Abstract

We prove a representation for the support of McKean-Vlasov Equations. To do so, we construct functional quantizations for the law of Brownian motion as a measure over the (non-reflexive) Banach space of Hölder continuous paths. By solving optimal Karhunen Loève expansions and exploiting the compact embedding of Gaussian measures, we obtain a sequence of deterministic finite supported measures that converge to the law of a Brownian motion with explicit rate. We show the approximation sequence is near optimal with very favourable integrability properties and prove these approximations remain true when the paths are enhanced to rough paths. These results are of independent interest.

The functional quantization results then yield a novel way to build deterministic, finite supported measures that approximate the law of the McKean-Vlasov Equation driven by the Brownian motion which crucially avoid the use of random empirical distributions. These are then used to solve an approximate skeleton process that characterises the support of the McKean-Vlasov Equation.

We give explicit rates of convergence for the deterministic finite supported measures in rough-path Hölder metrics and determine the size of the particle system required to accurately estimate the law of McKean-Vlasov equations with respect to the Hölder norm.

Keywords: McKean-Vlasov equations, Support Theorem, Gaussian Functional Quantization, Rough Paths

2010 AMS subject classifications:
Primary: 60Hxx, Secondary: 28C20

*The work of Thomas Cass is supported by EPSRC Programme Grant EP/S026347/1
†G. dos Reis acknowledges support from the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2019 (Centro de Matemática e Aplicações CMA/FCT/UNL).
1 Introduction

McKean Vlasov equations are Stochastic Differential Equations (SDE) with coefficients that depend on the law of the solution. This makes their analysis more involved than classical SDEs. They are sometimes referred to as mean-field or distribution dependent SDEs and were first studied in [McK66]. These equations describe a limiting behaviour of individual particles having diffusive dynamics and which interact with each other in a “mean-field” sense. Hence, the motion of a single particle is determined in terms of the motion of all other particles. The solutions of these mean field systems of equations are a powerful tool in understanding statistical mechanics such as Boltzmann Equations. Applications are numerous and vary from opinion dynamics [HK02], the dynamics of granular materials [BCCP98, BGG13, CGM08], molecular and fluid dynamics [Pop01], interacting agents in economics or social networks [CDL13], mathematical biology [KS71, BCM07], Galactic dynamics [BT11], droplet growth [CS19], Plasma Physics [Bit13], interacting neurons [DIRT15] and deep learning neural networks [HKR19]. See also [CD17a, CD17b] and references therein for a detailed exploration of the applications of McKean Vlasov Equations.

McKean Vlasov Equations have also been studied in the context of rough paths. In the first work [CL15], the authors treat the measure dependency in the drift term as a bounded variation Banach valued operator. Thus the interactive forces can be calculated using Banach valued Young integrals and there is no need to exploit the rough path structures beyond what is already necessary to incorporate the noise. Their approach is limited only by the assumption of no measure dependencies in the diffusion term. Later, in [BCD18] the authors develop the new framework of Probabilistic Rough Paths. This insightful development encodes the law of the noise into the rough path, allowing the noise to interact with the measure dependencies and opening up the collection of possible diffusion terms to include adequately regular measure dependencies. Other works that study McKean Vlasov Equations via rough paths include [DFMS18], [CDFM18] and [CN19].

The support of a measure is the smallest closed set of full measure. Thus the Support theorem for the law of an SDE characterises the set of admissible paths that the SDE can take with respect to a particular choice of topology. The first work studying the support of an SDE was [SV72] where the law of an SDE is characterised in terms of the supremum norm and the authors goal was to establish a Strong Maximum principle for a class of Elliptic Partial Differential Equations. This was later extended to a wide class of processes in [GP90]. Later, a support theorem with respect to the Hölder norm was established in [BAGL94], and for a much wider class of norms in [GNSS95]. These works laid the groundwork for the later publication [LQZ02] which studies the support of the solution law of a Rough Differential Equation driven by a Gaussian white noise. In [FLS06] it is shown that the continuity of the Itô-Lyons map means that the proof of a support theorem can be reduced to establishing a characterisation of the support of the driving noise in an adequately rich topology.

Support theorem results have been key in some other applications, for example, a support theorem for SDEs with jump noise was crucial in showing Exponential Ergodicity in [Kul09]. One of the conditions the authors require is Topological Irreducibility, that for any two points, there is a path of the jump process that passes between them in finite time. This can be verified by finding an expression for the support of the law. Support theorems are also central in the establishment of Stochastic Invariance principle. A stochastic process is said to be invariant of a closed set \( D \subseteq \mathbb{R}^d \) if the solution starts and remains on the set \( D \) \( \mathbb{P} \)-almost surely \( \forall t \in [0, T] \). This problem was first studied in [ADP90]. More recently, Stochastic Invariance has been studied in [Zab00, BQRT10] and [FTT14]. In general, support theorems continue to draw attention from a wide range of academics, see [CF18, CK19] and [HS19]. Lastly, a motivation to study support theorem results for McKean
Vlasov Equations is the recent link between this class of equations, deep learning (or rich learning) and ergodicity, see [HKR19].

The following useful method for proving a classical support theorem can be found in [MSS94].

**Theorem 1.1.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space containing a Brownian motion and let \(E\) be a separable Banach space. Let \(\mathcal{H}\) be the reproducing kernel Hilbert space of Brownian motion. Let \(X : \Omega \rightarrow E\) be a random variable and let \(\Phi : \mathcal{H} \rightarrow E\) be a measurable map.

1. Suppose there exists a sequence of random variables \(H_n : \Omega \rightarrow \mathcal{H}\) such that for any \(\varepsilon > 0\)
\[
\lim_{n \to \infty} \mathbb{P}\left[ \|X(\cdot) - \Phi(H_n(\cdot))\|_E > \varepsilon \right] = 0.
\] (1.1)
Then \(\text{supp}(L^X) \subset \Phi(\mathcal{H})^E\).

2. Suppose there exists a sequence of measure transforms \(T_n^h\) such that \(\mathbb{P} \circ T_n^h\) is absolutely continuous with respect to \(\mathbb{P}\) and for any \(\varepsilon > 0\)
\[
\limsup_{n \to \infty} \mathbb{P}\left[ \|X(T_n^h(\cdot)) - \Phi(h)\|_E < \varepsilon \right] > 0.
\] (1.2)
Then \(\overline{\Phi(\mathcal{H})}^E \subset \text{supp}(L^X)\).

If both (1.1) and (1.2) are satisfied, then \(\overline{\Phi(\mathcal{H})}^E = \text{supp}(L^X)\) and \(\Phi\) is called the Skeleton Process of the random variable \(X\), see [CFN97].

Equation (1.1) is sometimes referred to as the Wong Zakai implication due to its similarity with the Wong Zakai theorem. Equation (1.2) is sometimes referred to as the Cameron Martin implication because the proof involves exploiting the absolute continuity of Cameron Martin transforms on Wiener space.

**Our contribution**

Proving a support theorem for McKean Vlasov Equations is more challenging than verifying Equations (1.1) and (1.2). The knowledgeable reader will realise that for McKean Vlasov Equations, the Skeleton process is itself dependent on the law of the solution of the McKean Vlasov Equation so the law must be known exogenously in order to solve any Skeleton process path. This is in contrast to the Skeleton process used in [dRST19] where the measure dependency is replaced by a Dirac following the skeleton process driven by a constant 0 noise.

Before tackling the methods to represent the support of McKean Vlasov equations we address, separately and of independent interest, the Quantization problem for the law of a Brownian motion as a measure over the collection of Hölder continuous rough paths. The quantization problem for Gaussian measures for Hilbert spaces was first studied in [LP02], but for Banach spaces, the problem is more challenging with the optimal rate of convergence solved in [GLP03] and separately [DFMS03]. These methods rely on the small ball probabilities of Brownian motion, see [BR92], a tool to measure the compactness of the reproducing kernel Hilbert space unit ball contained in the Banach space.

Using functional analytic methods, we construct a quantization for the law of the Brownian motion that has a rate of convergence that is asymptotically equivalent to the optimal rate of convergence. Our quantization is not optimal, indeed such a quantization does not exist due to the
non-weak compactness of the Hölder unit ball. We choose to sacrifice optimality in order to retain certain key properties that allow us to estimate the law of the quantization accurately. To do this, we construct a Karhunen Loève expansion that optimally approximates the Brownian motion with respect to the Hölder norm. Although this representation for Brownian motion is well documented [HIPP14], it is not so well known that the wavelet representation comes from the spectral decomposition of the covariance kernel and so it embodies the optimal approximation by a finite dimensional Gaussian. These quantizations are then enhanced to rough paths and we prove that the rate of convergence for the enhanced quantization to the enhanced Brownian motion is asymptotically the same.

Quantization for rough paths has been first studied in [PS11]. The choice of Karhunen Loève expansion and method of construction used in this work, namely the trigonometric functions, best suits approximations of Brownian motion in the \( L^2([0,T];\mathbb{R}^d) \) norm. Although this is enough to ensure convergence in the Hölder norm, it is far from efficient and (to the best of our knowledge) no literature exists for rates of convergence. Our approach is demonstrated to be arbitrarily close to optimal and we provide upper and lower bounds on the rate of convergence.

The key advantage of this deterministic construction over the use of empirical distribution used in McKean Vlasov numerics is that we avoid all difficulties with characterising the support (a deterministic set) from random samples. For instance, the almost-sure rate of convergence for an Empirical distribution may, for a particular sample, be too poor to be of any effective use.

By solving the system of interacting Rough Differential Equations driven by a Hölder quantization of the Brownian motion and exploiting the continuity properties of Rough Differential Equations, we obtain a deterministic, finite support measure that approximates the law of the McKean Vlasov Equation without having to solve the law explicitly. One could equivalently obtain the solution law by solving the non-linear Fokker Planck equation, but a novelty of this work is to attain the law without having to resort to PDE methods. We initiate our study by developing our results entirely within the framework established in [CL15].

It is also worth emphasising that the rate of convergence that we obtain in Theorem 3.12 is, at face value, much slower than other well known methods for sampling a measure. The reason for this is we approximate in pathspace rather than for any fixed choice of time. Thus our quantization encodes both information about the path of a Brownian motion and the Hölder regularity.

