h, u_{h+1})}^{r_j} \right|^{p/m} \right) \right)^{m/r_j} \right)^{r_j/m} \right)$$

$$\leq \sum_{k=1}^{m} \sum_{\substack{(r_1,...,r_k) \\ \sum_1^k r_j = m}} c_3 \left(\sum_{l=1}^{k} \sum_{n=1}^{\infty} n^\gamma \sum_{h=0}^{2^{n-1}-1} \left| X_{t_h^n t_{h+1}^n}^{r_l} - \tilde{X}_{t_h^n t_{h+1}^n}^{r_l} \right|^{p/r_l} \right)^{r_l/m}$$

$$\prod_{\substack{j=1 \\ j \neq l}}^{k} \left(\sum_{n=1}^{\infty} n^\gamma \sum_{h=0}^{2^{n-1}-1} \left| X_{t_h^n t_{h+1}^n}^{r_j} \right|^{p/r_j} + \left| \tilde{X}_{t_h^n t_{h+1}^n}^{r_j} \right|^{p/r_j} \right)^{r_j/m}.$$

By the definition of η we have the desired result. $\qquad\square$

Corollary 4.2. *There exists a constant $c_{4.2}$ such that for all p*

$$\omega(0,1) \le c_{4.2} \sum_{m=1}^{\nu} \eta(X; m, p). \tag{19}$$

Proof. Rewriting Lemma 4.1 with $\tilde{\mathbf{X}} = 0$, we have

$$\omega(0,1) \le c_{4.1} \sum_{m=1}^{\nu} \sum_{k=1}^{m} \sum_{\substack{(r_1,\dots,r_k) \\ \sum_1^k r_j = m}} \sum_{l=1}^{k} \eta(X; r_l, p)^{r_l/m} \prod_{\substack{j=1 \\ j \ne l}}^{k} \eta(X; r_j, p)^{r_j/m}.$$

Now observe that if $\sum_{j=1}^{k} r_j = m$, then

$$\eta(X; r_l, p)^{r_l/m} \prod_{\substack{j=1 \\ j \ne l}}^{k} \eta(X; r_j, p)^{r_j/m} \le \sum_{j=1}^{k} \eta(X; r_j, p),$$

and apply this to get the result. □

We now move to Lie coordinates and $\nu = 2$.

Theorem 4.3. *The pair (X_t, A_{st}), of the path and its area process, have finite p-variation for all $p > \alpha$, $\mathbb{P}$-a.s.*

Proof. From our Assumption 3.2, we have

$$E|X_{t_i^n} - X_{t_{i+1}^n}|^p \le c_{3.1}(p)2^{-np/\alpha},$$

and using the moment estimates of Theorem 3.3, we have

$$E\left|A_{t_i^n t_{i+1}^n}\right|^{p/2} \le c_{3.2}(p/2)2^{-np/\alpha}.$$

Hence, using the Corollary 4.2, we can estimate the expectation of $\omega(0,T)$,

$$E\omega(0,T) \le c_1 \sum_{n=1}^{\infty} n^{\gamma} 2^n 2^{-np/\alpha}.$$

The right hand side is easily seen to be finite almost surely, for all $p > \alpha$.

We have simultaneously proved the existence and finite p-variation of the path (X, A). It is clear that it satisfies the algebraic relation of a multiplicative functional since it must do so at every basic time pair and they are dense on the line. We have completed the proof of the theorem. □

Let A_{st}^m be the area process associated to the path x^m.

Theorem 4.4. *The paths $\mathbf{x}^m$ converge to $\mathbf{X}$ in p-variation as $m \to \infty$ for all $p > \alpha$, $\mathbb{P}$-a.s.*

Proof. We use Lemma 4.1 with the approximating sequence $\mathbf{x}^m$,

$$\omega_{\mathbf{X}\mathbf{x}^m}(0,1) \le c_{4.1}\left(\eta(A-A^m;2,p) + \eta(X-x^m;1,p)^{1/2}(\eta(X;1,p)\right.$$
$$\left. + \eta(x^m;1,p))^{1/2} + \eta(X-x^m;1,p)\right).$$

We consider the terms separately and, noting that $x_{t_i^n}^m = X_{t_i^n}$ for $n \le m$ and also $A_{t_i^n t_{i-1}^n}^m = 0$ for $n \ge m$, we have

$$\eta(A-A^m;2,p) \le \sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n-1} \left| A_{t_i^n t_{i+1}^n} - A_{t_i^n t_{i+1}^n}^m \right|^{p/2} + \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| A_{t_i^n t_{i+1}^n} \right|^{p/2}$$

$$\eta(X-x^m;1,p) \le \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| X_{t_i^n t_{i+1}^n} - x_{t_i^n t_{i+1}^n}^m \right|^p$$

$$\le 2^{p-1} \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| X_{t_i^n t_{i+1}^n} \right|^p + 2^{p-1} \sum_{n=m}^{\infty} n^\gamma 2^{n-m} \sum_{i=0}^{2^m-1} \left| 2^{m-n} x_{t_i^m t_{i+1}^m}^m \right|^p$$

$$\le 2^{p-1} \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| X_{t_i^n t_{i+1}^n} \right|^p + c_1 m^\gamma \sum_{i=0}^{2^m-1} \left| x_{t_i^m t_{i+1}^m}^m \right|^p$$

$$= c_2 \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{p-1} \left| X_{t_i^n t_{i+1}^n} \right|^p. \tag{20}$$

We can take expectations and use Cauchy-Schwarz to control the cross term,

$$E\left(\eta(X-\tilde{X};1,p)\big(\eta(X;1,p)+\eta(\tilde{X};1,p)\big)\right)^{1/2}$$
$$\le \left(E\eta(X-\tilde{X};1,p)E\big(\eta(X;1,p)+\eta(\tilde{X};1,p)\big)\right)^{1/2}.$$

Using this, our moment assumptions and Corollary 3.6,

$$E\omega_{\mathbf{X}\mathbf{x}^m}(0,1) \le c_{4.1}\left(\sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n-1} E\left| A_{t_i^n t_{i+1}^n} - A_{t_i^n t_{i+1}^n}^m \right|^{p/2}\right)$$

$$+ c_{4.1}\left(\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} E\left| A_{t_i^n t_{i+1}^n} \right|^{p/2}\right) + c_{4.1}\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{p-1} E\left| X_{t_i^n t_{i+1}^n} \right|^p$$

$$+ c_{4.1}\left(\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{p-1} E\left| X_{t_i^n t_{i+1}^n} \right|^p\right)^{1/2} (E\eta(X;1,p)+E\eta(x^m;1,p))^{1/2}$$

$$\le c_3 \sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n-1} 2^{-np/\alpha} 2^{(m-n)(1-4/\alpha)p/4} + c_4 \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{-np/\alpha}$$

$$+ c_5 \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{p-1} 2^{-np/\alpha}$$

$$+ c_6 \left(\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{p-1} 2^{-np/\alpha}\right)^{1/2} (E\eta(X;1,p)+E\eta(x^m;1,p))^{1/2}.$$

Observing that $E\eta(x^m; 1, p) \le E\eta(X; 1, p) \le \sum_n n^\gamma 2^{n(1-p/\alpha)} = c_7$, we have

$$E\omega_{\mathbf{Xx^m}}(0, 1) \le c_8 m^{\gamma+1} 2^{m(1-p/\alpha)} + c_9 m^{(\gamma+1)/2} 2^{m(1-p/\alpha)/2},$$

where the constants depend upon p, α and require $\alpha < p \le 4$. If $\alpha < 4 < p$ we have the estimate

$$E\omega_{\mathbf{Xx^m}}(0, 1) \le c_{10} m^{\gamma+1} 2^{mp(1-4/\alpha)/4} + c_{11} m^{(\gamma+1)/2} 2^{m(1-p/\alpha)/2}.$$

Thus we have exponential control on the moments and a routine Borel-Cantelli argument shows that, in the case where $p \le 4$,

$$\limsup_{n\to\infty} \frac{\omega_{\mathbf{Xx^n}}(0, 1)}{n^{(3+\gamma)/2}(\log n)^{1+\epsilon} 2^{(1-p/\alpha)n/2}} = 0.$$

There is a similar result for the case when $p > 4$. Thus we have established the convergence in p-variation for all $p > \alpha$.   □

Finally we can state a version of the Wong-Zakai theorem for our reversible Markov processes.

Corollary 4.5. *Let X be a rough path satisfying Assumptions 3.1, 3.2 with $\alpha < 3$ and let x^m denote the sequence of dyadic piecewise linear approximations to X. If f is a Lip(γ) vector field with $\gamma > \alpha$ and y^m is the sequence of paths defined by solving*

$$dy^m = f(y^m)dx^m, \quad y_0^m = a,$$

then the sequence (x^m, y^m) (of paths and their iterated integrals) converges in p-variation to the pair of rough paths (X, Y) for every $p > \alpha$.

5. Higher order iterated integrals

Consider now the case where the sample paths of our process have a Hölder continuity of order at most $1/3$. In order to use the results of [17] to solve differential equations along the sample paths of these Markov processes it is sufficient to find a canonical construction, valid for every path, for a third iterated integral satisfying the appropriate algebraic and analytic properties.

We will discuss two natural approaches to this problem. The most direct method is to exploit the existence of the area process we have already constructed. The argument is analytically robust and would work for other choices of the area in addition to the canonical one we have introduced. We will prove that the solutions to the differential equations obtained by taking helical approximations on our dyadic partitions, chosen so that the curves over the time intervals $\left[k2^{-n}, (k+1)2^{-n}\right]$ have the same increment and area as the restriction of the underlying stochastic path to the same interval, converge. Then we will consider the case where we construct the third order term just using the polygonal paths, without knowledge of the exact area process as constructed in the previous section.

In order to give the extension of the Wong-Zakai theorem to the third order case we need convergence of the approximations. We will be able to show that if we take either approach and compute the third iterated integrals of these approximations, then these will converge in the p-variation sense required for continuity of the solutions to the differential equations.

5.1. Constructing the third order term, approach one

Our first approach is quite close to the original argument of [17] used in proving Theorem 2.3. The key point is that we do not try to show that the polygonal approximations converge with their iterated integrals of degree up to three. Rather we use the existence of the Lévy area and the reversibility to construct the third iterated integral directly.

In order to compute the third order iterated integral we follow the same approach as for the area; decompose the time interval at dyadic times and consider a limit of products in $T^{(3)}$, the enveloping algebra. For a dissection D we use our knowledge of the path and its area to define

$$\mathbf{x}_{0t}^{D,(3)} = e^{x_{u_0 u_1} + A_{u_0 u_1}} e^{x_{u_1 u_2} + A_{u_1 u_2}} \dots e^{x_{u_{r-1} u_r} + A_{u_{r-1} u_r}}, \tag{21}$$

where we take the obvious inclusion (not Lie!) of $\mathcal{A}^{(2)} \subset \mathcal{A}^{(3)}$ to place the product in $T^{(3)}$. Recall that we used the notation t^n for the dyadic partition. Our aim will be to show that for these t^n, the products $\mathbf{x}_{0t}^{t^n,(3)}$ converge. This is quite obvious at the level of one and two tensor components for, by the algebraic condition (3) satisfied by $\mathbf{x}_{0t}^{(2)}$, these components of $\mathbf{x}_{0t}^{D,(3)}$ do not depend on D. All the action, and our interest is in the level three component. As a result, if we have two partitions D, D', then $\mathbf{x}_{0t}^{D,(3)} - \mathbf{x}_{0t}^{D',(3)} = \log \mathbf{x}_{0t}^{D,(3)} - \log \mathbf{x}_{0t}^{D',(3)}$ and indeed lies in the centre of the algebra. This means we can look at terms coming from individual sections of the path and "add them up". We also remark that as we are considering products of elements in $G^{(3)}$ it is obvious that the limit if it exists will also be in $G^{(3)}$.