Finally, to prove the support of McKean Vlasov Equation we develop a novel method by considering the sequences of pairs \((H_n, \mathcal{L}_n)\) \(n \in \mathbb{N}\) and \((T_n^h, \mathcal{L}_n)\) \(n \in \mathbb{N}\) where \((\mathcal{L}_n)\) is a sequence of measures that converge to the law of the McKean Vlasov Equation. However, for each \(n \in \mathbb{N}\), the Skeleton process \(\Phi(h, \mathcal{L}_n)\) driven by \(\mathcal{L}_n\) and a reproducing kernel Hilbert space path \(h\) are not necessarily contained in the support even though they are a good approximation of a path that is contained in the support. Thus our statement for the support takes the form (see Theorem 5.8 below)

\[
\text{supp}(\mathcal{L}) = \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} \left\{ \Phi(h, \mathcal{L}_m) : h \in \mathcal{H} \right\}^{\alpha-\text{Höl}}.
\]

In this paper, we prove two support theorems, see Theorem 5.8 and Theorem 5.13. The first is for McKean Vlasov Equations where the initial condition is deterministic while the second is an extension of this result for McKean Vlasov Equations with random initial condition. The proof of the extension is simple and follows from [CFN97] so we focus predominantly on the first case.

Lastly, we highlight some similarities between our statement (and proof) of the support theorem and the brilliant approach by Hairer and Schönbauder in [HST19, Theorem 1.3] studying the support of the solution to Singular Stochastic Partial Differential Equations using Regularity Structures. Both
results are stated in this non-standard way (crucially due to the approach), namely, as the restriction to the limit points of a collection of smooth paths. Our work was developed independently and was first presented at the 10th Oxford-Berlin Young Researchers Meeting on Applied Stochastic Analysis in December 2018.

We point out that we make full use of rough paths techniques in our manuscript but restrict ourselves to McKean Vlasov equations driven by Brownian motion (as opposed to general Gaussian noises as in [CL15]). The reason for this is the challenges associated with constructing a Truncation and Quantization for a general Gaussian driving noise. In particular, for general Gaussian processes one loses the neat truncation properties given by the reproducing kernel Hilbert space being spanned by orthonormal Schauder functions. It is noteworthy to point out that there is no difference between the construction of the quantization of a Brownian motion and that of a Brownian bridge. This is because Schauder wavelets are also orthonormal in the reproducing kernel Hilbert space of a Brownian Bridge.

All in all, there are several works that address quantization of more general Gaussian processes, for instance [LP06], [DS06] and [LP08]. These works study functional quantization with respect to the $L^p$ norm and supremum norm over time rather than the Hölder norm and so do not properly encode all of the necessary regularity information to solve Rough Differential Equations efficiently. We will shortly address this problem.

This work is organized as follows. We recall several crucial definition and results in Section 2. In Section 3 we discuss the construction of a finite support measure that approximates the law of an enhanced Brownian motion as a measure over the space of geometric rough paths using functional quantization. In Section 4 we visit the construction of rough McKean Vlasov Equations. The support theorem for the class of McKean Vlasov equations addressed in this work in presented in Section 5.

2 Preliminaries

2.1 Notation and spaces

We denote by $\mathbb{N} = \{1, 2, \cdots \}$ the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$; $\mathbb{Z}$ and $\mathbb{R}$ denote the set of integers and real numbers respectively. $\mathbb{R}^+ = [0, \infty)$. By $\lfloor x \rfloor$ we denote the largest integer less than or equal to $x \in \mathbb{R}$. $\mathbb{1}_A$ denotes the usual indicator function over some set $A$. Let $e_j$ be the unit vector in the $j^{th}$ component.

For sequences $(f_n)_{n \in \mathbb{N}}$ and $(g_n)_{n \in \mathbb{N}}$, we denote

$$f_n \lesssim g_n \iff \limsup_{n \to \infty} \frac{f_n}{g_n} \leq C,$$

and

$$f_n \gtrsim g_n \iff \liminf_{n \to \infty} \frac{f_n}{g_n} \geq C,$$

where $C$ is a positive constant independent of the limiting variable. When $f_n \lesssim g_n$ and $f_n \gtrsim g_n$, we say $f_n \asymp g_n$. This is distinct from

$$f_n \sim g_n \iff \lim_{n \to \infty} \frac{f_n}{g_n} = 1.$$ 

Let $C^0([0, T]; \mathbb{R}^d)$ be the space of continuous functions over the interval $[0, T]$ taking values in the vector space $\mathbb{R}^d$ that start at 0 paired with the supremum norm. For $\alpha \in (0, 1)$, we define the $\alpha$-Hölder norm

$$\|\psi\|_\alpha = \sup_{s, t \in [0, T]} \frac{|\psi(t) - \psi(s)|}{|t - s|^\alpha}.$$
Let \( C^\alpha([0,T];\mathbb{R}^d) \) be the subset of \( C^0([0,T];\mathbb{R}^d) \) such that \( \| \cdot \|_\alpha \) is finite. For \( \alpha < \beta < 1 \), \( \beta \)-Hölder continuous paths are compactly embedded in the space of \( \alpha \)-Hölder continuous paths e.g. the spaces \( C^\beta([0,T];\mathbb{R}^d) \subseteq C^\alpha([0,T];\mathbb{R}^d) \). Although the space \( C^\alpha([0,T];\mathbb{R}^d) \) is not separable, the subset \( C^{\alpha,0}([0,T];\mathbb{R}^d) := \overline{C^\beta([0,T];\mathbb{R}^d)}^{\alpha \text{-Hölder}} \) is separable.

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space carrying a \( d \)-dimensional Brownian Motion on the interval \([0,T]\) where throughout \( T > 0 \). The filtration on this space satisfies the usual assumptions. We denote by \( \mathbb{E} \) and \( \mathbb{E}[\cdot|\mathcal{F}_t] \) the usual expectation and conditional expectation operator (with respect to \( \mathbb{P} \)) respectively. For a random variable \( X \) we denote its probability distribution (or law) by \( \mathbb{P}^X \), the law of a process \((X_t)_{t \in [0,T]}\) at time \( t \) is denoted by \( \mathbb{P}^X_t \).

For \( \mu \), a probability measure on \((E,\mathcal{E})\), we define the support of \( \mu \), denoted \( \text{supp}(\mu) \), to be the set of points \( x \in E \) such that every open neighbourhood of \( x \) has positive measure. Equivalently, it is the smallest closed set of full measure.

### 2.2 Gaussian Theory

We briefly summarise some standard results relating to Gaussian processes and Gaussian measures.

**Definition 2.1.** A centred Gaussian measure \( \mathcal{L} \) on a real separable Banach space \( E \) equipped with its Borel \( \sigma \)-algebra \( \mathcal{E} \) is a Borel probability measure on \((E,\mathcal{E})\) such that the law of each continuous linear functional on \( E \) is Gaussian with mean 0.

Let \( E \) be a separable Banach space. Then it is well known that the Borel \( \sigma \)-algebra and the cylindrical \( \sigma \)-algebra are the same (see for example [Bog98]). Let \( \mathcal{H} \) be the Reproducing Kernel Hilbert Space (RKHS) of the Gaussian measure. We denote the unit ball in the RKHS norm as \( \mathcal{K} \). It is well known that the set \( \mathcal{K} \) is compact in the Banach space topology of \( E \) and \( \mathcal{H} \) is dense in the support of \( \mathcal{L} \).

We consider the law of a Gaussian process as a measure on pathspace, that is a measure over the space of continuous paths starting at 0 in \( \mathbb{R}^d \). We are interested in the space of \( \alpha \)-Hölder continuous paths for \( \alpha < \frac{1}{2\varrho} \) and the topology induced by this norm where \( \varrho \in [1,2) \). For any choice of \( \alpha < \frac{1}{2\varrho} \), we can find \( \alpha < \alpha' < \frac{1}{2\varrho} \) for which the Gaussian process will be \( \alpha' \)-Hölder continuous. Therefore, we will always have that the Gaussian process takes values in \( C^{\alpha,0}([0,T];\mathbb{R}^d) \) and we do not concern ourselves with separability further.

**Definition 2.2 (Haar Functions).** Let \( t \in [0,T] \). For \( p \in \mathbb{N}_0 \) and \( m \in \{1,\ldots,2^p\} \), define the sequence of values

\[
t^0_{pm} = \frac{(m-1)T}{2^p}, \quad t^1_{pm} = \frac{(2m-1)T}{2^{p+1}}, \quad \text{and} \quad t^2_{pm} = \frac{mT}{2^p}.
\]

Define the functions \( H_{00}(t) = 1 \) and

\[
H_{pm}(t) = \begin{cases} \sqrt{2^p T}, & \text{if } t \in [t^0_{pm}, t^1_{pm}), \\ -\sqrt{2^p T}, & \text{if } t \in [t^1_{pm}, t^2_{pm}), \\ 0, & \text{otherwise}. \end{cases}
\]

These are called the Haar functions, an orthonormal collection of functions in \( L^2([0,T];\mathbb{R}) \).

The Schauder functions are similarly defined \( G_{pm}(t) = \int_0^t H_{pm}(s)ds \).

The Haar functions form an orthonormal basis on the space \( L^2([0,T];\mathbb{R}) \) with the canonical inner product. Therefore, we define the Fourier coefficients \( \psi_{pm} = \int_0^T H_{pm}(s)\psi(s)ds \) and the set

\[
\Lambda := \{(p,m) : p \in \mathbb{N}_0, m \in \{1,\ldots,2^p\}\} \cup \{(0,0)\}.
\]
We do not include the pair \((p,m) = (-1,0)\) as throughout we will be dealing with Gaussian processes which are \(0\) at \(t = 0\).

Next, for some continuous path \(\psi\) taking values in \(\mathbb{R}^d\), we define the Schauder Fourier coefficients to be
\[
\psi_{pm} := \langle H_{pm}, d\psi \rangle := \sqrt{2^{p}} [2\psi(t_{pm}^1) - \psi(t_{pm}^0) - \psi(t_{pm}^2)] \in \mathbb{R}^d, \quad \text{for } (p,m) \in \Lambda; \tag{2.1}
\]
additionally \(\psi_{00} := \langle H_{00}, d\psi \rangle = \psi(1) - \psi(0)\). Let us denote \(\Lambda_N = \{(p,m) \in \Lambda : p \leq N\}\) as a truncation of \(\Lambda\).