Define B_{0t}^m to be the third order component of $\log \mathbf{x}_{0t}^{t^m,(3)}$. First we consider this set of times to be fixed and construct the third order term as a sum in the same way as the area (5),

$$B^m = \sum_{n=0}^{m-1} \sum_{i=0}^{2^n-1} B_{t_i^n t_{i+1}^n}\left(x_{i,i+1}^{n+1}, A_{t_i^n t_{i+1}^n}\right). \tag{22}$$

From the Campbell-Baker-Hausdorff formula and (21), the term $B_{t_i^n t_{i+1}^n}(x_{i,i+1}^{n+1}, A_{t_i^n t_{i+1}^n})$ is given by

$$B_{t_i^n t_{i+1}^n}\left(x_{i,i+1}^{n+1}, A_{t_i^n t_{i+1}^n}\right) = \frac{1}{2}[\tilde{x}_1, \tilde{A}_2] + \frac{1}{2}[\tilde{x}_2, \tilde{A}_1] + \frac{1}{12}[\tilde{x}_1, [\tilde{x}_1, \tilde{x}_2]]$$
$$+ \frac{1}{12}[\tilde{x}_2, [\tilde{x}_2, \tilde{x}_1]], \tag{23}$$

where $\tilde{x}_1 = x_i^n - x_{2i+1}^{n+1}$, $\tilde{x}_2 = x_{2i+1}^{n+1} - x_{i+1}^n$, $\tilde{A}_1 = A_{t_i^n t_{2i+1}^{n+1}}$ and $\tilde{A}_2 = A_{t_{2i+1}^{n+1} t_{i+1}^n}$.

Note that all terms are third order and we can bound them above by

$$|B_{t_i^n t_{i+1}^n}| \leq \frac{1}{2}|x_i^n - x_{2i+1}^{n+1}||\tilde{A}_2| + \frac{1}{2}|\tilde{A}_1||x_{2i+1}^{n+1} - x_{i+1}^n|$$
$$+ \frac{1}{12}|x_i^n - x_{2i+1}^{n+1}|^2|x_{2i+1}^{n+1} - x_{i+1}^n| + \frac{1}{12}|x_i^n - x_{2i+1}^{n+1}||x_{2i+1}^{n+1} - x_{i+1}^n|^2.$$
(24)

Theorem 5.1. *For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumptions 3.1, 3.2, the third order iterated integral $B_{0T}(X, A)$, up to time T, defined by the limit in (22) exists and there are constants $c_{5.1}(\beta)$ such that for all $\beta \geq 1$,*

$$E|B_{0T}^\beta(X, A)| \leq c_{5.1}(\beta)T^{3\beta/\alpha}.$$

Proof. In order to demonstrate the existence of the third order iterated integral for a fixed pair of times we follow the same reasoning as for the area and show that the limit of B^m as $m \to \infty$ has finite mean. Let $\psi(\pi(X_{t_i^n}, X_{t_{i+1}^n})) = B_{t_i^n t_{i+1}^n}(x_{i,i+1}^{n+1}, A_{t_i^n t_{i+1}^n}(x^{n+1}))$. By the definition of $B_{t_i^n t_{i+1}^n}$ in (23) we see that it changes sign as the path is reversed. This, combined with the reversibility of the path, allows us to apply Lemma 3.4 to get the forward-backward martingale decomposition

$$\sum_{i=0}^{2^n-1} B_{t_i^n t_{i+1}^n}(x_{i,i+1}^{n+1}, A_{t_i^n t_{i+1}^n}) = \frac{1}{2}M_{2^n}^n + \frac{1}{2}\tilde{M}_0^n - E(B_{t_{2^n-1}^n t_{2^n}^n}|\tilde{\mathcal{G}}_0) + E(B_{t_1^n t_0^n}|\mathcal{G}_0),$$
(25)

where the forward martingale is defined to be

$$M_k^n = \sum_{i=0}^{k}(B_{t_i^n t_{i+1}^n} - E(B_{t_i^n t_{i+1}^n}|\mathcal{G}_{t_i^n})),$$
(26)

and the backward martingale is defined similarly, where the filtrations are those given in (6). We can then establish

$$E(|B_{t_i^n t_{i+1}^n}(x_{t_i^n t_{i+1}^n}^{n+1}, A_{t_i^n t_{i+1}^n})| \mid \mathcal{G}_{t_i^n}) \leq c_1 T^{3/\alpha}2^{-3n/\alpha},$$
(27)

following the proof of Lemma 3.5, using (24), Cauchy-Schwarz, the moment assumption on the increments and the moment estimate for the area process.

The next step is to estimate the brackets of the martingales and this is done in the same way as in the derivation of (12). We get pathwise control on the first term in (26) and use (27) to estimate the second. Taking expectations gives

$$E([M^n]_{2^n}) \leq c_2 T^{6/\alpha}2^{n(1-6/\alpha)}.$$

Thus we can control the forward and backward martingales provided $\alpha > 6$. Now we can let $m \to \infty$ in (22) and see that the mean of the sum for the third order term will converge, to establish the existence, provided $\alpha < 6$. The moment estimate will follow in exactly the same way. $\square$

In Sect. 3 we showed that, with probability one, the dyadic polygonal approximation to our path, when enhanced with its natural area process as in Definition 2.1, converges in p-variation. In analogy with our definition of canonical area for a piecewise smooth path we can define an n-th order canonical iterated integral for a piecewise smooth path. It is natural to ask if there is a similar theorem here. In other words we would like to identify a sequence of piecewise smooth paths x^m, with the same increment and area over the intervals $[t_i^n, t_{i+1}^n]$, as our path (X, A) constructed above, which converges in p-variation to (X, A, B).

The first point is to identify a sensible replacement for linear interpolation as straight line segments have no area. The appropriate quantity is a helix, Γ_t, defined by

$$
\Gamma_t = \begin{bmatrix}
\cos a_1 t & \sin a_1 t & 0 & 0 & 0 & 0 \\
-\sin a_1 t & \cos a_1 t & 0 & 0 & 0 & 0 \\
0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & 0 & 0 & 0 & \cos a_d t & \sin a_d t \\
0 & 0 & 0 & 0 & -\sin a_d t & \cos a_d t
\end{bmatrix}
\begin{bmatrix}
r_1 \\ 0 \\ \vdots \\ \vdots \\ r_d \\ 0
\end{bmatrix},
$$

if the dimension is even, or by

$$
\Gamma_t = \begin{bmatrix}
\cos a_1 t & \sin a_1 t & 0 & 0 & 0 & 0 & 0 \\
-\sin a_1 t & \cos a_1 t & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \ddots & \ddots & 0 & 0 & 0 \\
0 & 0 & \ddots & \ddots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cos a_d t & \sin a_d t & 0 \\
0 & 0 & 0 & 0 & -\sin a_d t & \cos a_d t & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & a_0 t
\end{bmatrix}
\begin{bmatrix}
r_1 \\ 0 \\ \vdots \\ \vdots \\ r_d \\ 0 \\ 1
\end{bmatrix},
$$

if the dimension is odd. In general, any curve is a helix if it can be isometrically embedded as a segment of Γ_t for suitable choices of a_i, r_i. The a_i are the angular velocities and the r_i are the radii associated with the helix.

It is an elementary piece of linear algebra that a helix with given increment and area always exists, indeed there are infinitely many but at most d have minimal length and generically the one of minimal length is unique. We will use these helical segments to interpolate and produce a piecewise smooth path with given increment and area over the intervals $[t_i^n, t_{i+1}^n]$.

In order to prove that the approximating paths converge to the limit path in p-variation we need to control the p-variation of rough paths up to the third order term. For the third order case the p-variation of our path $\mathbf{X}$ in Lie coordinates is

$$
\omega(s, t) = \sup_{\mathcal{D}} \sum_i \left(|x_{u_i u_{i+1}}|^p + |A_{u_i u_{i+1}}|^{p/2} + |B_{u_i u_{i+1}}|^{p/3} \right).
$$

From Corollary 4.2, we can control it using the values at the dyadics of the iterated integrals up to order 3,

$$\omega'(s, t) \leq c_{5.2} \sum_{i=1}^{3} \eta(X; i, p) \tag{28}$$

By taking expectations in (28), substituting our moment estimates and applying Borel-Cantelli, we have the following.

Lemma 5.2. *For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumptions 3.1, 3.2, the path $\mathbf{X} = (1, X, X^2, X^3)$ has finite p-variation, for all $p > \alpha$, $\mathbb{P}$-a.s.*

We can also establish the convergence of the polygonal approximations with their areas.

Lemma 5.3. *For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumptions 3.1, 3.2, the polygonal path $\mathbf{x}^m = (1, x^m, (x^m)^2, (x^m)^3)$, taken with the correct area, converges in p-variation to the path $\mathbf{X} = (1, X, X^2, X^3)$ for all $p > \alpha$, $\mathbb{P}$-a.s.*

Proof. We begin by considering the third order version of Lemma 4.1,

$$\omega_{\mathbf{X}\mathbf{x}^m}(0, 1)$$

$$\leq c_{4.1} \sum_{i=1}^{3} \sum_{k=1}^{i} \sum_{\substack{(r_1,\dots,r_k) \\ \sum_1^k r_j = i}} \sum_{l=1}^{k} \eta(X - x^m; r_l, p)^{r_l/i} \prod_{\substack{j=1 \\ j \neq l}}^{k} \left(\eta(X; r_j, p) + \eta(x^m; r_j, p) \right)^{r_j/i}.$$

Now writing out the control on the third term and using the fact that

$$\eta(x^m; r, p) \leq \eta(X; r, p),$$

we have

$$\begin{aligned}
\omega_{\mathbf{X}\mathbf{x}^m}(0, 1) \leq c_{4.1} \Big(&\eta(X - x^m; 1, p) + 2(2\eta(X; 1, p)\eta(X - x^m; 1, p))^{1/2} \\
&+ \eta(X - x^m; 2, p) + 3\eta(X - x^m; 1, p)^{1/3}(2\eta(X; 1, p))^{2/3} \\
&+ 2\eta(X - x^m; 1, p)^{1/3}(2\eta(X; 2, p))^{2/3} \\
&+ 2\eta(X - x^m; 2, p)^{2/3}(2\eta(X; 1, p))^{1/3} + \eta(X - x^m; 3, p) \Big)
\end{aligned} \tag{29}$$

From the relationship between the two coordinate charts for the group we can write

$$\eta(X; 3, p) \leq c_1 \left(\eta(X; 1, p) + \eta(A; 2, p) + \eta(B; 3, p) \right).$$

If we combine this with (29), take expectations and apply Hölders inequality we have

$$
\begin{aligned}
E\omega_{\mathbf{X}\mathbf{x}^{\mathbf{m}}}(0,1) \leq{} & c_2 \left(E\eta(X - x^m; 1, p) + E\eta(A - A^m; 2, p) + E\eta(B - B^m; 3, p) \right) \\
& + c_3 (E(\eta(X; 1, p)) E(\eta(X - x^m; 1, p)))^{1/2} + c_4 \left((E(\eta(X; 1, p)))^{2/3} \right. \\
& \left. + (E(\eta(A; 2, p)))^{2/3} \right) (E(\eta(X - x^m; 1, p)))^{1/3} + \\
& + c_5 (E(\eta(X; 1, p)))^{1/3} \left(E(\eta(A - A^m; 2, p)) + E(\eta(X - x^m; 1, p)) \right)^{2/3}
\end{aligned}
$$

$$(30)$$

Using the fact that there are constants such that $E\eta(X; 1, p) \leq c_6$ and $E\eta(X; 2, p) \leq c_7$, we see that it is enough to show the convergence of the terms $\eta(X - x^m; 1, p)$, $\eta(A - A^m; 2, p)$ and $\eta(B - B^m; 3, p)$. Firstly, observe that as the path and its dyadic approximation are equal at the dyadic time points, we can write

$$
\begin{aligned}
E\eta(X - x^m; 1, p) &= \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} E \left| X_{t_i^n t_{i+1}^n} - x_{t_i^n t_{i+1}^n}^m \right|^p \\
&\leq \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} 2^{p-1} E \left| X_{t_i^n t_{i+1}^n} \right|^p + \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} 2^{p-1} E \left| X_{t_i^n t_{i+1}^n}^m \right|^p,
\end{aligned}
$$

and similarly

$$
\begin{aligned}
E\eta(A - A^m; 2, p) &= \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} E \left| A_{t_i^n t_{i+1}^n} - A_{t_i^n t_{i+1}^n}^m \right|^{p/2} \\
&\leq \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} 2^{p-1} E \left| A_{t_i^n t_{i+1}^n} \right|^{p/2} + \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} 2^{p-1} E \left| A_{t_i^n t_{i+1}^n}^m \right|^{p/2}.
\end{aligned}
$$

The third order term is different as we have to take account of the initial terms

$$
\begin{aligned}
E\eta(B - B^m; 3, p) &= \sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n - 1} E \left| B_{t_i^n t_{i+1}^n} - B_{t_i^n t_{i+1}^n}^m \right|^{p/3} \\
&\quad + \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n - 1} E \left| B_{t_i^n, t_{i+1}^n} - B_{t_i^n t_{i+1}^n}^m \right|^{p/3} \\
&\leq \sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n - 1} E \left| B_{t_i^n t_{i+1}^n} - B_{t_i^n t_{i+1}^n}^m \right|^{p/3} \\
&\quad + \sum_{n=m}^{\infty} n^\gamma 2^{p-1} \sum_{i=0}^{2^n - 1} E \left| B_{t_i^n t_{i+1}^n} \right|^{p/3} \\
&\quad + \sum_{n=m}^{\infty} n^\gamma 2^{p-1} \sum_{i=0}^{2^n - 1} E \left| B_{t_i^n t_{i+1}^n}^m \right|^{p/3}.
\end{aligned}
$$