The following Theorem, often referred to as the Ciesielski Isomorphism, provides the link between wavelet theory and rough paths.

**Theorem 2.3** ([HIPP14]). For \(\alpha > 0\), let \(\| \cdot \|_\alpha\) be the \(\alpha\)-Hölder norm. Let \(\psi \in C^0([0,T]; \mathbb{R}^d)\). We have that \(\| \cdot \|_\alpha\) is equivalent to
\[
\|\psi\|_\alpha = \sup_{(p,m) \in \Lambda} 2^{(\alpha - 1/2)p} |\psi_{pm}|. \quad \tag{2.2}
\]
If, in addition, we have that
\[
\lim_{p \to \infty} 2^{p(\alpha - 1/2)} \sup_{1 \leq m \leq 2^p} |\psi_{pm}| = 0
\]
we say that \(\psi \in C^{\alpha,0}([0,T]; \mathbb{R}^d)\). This space is a separable subset of \(C^{\alpha}([0,T]; \mathbb{R}^d)\).

**Example 2.4** (Ciesielski Representation of Brownian motion). Due to the orthogonality of the Schauder functions in the RKHS of Brownian motion, we can represent Brownian motion as
\[
W_t = \sum_{(p,m) \in \Lambda} W_{pm} G_{pm}(t) \quad t \in [0,T] \tag{2.3}
\]
where \(W_{pm}\) is a sequence of \(d^d\)-dimensional, independent, standard normally distributed random variables. Thus
\[
\|W\|_\alpha = \sup_{(p,m) \in \Lambda} 2^p(\alpha - 1/2)|W_{pm}|.
\]

### 2.3 Measures and Approximation

For \(E\) a complete, separable metric space with Borel \(\sigma\)-algebra \(E\), let \(\mathcal{P}_r(E)\) be the set of all Borel measures over \((E,\mathcal{E})\) which have finite \(r^{th}\) moments.

**Definition 2.5.** Let \(\mu, \nu \in \mathcal{P}_r(E)\). We define the Wasserstein \(r\)-distance \(\mathcal{W}^{r}_{E,d}(\mu, \nu) : \mathcal{P}_r(E) \times \mathcal{P}_r(E) \to \mathbb{R}^+\) to be
\[
\mathcal{W}^{r}_{E,d}(\mu, \nu) = \left( \int_{E \times E} d(x,y)^r \gamma(dx,dy) \right)^{1/r} \tag{2.4}
\]
where \(\gamma\) is a joint distribution over \(E \times E\) which has marginals \(\mu\) and \(\nu\). When the space the measure is defined on is clear, we write \(\mathcal{W}^{r}_d(\mu, \nu)\) where \(d\) is the metric over \(E\).

The problem of finding a measure \(\gamma \in \mathcal{P}_2(E \times E)\) that minimises (2.4) is sometimes referred to as the Kantorovich problem and \(\gamma\) is called the transport plan of \(\mu\) and \(\nu\). The choice of \(r = 2\) is common throughout literature. However, we will also be interested in the case \(r = 1\). The \(r\)-Wasserstein distance induces the topology of weak convergence of measure as well as convergence in moments of order up to and including \(r\). The Wasserstein distance is a metric, but the metric does not induce a norm. The Wasserstein distance is homogeneous but not translation invariant. The space \(\mathcal{P}_2(E)\) is complete and separable with respect to the Wasserstein metric (see [Bo08]).
2.3.1 Quantization of Measures

We provide a brief introduction to the field of quantization. For further details, see [GL00].

**Definition 2.6.** Let $L$ be a measure on a separable Banach space $E$ endowed with the Borel $\sigma$-algebra such that $L \in \mathcal{P}_2(E)$ and all sets of codimension 1 have null $L$-measure.

Let $I$ be a countable index, let $\mathcal{S} := \{s_i, i \in I\}$ be a partition of $E$ and let $\mathcal{C} := \{c_i \in E, i \in I\}$ be a codebook. For any partition $\mathcal{S}$ and codebook $\mathcal{C}$, we define a quantization $q : E \to E$ by

$$q(x) = c_i \quad \text{for} \quad x \in s_i,$$

so that

$$L \circ q^{-1}(\cdot) = \sum_{i \in I} L(s_i) \delta_{c_i}(\cdot) \in \mathcal{P}_2(E).$$

The collection of all quantizations is denoted $\mathcal{Q}$.

**Definition 2.7.** Let $\mathcal{P} \subset [0, 1]^\mathbb{N}$ be the set of probability vectors e.g. for every $p = (p_i)_{i \in \mathbb{N}}$, we have $p_i \in [0, 1]$ and $\sum_{i \in \mathbb{N}} p_i = 1$.

Given a partition $\mathcal{S}$ of a space $E$, we have that the sequence $(L(s_i))_{s_i \in \mathcal{S}}$ is a probability vector.

**Definition 2.8 (Optimal Quantizers).** Let $n \in \mathbb{N}$ and $r \in [1, \infty)$. The minimal $n^{th}$ quantization error of order $r$ of a measure $L$ on a separable Banach space $E$ is defined to be

$$\mathcal{E}_{n,r}(L) = \inf \left\{ \left( \int_E \min_{i \in \mathcal{C}} \|x - c_i\|_r^r dL(x) \right)^{\frac{1}{r}} : \mathcal{C} \subset E, 1 \leq |\mathcal{C}| \leq n \right\}.$$

A codebook $\mathcal{C} = \{c_i, i \in I\}$ with $1 \leq |\mathcal{C}| \leq n$ is called an $n$-optimal set of centres of $L$ (of order $r$) if

$$\mathcal{E}_{n,r}(L) = \left( \int_E \min_{i=1,\ldots,n} \|x - c_i\|_r^r dL(x) \right)^{\frac{1}{r}}.$$

Given a finite collection of elements $(c_i)_{i=1,\ldots,n}$, the optimal way to choose the partition of $E$ is to use the nearest neighbour rule which corresponds to the Voronoi partition

$$s(c_i | (c_j)_{j=1,\ldots,n}) := \left\{ x \in E : \|x - c_i\| = \min_{j=1,\ldots,n} \|x - c_j\| \right\} \quad \text{(2.5)}$$

provided the boundary of the Voronoi sets has measure 0. Sets of the form (2.5) are called Voronoi sets. Similarly, given a finite partition $(s_i)_{i=1,\ldots,n}$ of $E$, the optimal choice of codebook is the centres of mass for the sets $s_i$ with respect to the measure $L$. For brevity of notation, we write $\mathcal{E}_n := \mathcal{E}_{n,2}$.

2.3.2 Stationary Quantization

A Stationary set is a codebook with a special property: the Voronoi sets generated by codebook have barycentres equal to the codebook.

**Definition 2.9.** Let $E$ be a separable Banach space with Borel $\sigma$-algebra $\mathcal{E}$, let $n \in \mathbb{N}$ and let $L$ be a measure on $(E, \mathcal{E})$ such that and all subsets of codimension 1 have null $L$-measure. Let $\mathcal{C} \subset E$ satisfy $|\mathcal{C}| = n$.

Suppose that the Voronoi partition $\mathcal{S}$ of $E$ generated by the elements of $\mathcal{C}$, containing the collection of sets $s_i := \{y \in E : \min_{j=1,\ldots,n} \|y - c_j\| = \|y - c_i\|\}$ satisfies that

$$\frac{1}{L(s_i)} \int_{s_i} y dL(y) = c_i.$$

Then we call the codebook $\mathcal{C}$ an $n$-stationary set of the law $L$. 

8
Theorem 2.10 ([Lal10, Theorem 2.1]). Let $E$ be a reflexive, separable Banach space and let $\mathcal{L}$ be a measure on $(E, \mathcal{E})$. For $c_i \in E$, define $A : E^n \to \mathbb{R}$ by

$$A(c_1, ..., c_n) = \int_{E} \min_{i=1,...,n} \|y - c_i\|_E^2 d\mathcal{L}(y)$$

e.g. $A(c_1, ..., c_n)$ is the mean square error between the measure $\mathcal{L}$ and the quantization with codebook $\{c_1, ..., c_n\}$ and partition equal to the Voronoi sets of the codebook.

Then $A$ admits at least one minimum, and so an $n$-stationary set exists.

Remark 2.11. The proof of the above result relies on the Assumption that the Banach space $E$ is reflexive. In particular, for a non-reflexive space the unit ball will be weak-$^\ast$ compact but not weak compact (see [FHH+01, Theorem 3.31]). By contrast, the functional $A$ can be shown to be weak lower semicontinuous but the proof does not extend to weak-$^\ast$ lower semicontinuity.

In particular, we are interested in Gaussian measures over the Banach space $C^{\alpha,0}([0, T]; \mathbb{R}^d)$, which is not reflexive and so Theorem 2.10 does not apply.

Lastly, it is not clear whether a stationary quantization exists in general.

Lemma 2.12. Let $\mathcal{L}$ be a centred Gaussian measure taking values on the Banach space $E$ and suppose that an $n$-stationary set exists. Let $C$ be an $n$-stationary set. Then $C \subset H$.

Proof. This proof is based on a similar argument first presented in [LP02] which focuses solely on Hilbert spaces. Using that the $n$-stationary set exists, we have that for any $c \in C$

$$c = \int_{E} x d\mathcal{L}(x) = \int_{E} x \cdot \frac{1}{\mathcal{L}(\mathbf{g})} d\mathcal{L}(x)$$

Next, we use that $\frac{1}{\mathcal{L}(\mathbf{g})}$ is a square integrable function with respect to $\mathcal{L}$ on $E$ and use Definition of the RKHS to conclude that the right hand side of this equation must be an element of $H$. Therefore $c \in H$.