The terms involving A^m, X^m can be controlled using the relationship between the length of the increment and the area of a helix. Firstly for the increment we use an isoperimetric inequality, that $|X^m_{t^n_i t^n_{i+1}}| \le c_8 |A^m_{t^m_i t^m_{i+1}}|^{1/2}$,

$$\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| X^m_{t^n_i t^n_{i+1}} \right|^p \le \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^m-1} 2^{n-m} \left(2^{m-n} c_8 \left| A^m_{t^m_i t^m_{i+1}} \right|^{1/2} \right)^p$$

$$\le c_9 m^{\gamma+1} \sum_{i=0}^{2^m-1} \left| A^m_{t^m_i t^m_{i+1}} \right|^{p/2}.$$

Similarly we have

$$\sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} \left| A^m_{t^n_i t^n_{i+1}} \right|^{p/2} \le c_{10} m^{\gamma+1} \sum_{i=0}^{2^m-1} \left| A^m_{t^m_i t^m_{i+1}} \right|^{p/2}.$$

Thus we see that, by our moment assumptions

$$E\eta(X - x^m; 1, p) \le c_{11} m^{\gamma+1} 2^m E \left| A^m_{t^m_i t^m_{i+1}} \right|^{p/2} = c_{12} m^{\gamma+1} 2^{m(1-p/\alpha)},$$

and similarly $E\eta(A - A^m; 2, p) \le c_{13} m^{\gamma+1} 2^{m(1-p/\alpha)}$.

The third order term can be controlled using

$$\left| B^m_{I_i} \right| \le \left| X^m_{I_i} \right| \left| A^m_{I_i} \right| + \frac{1}{6} \left| X^m_{I_i} \right|^3 \le c_{14} \left| A^m_{I_i} \right|^{3/2} + c_{15} \left| X^m_{I_i} \right|^3.$$

We observe that there is a rate of convergence result for the third order term arising from the proof of Theorem 5.1

$$E \left| B_{0T} - B^m_{0T} \right|^\beta \le c_{16} T^{3\beta/\alpha} 2^{m\beta(1-6/\alpha)/2}.$$

Now, using the above estimates and the rate of convergence result in (30), we get

$$E\omega_{\mathbf{Xx^m}}(0,1) \le c_{16} \sum_{n=1}^{m-1} n^\gamma \sum_{i=0}^{2^n-1} 2^{-np/\alpha} 2^{(m-n)(1-6/\alpha)p/6} + c_{17} \sum_{n=m}^{\infty} n^\gamma \sum_{i=0}^{2^n-1} 2^{-np/\alpha}$$

$$+ c_{18} m^{2(\gamma+1)/3} 2^{2m(1-p/\alpha)/3} + c_{19} m^{(\gamma+1)/2} 2^{m(1-p/\alpha)/2}$$

$$+ c_{20} m^{(\gamma+1)/3} 2^{m(1-p/\alpha)/3}$$

$$= c_{21} \left(m^{\gamma+1} 2^{m(1-p/\alpha)} + m^{(\gamma+1)/3} 2^{m(1-p/\alpha)/3} \right),$$

where the constant depends upon p, α and T and requires $\alpha < p \le 6$. If $\alpha < 6 < p$ we have the estimate

$$E\omega_{\mathbf{Xx^m}}(0,1) \le c_{22} \left(m^{\gamma+1} 2^{mp(1-6/\alpha)/6} + m^{(\gamma+1)/3} 2^{m(1-p/\alpha)/3} \right).$$

Thus we have convergence in p-variation norm. $\square$

5.2. Constructing the third order term, approach two

Thus far we have constructed the third order term, given that we already have the Lévy area process itself. In this section we will consider the question of constructing the third order term given that we only have the polygonal approximations to the path. In this situation we will not be able to add up the separate pieces, there will be a contribution to the third order iterated integral from pieces of the path outside the particular increment. However, we will see that this new piece will still undergo a change of sign as the path is reversed and hence we will be able to apply the forward-backward martingale decomposition to prove the existence of this part of the third iterated integral. As the techniques used here are essentially the same as in the previous section we will omit detailed proofs.

We consider building our polygonal approximations to the path by considering a particular dyadic interval $[t_i^n, t_{i+1}^n]$ in the interval $[0, T]$. We determine the effect on the rough path at level 3 of the polygonal approximation over this interval by comparing the second and third order terms for the path in that interval with the full path in $T^{(3)}$,

$$
\log(e^{x_{0t_i^n} + A_{0t_i^n} + B_{0t_i^n}} e^{x_{t_i^n t_{i+1}^n} + A_{t_i^n t_{i+1}^n} + B_{t_i^n t_{i+1}^n}} e^{x_{t_{i+1}^n T} + A_{t_{i+1}^n T} + B_{t_{i+1}^n T}})
$$
$$
= \log(e^{x_{0t_i^n} + A_{0t_i^n} + B_{0t_i^n}} e^{x_{t_i^n t_{i+1}^n}} e^{x_{t_{i+1}^n T} + A_{t_{i+1}^n T} + B_{t_{i+1}^n T}}) + A_{t_i^n t_{i+1}^n}
$$
$$
+ (B_{t_i^n t_{i+1}^n} + [x_{0t_i^n}, A_{t_i^n t_{i+1}^n}] + [A_{t_i^n t_{i+1}^n}, x_{t_{i+1}^n T}]).
$$

Thus the second order term required is the area, which we saw how to compute in the previous sections. The third order term is composed of two parts. There is our third order term $B_{t_i^n t_{i+1}^n}$, defined on the increment, and a new third order term, generated by terms outside of the increment, given by

$$
\hat{B}_{t_i^n t_{i+1}^n} = [x_{0t_i^n}, A_{t_i^n t_{i+1}^n}] + [A_{t_i^n t_{i+1}^n}, x_{t_{i+1}^n T}]
$$
$$
= [x_{0t_i^n} - x_{t_{i+1}^n T}, A_{t_i^n t_{i+1}^n}]
$$
$$
= [x_i^n + x_{i+1}^n, A_{t_i^n t_{i+1}^n}] - [x_0^n + x_T^n, A_{t_i^n t_{i+1}^n}].
$$

In approach one we have seen how to construct the third order term $B_{t_i^n t_{i+1}^n}$. Observing that the area of a line segment is 0 we have, from (23), that for purely polygonal paths,

$$
B_{t_i^n t_{i+1}^n}(x^{n+1}) = \frac{1}{12} \left([x_1, [x_1, x_2]] + [x_2, [x_2, x_1]]\right),
$$

where $x_1 = x_{t_i^n t_{2i+1}^{n+1}}^{n+1}$, $x_2 = x_{t_{2i+1}^{n+1} t_{i+1}^n}^{n+1}$.