Remark 2.13. In particular, if $q_n(W)$ denotes the quantized random variable $W$, then the Stationary quantization has the property that

$$q_n(W) = E[W|\mathcal{F}_n],$$

where $\mathcal{F}_n$ is the $\sigma$-algebra generated by the partition of $q_n$. This is a particularly useful property when it comes to establishing uniform integrability of quantizations due to the following simple argument:

Let $\phi$ be a convex function on a Banach space $E$. Then

$$\sup_{n \in \mathbb{N}} E\left[\phi(q_n(W))\right] = \sup_{n \in \mathbb{N}} E\left[\phi(E[W|\mathcal{F}_n])\right] \leq \sup_{n \in \mathbb{N}} E\left[\phi(W|\mathcal{F}_n)\right] = E\left[\phi(W)\right].$$

(2.6)

Lemma 2.14. Let $\mathcal{L}$ be a non-degenerate Gaussian measure over $E$ with RKHS $H$. Let $U$ be a finite dimensional subspace of $H$ and let $P_U$ be the orthogonal projection operator from $H$ to $U$ extended to $E = H^E$. Then $\forall r > 1$

$$\mathcal{E}_{n,r}(\mathcal{L}) \lesssim \mathcal{E}_{n,r}(\mathcal{L} \circ (P_U)^{-1}) + \left(\int_{E} \|x - P_U[x]\|_E^r d\mathcal{L}(x)\right)^{1/r}.$$

In particular, when the measure $\mathcal{L}$ is in some sense “concentrated” on a finite dimensional linear subspace of the Banach space $E$, then the quantization problem can be simplified to a finite dimensional problem.

Proof. See Appendix B.
2.3.3 Rate of Convergence for Quantization

In the finite dimensional setting, the minimal quantization error is well understood (see [GL00]). Let \( L \) be a measure over a \( d \)-dimensional vector space. Then
\[
E_{n,r}(L) \approx n^{1/d}.
\]
However, for a Gaussian measure over a Banach space \( E \), the limit \( d \to \infty \) is no longer meaningful. In both [DFMS03] and [GLP03], the authors investigate the relation between the minimal quantization error and the probabilities of small balls.

**Theorem 2.15** ([DFMS03], [GLP03]). Let \( L^W \) be a Gaussian measure over a Banach space \( E \). Let \( \mathcal{B}_W \) be the small ball probability of \( L^W \) defined by
\[
\mathcal{B}_W(\varepsilon) := -\log \mathcal{L}\{x \in E : \|x\|_E < \varepsilon\}.
\]
Then for any choice of \( r \geq 1 \)
\[
E_{n,r}(L^W) \approx \left(\mathcal{B}_W\right)^{-1}\left(\log(n)\right)
\]
as \( n \to \infty \). In particular, let \( L^W \) be the law of Brownian motion over \( C^{\alpha,0}([0,T];\mathbb{R}^{d'}) \). Then by the results of [BR92]
\[
E_{n,r}(L^W) \approx d'(\log(n^{1/d'}))^\alpha^{-1/2}.
\]

In particular, Equation (2.8) provides us with a lower bound that the error of the quantization for Brownian motion cannot outstrip. However, as already explained in Remark 2.11, there may not exist a stationary quantization that attains \( E_{n,r}(L^W) \).

A remarkable aspect of [DFMS03] is that the authors additionally prove that the mean square error between an empirical measure and the true Gaussian measure in the Wasserstein distance converges at the same rate as the optimal quantization error.

2.4 Rough Paths

Throughout this paper, we will use the notation for increments of a path \( X_{s,t} = X_t - X_s \) for \( s \leq t \). Rough paths were first introduced in [Lyo98]. For a detailed overview of rough path theory, see [FV10b], [FH14] and [LQ02]. For a self-contained exposition, the reader can find a primer on rough paths in Appendix A.

2.4.1 The lift of Gaussian Processes

In [FV10a], the authors prove that when the covariance operator of the Gaussian satisfies a \( p \)-variation condition, the path of the Gaussian can be lifted to a finite \( p \)-variation or \( \alpha \)-Hölder continuous rough path.

**Assumption 2.16.** Let \( L^W \) be the law of a continuous, centred Gaussian process with independent components taking values in \( \mathbb{R}^{d'} \) and covariance operator \( R \) such that \( \|R\|_{\ell^2|s,t|^2} \leq M|t - s|^{1/\phi} \) for \( \phi \in [1,2) \) and \( M < \infty \) with

3 Approximation of Brownian motion

The goal of this Section is to construct a finite support measure that approximates the law of an enhanced Brownian motion as a measure over the space of geometric rough paths.
3.1 Truncation of Brownian Motion

Using the Cielsielski representation for Brownian motion from Equation (2.3), we can obtain a finite dimensional Gaussian measure on \( C^{\alpha,0}([0, T]; \mathbb{R}^d) \)

\[
W_t^N = \sum_{(p,m) \in \Lambda_N} W_{pm} G_{pm}(t),
\]

which approximates Brownian motion. Let us briefly describe some of the properties this random variable:

- \( W^N \) is a Gaussian measure on \( C^{\alpha,0}([0, T]; \mathbb{R}^d) \) with RKHS \( \mathcal{H}_N = \langle \text{span}_{(p,m) \in \Lambda_N} \{ G_{pm} \} \rangle^{d'} \).
- As a finite dimensional Gaussian, the support of \( W^N \) is just \( \mathcal{H}_N \). This is equal to the space of \( d' \)-dimensional, piecewise linear paths over the dyadic intervals of size \( T 2^{-N} \).
- This is the optimal finite dimensional approximation of Brownian motion with respect to the \( \alpha \)-Hölder norm (2.2).
- The support of the measure \( \mathcal{L}_W^N \) is Reflexive, so by Theorem 2.10 a stationary quantization exists.

3.1.1 Optimality of the Truncation

We prove that the truncation chosen in (3.1) is the optimal choice with respect to the \( \alpha \)-Hölder norm. This is an application of the results of [BC19].

**Proposition 3.1.** Let \( \mathcal{L}_W^N \) be the law of Brownian motion over the Banach space \( C^{\alpha,0}([0, T]; \mathbb{R}^d) \). Then the \( d' \cdot 2^N \) dimensional projection \( P : C^{\alpha,0}([0, T]; \mathbb{R}^d) \rightarrow C^{\alpha,0}([0, T]; \mathbb{R}^d) \) that minimises the integral

\[
\mathbb{E} \left[ \left\| W - P[W] \right\|_{\alpha-\text{Höld}}^2 \right],
\]

is the projection

\[
P[W]_t = \sum_{(p,m) \in \Lambda_N} W_{pm} G_{pm}(t).
\]

**Proof.** See Appendix B.

3.1.2 Rate of Convergence of the Truncation

We measure the rate of convergence for a truncated Brownian motion with respect to the \( \alpha \)-Hölder norm. We point out that the Banach space \( C^{\alpha,0}([0, T]; \mathbb{R}^d) \) is not \( K \)-convex (see [Pis89, Definition 2.3]) so consequently the upper and lower bounds of the rate of convergence cannot be the same.

**Proposition 3.2.** Let \( W \) be a Brownian motion as expressed in (2.3) and let \( W^N \) be truncated Brownian motion (3.1). Then for \( r > 1 \) we have

\[
c \cdot d' \cdot N^{1/2-\alpha} \cdot 2^{(\alpha-1/2)N} \leq \mathbb{E} \left[ \left\| W - W^N \right\|_{\alpha}^{1/r} \right]^{1/r} \leq C \cdot d' \cdot \sqrt{N} \cdot 2^{(\alpha-1/2)N},
\]

where the constants \( c \) and \( C \) dependent only on \( \alpha \) and \( r \).
Proof. Using Theorem 2.3 and the methods of [BR92], we have
\[ \mathcal{B}_W(\varepsilon) = -\log \left( \mathbb{P} \left( \|W\|_\alpha < \varepsilon \right) \right) \approx d' \cdot \left( \frac{\varepsilon}{\pi} \right)^{1/2-\alpha} \] and \[ \mathcal{B}_W(\varepsilon) \lesssim \mathcal{B}_W(2\varepsilon). \]

Then by [LL99] Proposition 4.1, this implies
\[ d' \cdot N^{1/2-\alpha} \cdot 2^{(\alpha-1/2)N} \lesssim \mathbb{E} \left[ \sup_{(p,m) \in \Lambda \setminus \Lambda_N} |W_{pm}|^{2p(2\alpha-1)} \right]^{1/2} \lesssim d' \cdot \sqrt{N} \cdot 2^{(\alpha-1/2)N}, \]
as \( N \to \infty \) since \( W^N \) is a \( d \cdot 2^{N+1} \)-dimensional Gaussian random variable.

The Gaussian random variables \( W - W^N \) can be dominated by \( W \). By using a concentration inequality and a standard hypercontractivity argument, we can find a constant \( C = C(r) \) such that
\[ \mathbb{E} \left[ \|W - W^N\|_\alpha^r \right] \leq C(r) \mathbb{E} \left[ \|W - W^N\|_\alpha^2 \right]^{\frac{r}{2}}. \]

Thus the rate of convergence in mean square is equivalent to the rate of convergence for any choice of \( r \).

\[ \square \]

3.1.3 Enhanced Truncated Brownian Motion

Finally, we prove that the rate of convergence of the enhanced truncated Brownian motion to the enhanced Brownian motion is the same when the process is lifted to a rough path and studied with respect to the inhomogeneous metric.

The rate of convergence for an enhanced piecewise linear approximation of a Brownian motion has already been studied in [FR11]. Our contribution is a sharper rate of convergence.

Proposition 3.3. Let \( N \in \mathbb{N} \) and let \( M \geq 2 \). Let \( L^N_W \) be the law of the truncated Brownian motion over the Banach space \( C^{\alpha,0}([0,T];\mathbb{R}^d) \). Then \( L^N_W \) satisfies Assumption 2.16 hence \( W^N \) can be lifted to an enhanced Gaussian rough path \( W^N = S_M(W^N) \) taking values on the Group \( G^M(\mathbb{R}^d) \) for \( M \geq 2 \). Further, for the enhanced Brownian motion \( W \) taking values in \( G\Omega_\alpha(\mathbb{R}^d) \), there exists a constant \( C = C(M, d', \alpha) \) such that
\[ \mathbb{E} \left[ \rho_i(W^N_{s,t}, W_{s,t})^2 \right] \leq C N^2 (2\alpha-1)N |t - s|^i, \quad (3.3) \]
where \( i \in \{1, \ldots, M\} \) and \( \rho_i \) is the tensor pseudo-metric (A.5) over \( T^M(\mathbb{R}^d) \).

Proof. The case \( i = 1 \) is immediate. We address \( i = 2 \) briefly. For \( j, k \in \{1, \ldots, d'\} \) and \( j \neq k \)
\[ \mathbb{E} \left[ \int_s^t \langle W_{s,r}, e_j \rangle \circ d\langle W_r, e_k \rangle - \int_s^t \langle W^N_{s,r}, e_j \rangle \circ d\langle W^N_r, e_k \rangle \right]^2 \]
\[ \leq \int_s^t \int_s^t R_{\langle W-W^N, e_j \rangle} \left( \frac{s}{u}, \frac{s}{v} \right) dR_{\langle W, e_k \rangle}(u, v) \]
\[ + \int_s^t \int_s^t R_{\langle W^N, e_j \rangle} \left( \frac{s}{u}, \frac{s}{v} \right) dR_{\langle W^N-W, e_k \rangle}(u, v) \]
\[ \leq C |t - s|^2 \cdot \mathbb{E} \left[ \|W - W^N\|_\alpha^2 \right] \cdot \mathbb{E} \left[ \|W\|_\alpha^2 \right]. \]
Compiling these terms by summing over \( j \) and \( k \) completes the \( i = 2 \) case.

For \( i > 2 \), we argue by induction. For a word \( A \) such that \( |A| = i \) and letter \( a \in \mathcal{A} \), we have

\[
\mathbb{E} \left[ \int_s^t \langle W_{s,r}, e_{(A,a)} \rangle \circ d\langle W_r, e_{(A,a)} \rangle - \int_s^t \langle W_{s,r}^N, e_{(A,a)} \rangle \circ d\langle W_r, e_{(A,a)} \rangle \right]^2
\]

\[
\leq \int_s^t \int_s^t \mathbb{E} \left[ \rho_i(W_{s,u}, W_{s,u}^N) \rho_i(W_{s,v}, W_{s,v}^N) \right] dR_{(W, e_{a})}(u, v)
\]

\[
+ \int_s^t \int_s^t \mathbb{E} \left[ (W_{s,u}, e_{A}) \cdot (W_{s,v}, e_{A}) \right] dR_{(W^N - W, e_{a})}(u, v)
\]

\[
\leq C |t - s|^{i+1} \cdot \mathbb{E} \left[ \|W - W^N\|_{\alpha}^2 \right],
\]

which implies the inductive hypothesis.

\[\square\]

**Theorem 3.4.** Let \( N \in \mathbb{N} \) and \( M \geq 2 \). Let \( r > 1 \). Let \( W^N \) be the enhanced truncated Brownian motion and let \( W \) be the enhanced Brownian motion over \( G^M(\mathbb{R}^d) \). Then

\[
\mathbb{E} \left[ \rho_{\alpha-\text{H"{o}l}}(W, W^N)^r \right]^{1/r} \lesssim \sqrt{N} \cdot 2^{(\alpha-1/2)N}
\]

as \( N \to \infty \). Also

\[
\mathbb{E} \left[ d_{\alpha-\text{H"{o}l}}(W, W^N)^r \right]^{1/r} \lesssim \max \left\{ \sqrt{N}2^{(\alpha-1/2)N}, (\sqrt{N}2^{(\alpha-1/2)N})^{1/M} \right\}. \tag{3.5}
\]

**Proof.** Firstly, it should be clear that we have

\[
\mathbb{E} \left[ \rho_{\alpha-\text{H"{o}l}}(W, 1)^2 \right] < C \quad \text{and} \quad \mathbb{E} \left[ \rho_{\alpha-\text{H"{o}l}}(W^N, 1)^2 \right] < C.
\]

Then, we apply [FV10b][Theorem 15.24] with Proposition 3.3 to get Equation (3.4) in the case \( r = 2 \).

For (3.5), we use the well known fact that the the identity operator is \( \frac{1}{M} \)-Hölder continuous from the space of rough paths paired with the Inhomogeneous metric to the space of rough paths paired with the homogeneous metric and \( r = 2 \).

Now for the case \( r \neq 2 \). Following [Rie17, Corollary 3.2], we can conclude that the pushforward of \( d_{\alpha}(W, W^N) \) with respect to the measure \( L^W \) has a Gaussian tail uniformly on \( N \) since the covariance of \( W - W^N \) can be dominated by the covariance of \( W \). Then we use a hypercontractivity argument to conclude that

\[
\mathbb{E} \left[ d_{\alpha}(W, W^N)^r \right] \leq C(r) \mathbb{E} \left[ d_{\alpha}(W, W^N)^2 \right]^{r/2}.
\]

Thus the rate of convergence in mean square is equivalent to the rate of convergence for any choice of \( r \). \[\square\]

### 3.2 Quantization of Brownian Motion

We perform a truncation to obtain a finite dimensional Gaussian that represents an optimal finite dimensional approximation of the Brownian motion. Here, we study how the choice of truncation affects the asymptotic rate of convergence of the quantization error.
Remark 3.5. $L^{W,N}$ is a non-degenerate measure over the (finite dimensional) vector space $(\mathcal{H}_N, \| \cdot \|_\alpha)$. Therefore by Theorem 2.10 we know that there exists a codebook $\mathcal{C}_n = \{c_1, \ldots, c_n\}$ and a partition $\hat{\mathcal{S}}_n = \{\hat{s}_1, \ldots, \hat{s}_n\}$ of $\mathcal{H}_N$ such that the quantization $\hat{q}_n$, satisfies

$$W^{(2)}_{\mathcal{H}_N, \| \cdot \|_\alpha} \left( L^{W,N} \mid_{\mathcal{H}_N}, L^{W,N} \circ \hat{q}^{-1}_n \mid_{\mathcal{H}_N} \right) = E \left[ \| W^{N} - \hat{q}_n(W^{N}) \|_{\alpha}^2 \right]^{1/2} = \mathcal{C}_n$$

However, the measure $L^{W,N}$ is degenerate over the whole space $C^{\alpha,0}([0,T]; \mathbb{R}^{d'})$ so constructing an optimal quantization becomes analytically problematic.

Definition 3.6. Let $N \in \mathbb{N}$ be fixed for the moment and let $n \in \mathbb{N}$. Let $L^{(W)}$ be the law of Brownian motion over $C^{\alpha,0}([0,T]; \mathbb{R})$ and let $L^{(W,N,\varepsilon_1)}$ be the law of the 1-dimensional truncated Brownian motion with RKHS $\mathcal{H}_N^{(1)}$. Let $\mathcal{C}_n^{(1)} = \{c_1^{(1)}, \ldots, c_n^{(1)}\}$ and $\hat{\mathcal{S}}_n^{(1)} = \{\hat{s}_1^{(1)}, \ldots, \hat{s}_n^{(1)}\}$ be the codebook and partition of the stationary quantization of $L^{(W,N,\varepsilon_1)}$ over $\mathcal{H}_N^{(1)}$.

Let $\mathcal{C}_n := (\mathcal{C}_n^{(1)})^{\times d'}$ and $\hat{\mathcal{S}}_n := (\hat{\mathcal{S}}_n^{(1)})^{\times d'}$. Thus $\mathcal{C}_n$ and $\hat{\mathcal{S}}_n$ form a quantization of the truncated Brownian motion over $\mathcal{H}_N^{(1)}$ with independent components. Let $P_N : \mathcal{H} \rightarrow \mathcal{H}_N$ be the orthogonal projection and let us continuously extend $P_N$ to $\mathcal{H}$.

We define the new partition of $\mathcal{H}$ to be

$$s_i := (P_N)^{-1}[\hat{s}_i], \quad \hat{\mathcal{S}}_n := \{s_1, \ldots, s_n\}. \quad (3.6)$$

Pairing the partition $\hat{\mathcal{S}}_n$ with the codebook $\mathcal{C}_n$, we obtain a quantization for the truncated Brownian motion over $\mathcal{H}_N^{\alpha-\text{Hol}}$.

It is worth noting that the codebook $|\mathcal{C}_n| = n^{d'}$. We should also emphasise that the quantization constructed in Definition 3.6 is not an optimal quantization of the measures $L^{W}$ or $L^{W,N}$ over the whole space. The reason for this approach is that this quantization exists and is solvable.

Lemma 3.7. Let $n, N \in \mathbb{N}$. Let $L^W$ be the law of a Brownian motion over $C^{\alpha,0}([0,T]; \mathbb{R}^{d'})$ with quantization $q_n$ as defined in Definition 3.6.

Let $i \neq j \in \{1, \ldots, d'\}$. Then $\langle q_n(W), e_i \rangle$ and $\langle q_n(W), e_j \rangle$ are independent.

Proof. For any two sets $C, D \in C^{\alpha,0}([0,T]; \mathbb{R})$, we have

$$\mathbb{P} \left[ \langle q_n(W), e_i \rangle \in C, \langle q_n(W), e_j \rangle \in D \right] = L^W \left[ \left( \bigcup_{\langle \tilde{k}, e_i \rangle \in C \cap \mathcal{H}_N} \tilde{s}_k \right) \cap \left( \bigcup_{\langle \tilde{e}_l, e_j \rangle \in D \cap \mathcal{H}_N} \tilde{s}_l \right) \right] = L^{W,N} \left[ \left( \bigcup_{\langle \tilde{k}, e_i \rangle \in C \cap \mathcal{H}_N} \tilde{s}_k^{(1)} \right) \cap \left( \bigcup_{\langle \tilde{e}_l, e_j \rangle \in D \cap \mathcal{H}_N} \tilde{s}_l^{(1)} \right) \right] = L^{(W,N,\varepsilon_1)} \left[ \bigcup_{\langle \tilde{k}, e_i \rangle \in C \cap \mathcal{H}_N^{(1)}} \tilde{s}_k^{(1)} \right] \cdot L^{(W,N,\varepsilon_1)} \left[ \bigcup_{\langle \tilde{e}_l, e_j \rangle \in D \cap \mathcal{H}_N^{(1)}} \tilde{s}_l^{(1)} \right] = \mathbb{P} \left[ \langle q_n(W), e_i \rangle \in C \right] \cdot \mathbb{P} \left[ \langle q_n(W), e_j \rangle \in D \right].$$

$\square$
3.2.1 Asymptotic rate of convergence for Quantization

Next, we apply Theorem 2.15 with Proposition 3.2 in order to demonstrate the rate of convergence of the quantization we construct.