We now need to ensure that the other component of the third order term, constructed from pieces of path outside the increment, is also well behaved. The piece we need is $\hat{B}_{0T}(X)$ which we obtain as the limit as $m \to \infty$ of

$$
\hat{B}^m = \sum_{n=1}^{m} \sum_{i=0}^{2^n - 1} \hat{B}_{t_i^n t_{i+1}^n}(x^{n+1}),
$$

where

$$\hat{B}_{t_i^n t_{i+1}^n}(x^{n+1}) = \left[x_i^n + x_{i+1}^n, A_{t_i^n t_{i+1}^n}(x^{n+1})\right] - \left[x_0 + x_T, A_{t_i^n t_{i+1}^n}(x^{n+1})\right].$$

If we run the path backward, then both increment and area are reversed in sign and hence the term $\hat{B}_{t_i^n t_{i+1}^n}(x^{n+1}) = -\hat{B}_{t_i^n t_{i+1}^n}(x^{n+1})$. Thus this function of the path segment satisfies the conditions needed to apply our forward-backward martingale decomposition, Lemma 3.4. From the definition of the new piece of the third iterated integral we have that there exists a constant $c_{5.3}$ such that

$$\left|\hat{B}_{t_i^n t_{i+1}}(x^{n+1})\right| \leq c_{5.3}\left|x_0 + x_T - x_i^n - x_{i+1}^n\right|\left|x_i^n - x_{2i+1}^{n+1}\right|\left|x_{2i+1}^{n+1} - x_{i+1}^n\right|.$$

We can control both the global and local parts of this expression with our assumption on the Hölder continuity and hence we can construct a canonical choice of third order iterated integral from the polygonal approximations to the path.

Theorem 5.4. *For a reversible Markov process $X = \{X_t; t \in [0, T]\}$, satisfying Assumptions 3.1, 3.2, the third order iterated integral, constructed purely from the polygonal path approximations, $B_{0T}(X) + \hat{B}_{0T}(X)$, up to time T, exists and there are constants $c_{5.4}(\beta)$ such that for all $\beta \geq 1$,*

$$E(B_{0T}(X) + \hat{B}_{0T}(X))^\beta \leq c_{5.4}(\beta)T^{3\beta/\alpha}.$$

Using the same approach as above we can combine the above results to establish our final result.

Theorem 5.5. *The path $\mathbf{X}$ has finite p-variation, for all $p > \alpha$, $\mathbb{P}$-a.s. and the polygonal path $\mathbf{x^m} = (1, x^m, (x^m)^2, (x^m)^3)$ converges in p-variation to the path $\mathbf{X} = (1, X, X^2, X^3)$, for all $p > \alpha$, $\mathbb{P}$-a.s.*

Note that the rate of convergence in the two approaches differs. As can be seen from the calculations, for the second approach we have for $\alpha < p \leq 4$,

$$E\omega_{\mathbf{X}\mathbf{x^m}}(0, 1)$$
$$\leq c_{5.4}\left(m^{\gamma+1}2^{mp(1-4/\alpha)/4} + m^{2(\gamma+1)/3}2^{2m(1-p/\alpha)/3} + m^{(\gamma+1)/3}2^{m(1-p/\alpha)/3}\right).$$

6. Examples and open problems

We illustrate our results by considering some examples. When we can construct the stochastic area process, or, if necessary, the third order iterated integral process, we can make sense of differential equations driven by the path of our reversible Markov process. This allows us to solve differential equations driven by processes which are not necessarily semi-martingales. In particular we can consider paths which arise from divergence form operators or diffusion processes on such fractals as the Sierpinski carpet.

1. Elliptic operators

We assume that the diffusion is generated by a second order elliptic operator in divergence form. Let $U \subset \mathbb{R}^d$ be open, and take a symmetric matrix a_{ij} which is positive definite and has $a_{ij}, (\partial/\partial x_i)a_{ij} \in L^2_{loc}(U, dx)$ for $1 \leq i, j \leq d$ (where we interpret the derivative in the sense of Schwarz distributions). Define the operator $\mathcal{L}$ on $L^2(U, dx)$ by

$$\mathcal{L}u = \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial}{\partial x_j} \right) u.$$

There is a Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(U, dx)$, defined by $\mathcal{E}(u, v) = (-\mathcal{L}u, v)$ for $u, v \in \mathcal{D}(\mathcal{E})$, where $\mathcal{D}$ is formed by closing $C_0^\infty(U)$ on $L^2(U, dx)$. The Dirichlet form is local and regular and by the theory of Dirichlet forms, there is an associated continuous Hunt process $X = \{X_t; t \geq 0\}$ on U. There is also a continuous heat kernel which enables us to define the Hunt process started from every point. This process is reversible and Markov and hence we can apply our results if there is a uniform Hölder condition on the paths. By standard results, such as Aronson's heat kernel bounds, we have that there exist constants $c_{6.1}(k)$, such that

$$\sup_{x \in U} E^x |X_t - X_0|^k \leq c_{6.1}(k)t^{k/2}, \quad 0 < t < 1,$$

and hence there is an area process associated with the path.

2. Brownian motion on nested fractals

The first fractal to have a Brownian motion defined was the Sierpinski gasket. The construction of the process was facilitated by the finite ramification of the fractal, that any connected subset of the fractal can be disconnected from the rest of the fractal by the removal of only a finite number of points. Nested fractals were proposed in [16] in order to define a class of fractals with the property of finite ramification, symmetry and exact self-similarity. Lindstrøm was able to prove the existence of a Brownian motion on these sets. They are defined as fixed points of a family of similitudes. For $\alpha > 1$, an α-*similitude* is a map $\psi : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\psi(x) = \alpha^{-1}U(x) + x_0, \tag{31}$$

where U is a unitary, linear map and $x_0 \in \mathbb{R}^d$. Let $\Psi = \{\psi_1, \ldots, \psi_m\}$ be a finite family of α-similitudes. For $B \subset \mathbb{R}^d$, define $\Phi(B) = \cup_{i=1}^m \psi_i(B)$, and let $\Phi_n(B) = \Phi \circ \cdots \circ \Phi(B)$. The map Φ, on the set of compact subsets of $\mathbb{R}^d$, has a unique non-void fixed point F, which is a self-similar set satisfying $F = \Phi(F)$.