**Proposition 3.8.** Let $L^W$ be the law of a Brownian on $C^{\alpha,0}([0,T];\mathbb{R}^d)$ and let $L^{W,N}$ be the law of the truncated Brownian motion. Choose $N$ to satisfy

$$N \approx W\left(\frac{\ln(2^{2\alpha - 1}) \cdot \log(n)^{2\alpha - 1}}{\ln(2^{2\alpha - 1})}\right),$$

(3.7)

where $W$ is the Lambert-$W$ function (see [BL16]), the inverse function of $y = xe^x$.

Then $\forall r > 1$, the quantization constructed in Definition 3.6 satisfies

$$\left(\int_{C^{\alpha,0}([0,T];\mathbb{R}^d)} \|x - q_n(x)\|_\alpha^r dL^W(x)\right)^{1/r} \lesssim d' \cdot \left(\log(n)\right)^{\alpha - 1/2}$$

(3.8)

as $n \to \infty$.

**Proof.** It should be clear that the partition as defined in Equation (3.6) is not the collection of Voronoi sets generated by the codebook $\mathcal{C}_n$ over $H^{\alpha,\text{H"ol}}$. Thus

$$\mathbb{E}\left[\|W - q_n(W)\|_\alpha^r\right]^{1/r} \geq \left(\int_{\mathbb{R}^d} \min_{\epsilon_j \in \mathcal{C}} \|x - \epsilon_j\|_\alpha dL^W(x)\right)^{1/r}.$$

We can further improve this lower bound by minimizing over the all possible codebooks $\mathcal{C}$ which yields the lower bound

$$d' \left(\log(n)\right)^{\alpha - 1/2} \lesssim \mathcal{E}_{n,d',r}(L^W) \leq \mathbb{E}\left[\|W - q_n(W)\|_\alpha^r\right]^{1/r}.$$

For the upper bound, we apply Lemma 2.14 and Proposition 3.2 to get

$$\mathbb{E}\left[\|W - q_n(W)\|_\alpha^r\right]^{1/r} \leq \mathbb{E}\left[\|W - q_n(W)\|_\alpha^r\right]^{1/r} + \mathbb{E}\left[\|W - W^N\|_\alpha^r\right]^{1/r} \lesssim \mathcal{B}_{W}^{-1}\left(\log(n)\right) + \sqrt{N} \cdot 2^{(\alpha - 1/2)N}.$$

By Theorem 2.15, we have asymptotic upper and lower bounds on the quantization error for both measures $L^W$ and $L^{W,N}$.

Due to the nice way in which the truncation and the Hölder norm overlap, we have that

$$\mathbb{P}\left[\|W^N\|_\alpha \leq \epsilon\right] \geq \mathbb{P}\left[\|W\|_\alpha < \epsilon\right],$$

or equivalently

$$- \log\left(\mathbb{P}\left[\|W^N\|_\alpha \leq \epsilon\right]\right) = \mathcal{B}_{W,N}(\epsilon) \leq \mathcal{B}_{W}(\epsilon) = - \log\left(\mathbb{P}\left[\|W\|_\alpha < \epsilon\right]\right).$$

This is true for any choice of truncation level $N$. Taking the inverse of these bijective, increasing functions gives

$$\mathcal{B}_{W,N}^{-1}(n) \leq \mathcal{B}_{W}^{-1}(n).$$
Thus, for any choice of $N \in \mathbb{N}$,

$$\mathbb{E}\left[\|W - q_n(W)\|_\alpha^r\right]^{1/r} \lesssim d' \left( \log(n) \right)^{\alpha-1/2} + d' \sqrt{N} \cdot 2^{(\alpha-1/2)N}.$$  

Finally, we note that the asymptotic relation of Equation (3.7) is equivalent to

$$\sqrt{N} \cdot 2^{(\alpha-1/2)N} \approx \left( \log(n) \right)^{\alpha-1/2}$$

which yields the conclusion.

\[\square\]

**Remark 3.9.** We know by results such as [DFMS18] that by sampling a Brownian motion in pathspace, the empirical law will be a good approximation for the law of Brownian motion.

The difference with this method is that sampling produces a convergence in measure type result. This is a deterministic and not probabilistic result.

### 3.2.2 Quantization for a Gaussian Rough Paths

For this Section, we explore lifting our quantized Brownian motion to a rough path. Quantization for rough paths was first studied in [PS11]. In their paper, the authors treat the law of Brownian motion as a measure over the Hilbert space $L^2([0, T]; \mathbb{R}^{d'})$. In particular, as a measure over a Hilbert space the authors are able to obtain a stationary quantization, see [LP02]. The Karhunen Loève expansion is obtained using an expansion of trigonometric functions and the authors use well understood pathspace results to establish pointwise convergence of the paths followed by convergence in $p$-variation. To the best of our knowledge, this is the only work studying quantization in a rough path framework so this chapter is new and of independent interest.

We perform quantization for a Brownian rough path with respect to the pathspace Hölder norm. Due to the nature of the $L^2$ norm with which the quantization is constructed in [PS11], the approximation with respect to the Hölder norm is far from optimal. By contrast, our approximation is arbitrarily close to optimal. In this Section, we prove that this remains true when the study is carried out with respect to the rough path Hölder norm.

As proved in Lemma 2.12, the sets $\mathcal{C} \subset \mathcal{H}$ so have a canonical Young integral signature $c = S_M(c)$ for each $c \in \mathcal{C}$.

**Definition 3.10.** Let $M \geq 2$. Let $\mathcal{L}^W$ be the law of a Brownian motion over $C^{\alpha, 0}([0, T]; \mathbb{R}^{d'})$ and let $\mathcal{L}^W$ be the law of the enhanced Brownian motion over $G\Omega_{\alpha}^0(\mathbb{R}^{d'})$. Let $q_n$ be the sequence of quantizations as defined in Definition 3.6 for the truncated Brownian motion with $N$ chosen to satisfy Equation (3.7) and codebooks $\mathcal{C}_n$ and partitions $\mathcal{S}_n$.

Define the sets

$$\mathcal{S} = \left\{ s_1, \ldots, s_n \right\}, \quad s_i := \left\{ h = S_M(h) : h \in s_i \cap \mathcal{H} \right\}_{Fr\text{-Hö}}^{\text{Ab}}.$$

These form a partition over the space $G\Omega_{\alpha}(\mathbb{R}^{d'})$ (up to boundary sets of measure 0). Similarly, define the codebook

$$\mathcal{C} = \{ c_1, \ldots, c_n \}, \quad c_i := S_M[c_i].$$

By combining the enhanced codebook with the partition $\mathcal{S}_n$, we obtain the enhanced quantization $q_n : G\Omega_{\alpha}(\mathbb{R}^{d'}) \to G\Omega_{\alpha}(\mathbb{R}^{d'})$

$$q_n(X) = c_i \quad \text{for} \ X \in s_i, \quad q_n \left( G\Omega_{\alpha}(\mathbb{R}^{d'}) \right) = C_n. \ \ (3.9)$$
The next result is an extension of Proposition 3.8 to the rough path setting. We follow the same methods as in Section 3.1.3.

**Proposition 3.11.** Let $M \geq 2$. Fix $N, n \in \mathbb{N}$. Let $L^W$ be the law of the enhanced Brownian motion. Then there exists a constant $C = C(M, d', \alpha)$ such that

$$E\left[\rho_n(W^N_{s,t}, q_n(W^N_{s,t}))^2\right] \leq C \left(\log(n)\right)^{2\alpha - 1} |t - s|^t. \quad (3.10)$$

**Proof.** The case $i = 1$ is already proved in Proposition 3.8. $i > 2$ can be addressed via an induction argument as in Proposition 3.3. Therefore, we only prove the case $i = 2$. Thus for $\mathcal{F}_n$, equal to the $\sigma$-algebra generated by the partition of $\mathcal{C}_n$, we have

$$\mathbb{E} \left[ \int_s^t \int_s^t \mathbb{E} [W^N_{s,r}, e_i] d\mathbb{E} [W^N_{s,r}, e_i] - \int_s^t \mathbb{E} [W^N_{s,r}, e_i] d\mathbb{E} [W^N_{s,r}, e_i] \right]^2 \leq 2 \int_s^t \int_s^t \mathbb{E} \left[ \mathbb{E} [W^N_{s,r}, e_i] - \mathbb{E} [W^N_{s,r}, e_i] \right] \cdot \mathbb{E} \left[ \mathbb{E} [W^N_{s,r}, e_i] - \mathbb{E} [W^N_{s,r}, e_i] \right] d\mathbb{E} \left[ \mathbb{E} [W^N_{s,r}, e_i] - \mathbb{E} [W^N_{s,r}, e_i] \right] \cdot \mathbb{E} \left[ \mathbb{E} [W^N_{s,r}, e_i] - \mathbb{E} [W^N_{s,r}, e_i] \right],$$

using Lemma 3.7 and the same Young Estimates as in Proposition 3.3.

$$\leq (t - s)^2 \mathbb{E} \left[ |W^N|_{\mathcal{F}_n}^2 \right] \cdot \mathbb{E} \left[ |W^N|_{\alpha}^2 \right] \leq C(t - s)^2 \left(\log(n)\right)^{2\alpha - 1}.$$  

□

**Theorem 3.12.** Let $r > 1$. Let $L^W$ be the law of Brownian motion on $G\Omega_\alpha(\mathbb{R}^{d'})$ and let $L^W$ be the law of the of the enhanced Brownian motion over $G\Omega_\alpha(\mathbb{R}^{d'})$. Let $q_n$ be the sequence of quantizations constructed in Definition 3.10. Then

$$\left(\int_{G\Omega_\alpha(\mathbb{R}^{d'})} \rho_{\alpha - H^\mathbb{R}}(0, T) \left( X, q_n(X) \right) dL^W(X) \right)^{1/r} \leq \left(\log(n)\right)^{\alpha - 1/2}. \quad (3.11)$$

**Proof.** The lower bound of Equation (3.11) is actually immediate from Equation (3.8). The $\rho_{\alpha - H^\mathbb{R}}$ metric can be lower bounded by the projection onto the first level of the Signature so that

$$E \left[ \rho_{\alpha - H^\mathbb{R}}(0, T) \left( W, q_n(W) \right) \right]^2 \geq E \left[ |W - q_n(W)|_\alpha^2 \right].$$

Also, by Theorem 3.4, we know the rate of convergence for

$$E \left[ \rho_{\alpha - H^\mathbb{R}}(W, W^N)^{1/r} \right] \leq \sqrt{N}, \frac{2^{(\alpha - 1/2)N}}{\left(\log(n)\right)^{\alpha - 1/2}}.$$  

where $N$ is the dimension of the linear span of the codebook $\mathcal{C}_n$ and the choice of Equation (3.7) provides the second step. It is clear that

$$E \left[ \rho_{\alpha - H^\mathbb{R}}(W, 1)^2 \right] < C, \quad E \left[ \rho_{\alpha - H^\mathbb{R}}(W^N, 1)^2 \right] < C.$$
We can then apply [FV10b] (Theorem A.13) with Proposition 3.11. We remark that although this method has been used to prove the regularity of enhanced gaussian rough paths before, there is no part of this method that requires the Gaussian structures, only regularity properties in all moments. Thus
\[ E \left[ \rho_{\alpha}^{-1}(0, T) \left( W^N, q_\alpha(W^N) \right)^{r_1} \right] \lesssim \left( \log(n) \right)^{\alpha - 1/2}. \]

4 Mean Field Rough Differential Equations

In the first Section, we address the approach of [CL15] to solve McKean Vlasov Rough Differential Equations driven by a Brownian rough path. There, the authors prove Existence, Uniqueness and a Propagation of Chaos result for McKean Vlasov Rough Differential Equations of the form
\[ dX_t = \sigma(X_t) dW_t + b(X_t) d\gamma^X_t, \quad X_0 = \xi, \quad t \in [0, T], \quad (4.1) \]
where the path \( \gamma^X_t = \int_0^t \mu_t ds \) represents the measure dependency in the drift term. [CL15] includes an explanation as to why the authors were unable to include a measure dependency in the diffusion terms.

Rough Differential Equations with a measure dependency in the drift term are addressed in the more recent preprints [BCD18] and [BCD20]. We choose to present this work in the framework of [CL15] to reduce the complexity and avoid obfuscated algebraic argument.

4.1 Controls and the Accumulated p-Variation

In this first Section, we establish a key condition for the integrability of our quantization. For notational simplicity, we denote \( p = \frac{1}{\alpha} \).

Definition 4.1. Let \( \beta > 0 \) and suppose that \( \omega : \Delta_T \to \mathbb{R}^+ \) is a control (recall Definition A.4). We define the Accumulated \( \beta \)-local \( \omega \)-variation by
\[ M_\beta(\omega) := \sup_{\omega(t_i, t_{i+1}) \leq \beta} \sum_{i : t_i \in D} \omega(t_i, t_{i+1}). \]

The Accumulated \( \beta \)-local controls were first introduced in [CLL13]. We are interested in the specific case where the control is induced by a weakly geometric rough path.

Definition 4.2. Let \( \beta > 0 \). Let \( p > 2 \) and let \( W \in G_{\alpha}(\mathbb{R}^d) \). We define the Accumulated \( \beta \)-local \( p \)-variation of a geometric rough path to a non-negative function defined by
\[ M_{\beta,p}(W) := M_{\beta}(\omega_W p). \]

We define the nondecreasing sequence \( (\tau_i(\beta, p, W))_{i \in \mathbb{N}} \) by
\[ \tau_0(\beta) = 0, \quad \tau_{i+1}(\beta) = \inf\{ t > \tau_i(\beta) : \|W\|_{p\text{-var}; [\tau_i(\beta), t]} \geq \beta \} \land T. \quad (4.2) \]
This is sometimes referred to as the Greedy sequence. Finally, we define the function \( N_{\beta,p,[0,T]} : G_{\alpha}(\mathbb{R}^d) \to \mathbb{N} \cup \{\infty\} \) given by
\[ N_{\beta,p,[0,T]}(W) := \sup\{ n \in \mathbb{N} \cup \{0\} : \tau_n(\beta) < T \}. \]
While stopping time arguments become problematic for McKean Vlasov Equations due to the presence of the measure dependency, we emphasise that the greedy sequence \((4.2)\) is dependent only on the driving noise and not the solution.

It is immediate from the definition that \(M_{\beta,p}(W) \leq \|W\|_{p-\text{var};[0,T]}^p\). However, when \(W\) is a Gaussian rough path and \(p > 2\), we have \(|W_{0,T}|^p \leq \|W_{0,T}\|_{cc}^p \leq \|W\|_{p-\text{var};[0,T]}^p\) and \(W_{0,T} \sim N(0,T)\) so

\[
E\left[ \exp\left(\|W\|_{p-\text{var};[0,T]}^p\right) \right] = \infty.
\]

**Remark 4.3.** The Accumulated \(p\)-variation is a way of restricting the size of the \(p\)-variation in the event that the \(p\)-variation becomes large. When the \(p\)-variation of a Gaussian is large, by far the most probable event is that there is a single large increment of the process. While the \(p\)-variation will increase proportionally to this steep increment, the Accumulated \(\beta\)-local \(p\)-variation is restricted to partitions where the increments cannot be larger than \(\beta\) so the one increment does not make a proportional contribution.

The following Proposition is key to the construction of McKean Vlasov Rough Differential Equations driven by Gaussian processes.

**Proposition 4.4.** Let \(W\) be a continuous, centred Gaussian rough path that satisfies Assumption 2.16 Then \(\forall \beta > 0\), the random variable \(M_{\beta,p}(W)\) has well defined Moment Generating Function

\[
[0, \infty) \ni \theta \mapsto E\left[ \exp(\theta M_{\beta,p}(W)) \right] < \infty.
\]

**Proof.** See [CLL13, Theorem 6.3] for tail estimates of the law of the Accumulated \(p\)-variation.

The existence of a moment generating function for the Accumulated \(p\)-variation of the driving noise for the McKean Vlasov Rough Differential Equation is a key Assumption of [CL15], see below. In order to prove propagation of chaos of a sequence of measures, the authors prove that the sequence of empirical measures each has a moment generating function and that the empirical laws converge weakly to the law of the driving noise. We verify the quantization also satisfies this condition:

**Lemma 4.5** ([FV10a]). Let \(\mathcal{L}W\) be the law of a Brownian motion over \(C^{0,0}(\alpha,T;\mathbb{R}^d)\). Let \(h_1, ..., h_n\) be a collection of orthonormal elements of \(\mathcal{H}\). Let \(W^n\) be a finite Karhunen Loève expansion of \(W\) generated by the set \(\{h_1, ..., h_n\}\) so that

\[W^n = \mathbb{E}[W|\mathcal{F}^n],\]

where \(\mathcal{F}^n\) is the \(\sigma\)-algebra generated by the functionals \(f_j = (i^*)^{-1}[h_j]\) for each \(j = 1, ..., n\).

Then the Brownian rough path \(W = S_2(W)\) satisfies the martingale formula

\[
E\left[ \log_{\mathbb{E}}(W_{s,t}) \left| \mathcal{F}^n \right. \right] = \log_{\mathbb{E}}(W^n_{s,t}),
\]

where \(W^n_{s,t} = S_2(W^n)_{s,t}\).

The martingale formula yields a very brief proof that the quantized Gaussians are adequately integrable. This first Lemma recasts the well known result mentioned earlier in Equation \((2.6)\).

**Lemma 4.6.** Let \(\mathcal{L}W\) be the law of Brownian motion over \(C^{0,0}(\alpha,T;\mathbb{R}^d)\). Let \(\mathcal{F}\) be a sub-\(\sigma\) algebra of the Borel sigma algebra over \(C^{0,0}(\alpha,T;\mathbb{R}^d)\) that is component-wise conditionally independent. Define
\[ \dot{W} = E[W|\mathcal{F}] \text{.} \] Let \( W \) be the Gaussian rough path of \( \mathcal{L}^W \) and \( \dot{W} \) be the lift of the random variable \( \dot{W} \) to a rough path.

Then, for a constant \( C_1 = C_1(d', p) \) dependent only on \( d' \) and \( p \), we have

\[
\| \dot{W} \|^p_{p\text{-var};[0,T]} \leq C_1 E \left[ \| W \|^p_{p\text{-var};[0,T]} \right].
\]

**Proof.** Firstly, we work with the homogeneous norm (A.1) for \( G^2(\mathbb{R}^{d'}) \) rather than the Carnot Carathéodory norm in order to evaluate the increments explicitly.

By component-wise conditional independence (for the 2nd equality) we have

\[
\| \dot{W} \|^p_{p\text{-var};[0,T]} = \sup_{D=(t_i), t_i \in D} \sum \left( \sum_{j=1}^{d'} \left( E[W_{t_i,t_{i+1}}|\mathcal{F}], e_j \right) + \sum_{j,k=1}^{d'} \int_{t_i}^{t_{i+1}} \left( E[W_{u}, e_j]d\left(W_{u}, e_k\right) \right)^{1/2} \right)^p
\]

\[
= \sup_{D=(t_i), t_i \in D} \sum \left( \sum_{j=1}^{d'} \left( E[W_{t_i,t_{i+1}}|\mathcal{F}], e_j \right) + \sum_{j,k=1}^{d'} E \left[ \int_{t_i}^{t_{i+1}} \left( W_{u}, e_j \right)d(W_u, e_k) \right]^{1/2} \right)^p
\]

\[
\leq d'(2p-2)/p \sup_{D=(t_i), t_i \in D} \sum E \left[ \left( \sum_{j=1}^{d'} \left( W_{t_i,t_{i+1}}, e_j \right) + \sum_{j,k=1}^{d'} \int_{t_i}^{t_{i+1}} \left( W_{u}, e_j \right)d(W_u, e_k) \right)^{1/2} \right]^{p} \cdot |\mathcal{F}|
\]

where we use a finite dimensional norm equivalence for the first inequality. There is a further multiplicative constant that appears from translating this result back to the Carnot Carathéodory norm which is dependent only on \( d' \).

This result does not follow immediately via the same convexity argument used in Equation (2.6) because the Expectation of a Group element may not be a Group element itself.

**Proposition 4.7.** Let \( n, N \in \mathbb{N} \). Let \( \mathcal{L}^W \) be the law of a Brownian motion on \( C^{[0,0)}([0,T];\mathbb{R}^{d'}) \) and let \( W^N \) be the truncated Brownian motion. From Definition 4.2 let \( \tau_{i}(\beta) \) be the greedy sequence of the Brownian rough path \( W \), let \( \tau_{i}(\beta) \) be the greedy sequence of the enhanced truncated Brownian motion \( W^N \) and let \( \tilde{\tau}_{i}(\beta) \) be the greedy sequence of the enhanced quantization \( q_{0,\beta}(W) \) as introduced in Definition 3.10. Let \( \tilde{\beta} = C_1 \beta \) and \( \tilde{\beta} = C_1 \beta \) where \( C_1 \) is the constant introduced in Lemma 4.6.