Let F' be the set of fixed points of the mappings ψ_i, $1 \leq i \leq m$. A point $x \in F'$ is called an essential fixed point if there exist $i, j \in \{1, \ldots, m\}$, $i \neq j$ and $y \in F'$ such that $\psi_i(x) = \psi_j(y)$. We write F_0 for the set of essential fixed points. Now define

$$\psi_{i_1,\ldots,i_n}(B) = \psi_{i_1} \circ \cdots \circ \psi_{i_n}(B), \quad B \subset \mathbb{R}^d.$$

We will call the set $\psi_{i_1,\dots,i_n}(F_0)$ an n-cell. The lattice of fixed points F_n is defined by $F_n = \Phi_n(F_0)$, and the set F can be recovered from the essential fixed points by setting $F = cl(\cup_{n=0}^{\infty} F_n)$.

We can now define a nested fractal as follows.

Definition 6.1. *The set F is a nested fractal if $\{\psi_1, \dots, \psi_m\}$ satisfy:*
(A1) (Connectivity) For any 1-cells C and C', there is a sequence $\{C_i : i = 0, \dots, n\}$ of 1-cells such that $C_0 = C$, $C_n = C'$ and $C_{i-1} \cap C_i \neq \emptyset$, $i = 1, \dots, n$.
(A2) (Symmetry) If $x, y \in F_0$ then reflection in the hyperplane $H_{xy} = \{z : |z-x| = |z-y|\}$ maps F_n to itself.
(A3) (Nesting) If $\{i_1, \dots, i_n\}, \{j_1, \dots, j_n\}$ are distinct sequences then

$$\psi_{i_1,\dots,i_n}(F) \bigcap \psi_{j_1,\dots,j_n}(F) = \psi_{i_1,\dots,i_n}(F_0) \bigcap \psi_{j_1,\dots,j_n}(F_0).$$

(A4) (Open set condition) There is a non-empty, bounded, open set V such that the $\psi_i(V)$ are disjoint and $\cup_{i=1}^{m} \psi_i(V) \subset V$.

The existence of a continuous strong Markov process, $X = \{X_t; t \geq 0\}$, reversible with respect to the Hausdorff measure on the fractal, was proved via Brouwer's fixed point theorem in [16]. Recently the uniqueness of this process has been demonstrated in [21]. The transition density estimates for these fractals were established in [14].

From [14] Theorem 6.1, there exist constants $c_{6.2}(k)$, d_w such that

$$E^x|X_t - x|^k \leq c_{6.2}(k)t^{k/d_w}, \quad 0 < t < 1. \tag{32}$$

The exponent for the Hölder continuity is given by d_w, an exponent called the walk dimension for the fractal satisfying $d_w > 2$. Hence we can construct the area process for diffusion processes on nested fractals for which $d_w < 4$ and, in the case where $3 \leq d_w < 4$, we can construct the third order iterated integral and thus give a meaning to differential equations driven by these processes provided that $d_w < 4$. It is also straightforward to extend this to affine nested fractals as defined in [7].

Note that these results cover those obtained in [8] for the Sierpinski gasket. If we consider the d-dimensional Sierpinski gasket, formed from $d + 1$, d-dimensional tetrahedra, then the walk dimension can be computed and is given by $d_w = \log(d + 3)/\log 2$. Thus for $2 \leq d < 5$ we have $2 < d_w < 3$, and we can construct the area and solve differential equations. If $5 \leq d < 13$ we have $3 \leq d_w < 4$ and we need to construct the third order iterated integral. In the case where $d \geq 13$ our technique does not allow us to construct a canonical choice of area.

3. Sierpinski carpets

We consider a further class of fractals which do not have the property of finite ramification. These are Sierpinski carpets as considered in [1], [2], [3]. We take the unit cube F_0 in dimension $d > 1$ and take a suitable family of similitudes $\psi = \{\psi_i; 1 \leq i \leq N\}$. We define the map $\Psi = \cup_{i=1}^{N} \psi_i$. Let $F_n = \Psi_n(F_0)$ denote

the $n - th$ approximation to the Sierpinski carpet, then the carpet E, the fixed point of the set of similitudes, is given by $E = \cap F_n = \Psi(E)$.

We call E a generalized Sierpinski carpet if it satisfies the following.

Definition 6.2. (A1) (Symmetry) F_1 is preserved by all the isometries of the unit cube F_0.
(A2) (Connectedness) E is connected and contains a path connecting the hyperplanes $\{x_1 = 0\}$ and $\{x_1 = 1\}$.
(A3) (Non-diagonality) Let B be a cube in F_0 with length $2/l$ and with vertices on $l^{-1}\mathbb{Z}$. Then if $Int(F_1 \cap B)$ is non-empty, it is connected.

The recent results of [3] show how to construct a Brownian motion on the generalized Sierpinski carpet in any dimension. Estimates on the heat kernel for the resulting process are also obtained which give the uniform Hölder condition in terms of a walk dimension, d_w. This is the same bound as in (32), however there is no explicit expression for the walk dimension. As in the case of nested fractals it is possible for the carpet to have walk dimension satisfying $3 \le d_w < 4$, where we need to consider the third iterated integral. It is also possible to have $d_w \ge 4$, in which case we do not know how to make a canonical choice of area. Thus we can solve Stratonovich stochastic differential equations driven by the paths of diffusions on generalized Sierpinski carpets with $d_w < 4$.

4. Non-existence of the canonical area
We conclude with an example which demonstrates that it may not be possible to find a canonical choice of area for diffusions with a Hölder condition, if they are not reversible. For this we consider a class of non-reversible processes on the Sierpinski gasket, called p-stream diffusions, constructed by Kumagai [15]. The diffusion process is constructed as a limit of random walks, where the walks have a preferred direction of rotation around the Sierpinski gasket. Consider the random walk on the lattice approximation to the Sierpinski gasket and allow it to make transitions in a clockwise direction with probability p^2 on the outer edge and $p(1 - p)$ toward the centre. In the anticlockwise direction we allow it to move to the centre with probability $p(1 - p)$ otherwise $(1 - p)^2$. The time scaling factor required to rescale the random walks can be calculated and is given by $\tau_p = 2(1 + p(1 - p))/(1 - 2p(1 - p))$. It is known that the resulting diffusion is not symmetric for $p \ne 1/2$ and has paths with $d_w = \log \tau_p / \log 2$ variation. Note that for $p = 1/2$ we obtain the usual Brownian motion on the Sierpinski gasket, for $p = (1 + 1/\sqrt{5})/2$, we have $\tau_p = 4$ and $d_w = 2$ and for $p = 1$ we have $\tau_p = 2$ and, if we start from the origin, the process is uniform clockwise motion around the outer triangle.

A description of the length of a path which crosses the unit triangle can be provided by a multitype branching process. As in [8], the growth rate of the area can be calculated as the new levels of random walk path are added. A careful analysis of this process shows that our canonical area does not exist for the p-stream diffusion in the case where the walk dimension $d_w \ge 2$. For the case where $d_w < 2$ we can construct the area from the path itself in a deterministic way.

6.1. Open problems and further remarks

1. In order to prove existence of the Lévy area we require that the path has uniform p-variation with $p < 4$. Is $p = 4$ a fundamental block - or is there still a canonical "Levy Area" and rough path above the spatial path, perhaps constructed via some averaging procedure.

2. Finally, in this paper we show the existence of the Lévy area for almost every path when we start the reversible Markov process with the stationary measure making it reversible. We have been unable to improve the result to the case of the process started at a single point. This difficulty is analogous to the problem of extending the forward-backward martingale decomposition. These were overcome in [18] for the martingale decomposition.

3. The initial success in understanding (1) for rough paths was that of Itô in the case where x was a Brownian motion. Brownian motion can be viewed as a Markov process, and we have shown here that (1) can be given meaning for a wide class of Markov processes. Brownian motion can also be viewed as a vector valued Gaussian process. Do other vector valued Gaussian processes also possess an area and higher order iterated integrals?

    Since the completion of this paper [5,6] have extended this approach to construct Levy area for Gaussian processes with suitable control on their correlation. In the case of fractional Brownian motion, if the Hurst index is at least $1/2$, it easily fits the theory of rough paths, because it is either Brownian motion ($H = 1/2$) or has sample paths of finite p variation less than 2. If the Hurst index $H < \frac{1}{2}$, then [5] shows that a canonical area can be constructed for $\frac{1}{4} < H \le \frac{1}{2}$.

References

1. M.T. Barlow, R.F. Bass: The construction of Brownian motion on the Sierpinski carpet. Ann. Inst. H. Poincare **25**, 225–257 (1989)
2. M.T. Barlow, R-F. Bass: Transition densities for Brownian motion on the Sierpinski carpet. Prob. Theory Rel. Fields **91**, 307–330 (1992)
3. M.T. Barlow, R-F. Bass: Harmonic analysis and Brownian motion on the Sierpinski carpet. Canad. J. Math. **51**, 673–744 (1997)
4. J.M.C. Clark: The representation of non-linear stochastic systems with applications to filtering. PhD thesis, Imperial College, University of London, April 1966
5. L. Coutin, Z.M. Qian: Stochastic differential equations driven by fractional Brownian motions. C. R. Acad. Sci. Paris Ser. I Math. **331**, 75–80 (2000)
6. L. Coutin, Z.M. Qian: Stochastic analysis, rough path analysis and fractional Brownian motions. Prob. Theory Rel. Fields **122**, 108–140 (2002)
7. P.J. Fitzsimmons, B.M. Hambly, T. Kumagai: Transition density estimates for diffusion on affine nested fractals. Comm. Math. Phys. **165**, 595–620 (1994)
8. B.M. Hambly, T.J. Lyons: Stochastic area for Brownian motion on the Sierpinski gasket. Ann. Probab., **26**, 132–148 (1998)
9. K. Ikeda, S. Watanabe: Stochastic differential equations and diffusion processes North-Holland, Amsterdam 1989
10. K. Itô: Differential equations determining Markov processes (in Japanese). J. Pan-Japan Math. Coll. **1077**, 1352–1400 (1942)
11. K. Itô: Stochastic integral Proc. Imperial Acad. Tokyo **20**, 519–524 (1944)

12. K. Itô: Selected papers, S.R.S. Varadhan, D.W. Stroock (eds.): New York-Berlin: Springer 1987
13. S. Kobayashi, K. Nomizu: Foundations of differential geometry, Vol. I. New York: Wiley 1963
14. T. Kumagai: Estimates of the transition densities for Brownian motion on nested fractals. Probab. Theory Relat. Fields **96**, 205–224 (1993)
15. T. Kumagai: Construction and some properties of a class of non-symmetric diffusion processes on the Sierpiski gasket. Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990), Longman, Pitman Res. Notes Math. Ser. **283**, 219–247 (1993)
16. T. Lindstrøm: Brownian motion on nested fractals. Memoirs Am. Math. Soc., No. 420, **83** (1990)
17. T.J. Lyons: Differential equations driven by rough signals. Rev. Math. Iberioamericana **14**, 215–310 (1998)
18. T.J. Lyons, L. Stoica: The limits of stochastic integals of differential forms. Ann. Probab. **27**, 1–49 (1999)
19. T.J. Lyons, O. Zeitouni: Conditional exponential moments for Iterated Wiener Integrals, and application to the Onsager-Machlup functional of diffusions Ann. Probab. **27**, 1738–1749 (1999)
20. T.J. Lyons, W. Zheng: A crossing estimate for the canonical process on a Dirichlet space and a tightness result, Colloque Paul Lévy, Astérisque No. 157–158, 249–271 (1988)
21. C. Sabot: Existence and uniqueness of diffusions on finitely ramified self-similar fractals. Ann. Scient. Ecole. Norm. Sup. **30**, 605–673 (1997)
22. D.W. Stroock, S.R.S. Varadhan: Multidimensional diffusion processes. Springer 1979
23. E. Wong, M. Zakai: On the relationship between ordinary and stochastic differential equations. Int. J. Eng. Science **3**, 213–229 (1965)
24. E. Wong, M. Zakai: On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Stat. **36**, 1560–1564 (1965)