Let \( N_{\beta,p,[0,T]}(W), \tilde{N}_{\beta,p,[0,T]}(W^N) \) and \( \tilde{N}_{\beta,p,[0,T]}(q_{0,\beta}(W)) \) be the number of elements of each of the respective greedy sequences over the interval \([0,T]\). Then

\[
N_{\beta,p,[0,T]}(q_{0,\beta}(W)) \leq \tilde{N}_{\beta,p,[0,T]}(W^N) \leq N_{\beta,p,[0,T]}(W).
\]

**Proof.** This proof relies on the choice of quantization, and we choose \( q(W^N) \) to be the optimal quantization of the finite dimensional Gaussian random variable \( W^N \) as a measure over the set \( \mathcal{H}^N \) with independent spatial components, see Lemma 3.7. Let \( \mathcal{F} \) be the \( \sigma \)-algebra generated by the partition of the quantization \( \tilde{\mathcal{F}} = \sigma(\tilde{\mathcal{G}}) \) and let \( \mathcal{F} \) be the cylindrical sigma algebra generated by the functionals \( (i^*)^{-1}[\mathcal{H}_N] \). Then we have \( q(W^N) = E[W^N|\mathcal{F}] \) and \( W^N = E[W|\mathcal{F}] \).
By Lemma 4.6, we therefore have that for any subinterval \([s, t]\)
\[
\|q_n(W)|^p_{p-\text{var};[s,t]} \leq C_1E\left[\|W^N|^{p}_{p-\text{var};[s,t]}\right], \quad \|W^N|^{p}_{p-\text{var};[s,t]} \leq C_1E\left[\|W|^{p}_{p-\text{var};[s,t]}\right].
\]
In particular, for the intervals \([0, \tau_1(\beta)]\) and \([0, \tau(\beta)]\) we have
\[
\|q_n(W)|^p_{p-\text{var};[0,\tau_1(\beta)]} \leq C_1E\left[\|W^N|^{p}_{p-\text{var};[0,\tau_1(\beta)]}\right] = C_1\beta,
\]
\[
\|W^N|^{p}_{p-\text{var};[0,\tau_1(\beta)]} \leq C_1E\left[\|W|^{p}_{p-\text{var};[0,\tau_1(\beta)]}\right] = C_1\beta.
\]
However, by definition we also have \(\|q_n(W)|^p_{p-\text{var};[0,\tau_1(\beta)]} = \tilde{\beta}\) and \(\|W^N|^{p}_{p-\text{var};[0,\tau_1(\beta)]} = \beta\), so we conclude that \(0 < \tilde{\tau}_1(\tilde{\beta}) \leq \tau_1(\beta) \leq \tau_1(\beta).

Next, arguing via induction we suppose that \(\tilde{\tau}_k(\tilde{\beta}) \leq \tau_k(\beta)\). Then
\[
\|W^N|^{p}_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)\setminus \tau_k(\beta)]} \leq C_1E\left[\|W|^{p}_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)\setminus \tau_k(\beta)]}\right] = C_1\beta,
\]
\[
\|q_n(W)|^p_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)\setminus \tau_k(\beta)]} \leq C_1E\left[\|W^N|^{p}_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)\setminus \tau_k(\beta)]}\right] = C_1\beta.
\]
However, \(\|W^N|^{p}_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)]} = \beta\) and
\[
\|q_n(W)|^p_{p-\text{var};[\tilde{\tau}_k(\tilde{\beta}),\tau_{k+1}(\beta)]} = \tilde{\beta}\) so we conclude \(\tilde{\tau}_{k+1}(\tilde{\beta}) \leq \tau_{k+1}(\beta)\).

Next, suppose that \(N_{\beta,p,[0,T]}(W) = k\) for some \(k \in \mathbb{N}\). Then \(T < \tau_{k+1}(\beta) \geq \tau_{k+1}(\tilde{\beta}) \geq \tau_{k+1}(\beta)\).

Thus \(k\) is an upper bound for \(N_{\beta,p,[0,T]}(W^N)\) and \(\tilde{N}_{\beta,p,[0,T]}(q_n(W))\). \(\square\)

Finally, we establish the uniform integrability of the quantizations.

**Proposition 4.8.** Let \(\mathcal{L}^W\) be the law of an enhanced Brownian motion and let \(\mathcal{L}^W \circ q_n^{-1}\) be the law of the quantized Brownian motion.

Then the Moment Generating function of the Accumulated \(p\)-variation of \(q_n(W)\) is well defined and bounded by the Moment Generating function of the Accumulated \(p\)-variation of \(W\).

**Proof.** From [CLL13, Proposition 4.11], we have
\[
\beta N_{\beta,p,[0,T]}(\omega) \leq M_\beta(\omega) \leq \beta \left(2N_{\beta,p,[0,T]}(\omega) + 1\right),
\]
for any control \(\omega\) so the existence of a Moment Generating Function for \(N\) is equivalent to the existence of a Moment Generating Function for \(M\).

Therefore, by Proposition 4.4 we have that \(\forall \theta, \beta > 0\) that
\[
\mathbb{E}\left[\exp\left(\theta N_{\beta,p,[0,T]}(W)\right)\right] < \infty.
\]
Applying Proposition 4.7 we get that
\[
\exp\left(\theta \tilde{N}_{\beta,p,[0,T]}(q_n(W))\right) \leq \exp\left(\theta N_{\beta,p,[0,T]}(W)\right).
\]
We take expectations to conclude. \(\square\)
4.2 Existence, Uniqueness and the Occupation Measure Path

In this Subsection, we overview some of the key details of [CL15] to establish the link between particle systems and McKean Vlasov Equations and the existence and uniqueness of the solution law of McKean Vlasov Equations.

The space of measures $\mu$ over the metric space $(E,d,E)$ is not a Banach space. However, a measure can be thought of as a functional over the space of Lipschitz functions on $E$.

**Definition 4.9.** For $\mu \in \mathcal{P}_2(E)$, we define $\gamma^\mu \in \text{Lip}_1^1(E)^*$ to be the linear functional such that for any $f \in \text{Lip}_1^1(E)$,

$$\gamma^\mu[f] = \int_E f d\mu.$$ 

Similarly, for a collection of measures $(\mu_t)_{t \in [0,T]}$, we define the Occupation measure path $\gamma^\mu_t$.

First introduced in [CL15], it is further proved that for the law of an SDE $\mu_t$, the Occupation measure path $\gamma^\mu$ is bounded variation in the Banach norm and so has a canonical Young Signature. The existence and uniqueness of a solution to equation (4.4) comes immediately from [FV10b, Chapter 12].

**Assumption 4.10.** Let $\varsigma > \frac{1}{3}, \gamma > 1$ and $M = \lfloor \frac{1}{\gamma} \rfloor$. Let

$$\sigma \in \text{Lip}^\varsigma \left( \mathbb{R}^d, L(\mathbb{R}^d, \mathbb{R}^d) \right) \quad \text{and} \quad b \in \text{Lip}^\gamma \left( \mathbb{R}^d, L(\text{Lip}_1^1(G^M(\mathbb{R}^d))^*, \mathbb{R}^d) \right).$$

**Definition 4.11.** Suppose $b$ and $\sigma$ satisfy Assumption 4.10. Let $\mu \in \mathcal{P}_1(G\Omega_\alpha(\mathbb{R}^d), \xi \in \mathbb{R}^d$ and $W = G\Omega_\alpha(\mathbb{R}^d)$.

Then the operator $\Theta_{b,\sigma} : \mathcal{P}_2 \left( G\Omega_\alpha(\mathbb{R}^d) \right) \times \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d) \to G\Omega_\alpha(\mathbb{R}^d)$ maps $(\mu, \xi, W)$ to the rough path that is the solution of the Rough Differential Equation

$$dX_t = b(X_t)d\gamma^\mu_t + \sigma(X_t)dW_t, \quad X_0 = \xi,$$

$$(\mu, \xi, W) \mapsto \Theta(\mu, \xi, W) = X. \quad (4.4)$$

4.2.1 Particle Approximations and Finite Support Laws

Firstly, we address the existence and uniqueness of a solution to the system of interacting particles that the McKean Vlasov equation models. Let $\mathcal{C}$ be a codebook for a quantization of the law of the Brownian motion $L^W$ as a measure over the Banach space $C^\alpha_\mathcal{H}([0,T] ; \mathbb{R}^d)$ containing $n$ elements $h^i$. Each $h^i$ is a RKHS path. Associated to each path is a component of the probability vector $p = (p_j)$ such that $p_j = L^W(s_j)$ where $s_j \in \mathcal{S}$ is the element of the partition associated to $h^i$.

By the nature of $\mathcal{H}$, we know that each path $h^i$ is a 1-variation path. Hence one can construct a canonical lift from $h^i$ to a rough path $h^i$ using Young Integration over the interval $[0,T]$. Thus for $t \in [0,T]$ where $M$ is the largest integer such that $M\alpha < 1$ we have

$$h^i_t = S_M(h^i)_{0,t}.$$ 

We know that $n$ is a finite integer, so we can denote the single path $h := \times_{j=1}^n h^i$ which takes values in $\mathbb{R}^{d \times n}$. This path is still 1-variation with respect to the canonical norm on $\mathbb{R}^{d \times n}$. Therefore, we can similarly construct

$$h_t = S_M \left( \times_{j=1}^n h^i \right)_{0,t}.$$ 

22
For clarity, we emphasise that this is a rough path taking values in $T^M(\mathbb{R}^d \times n)$ and it is not the same as $\oplus_{j=1}^{n} h^j = \oplus_{j=1}^{n} S_M(h^j)$ which takes values in $T^M(\mathbb{R}^d) \oplus n$.

When working on the tensor algebra $T^M(V)$. We refer to the Alphabet $A$, which in the case $V = \mathbb{R}^d$, is the letters $\{1, ..., d\}$. However, when working on the tensor algebra $T^M(\mathbb{R}^{d \times n})$, we have the Alphabet $A$ containing all the pairs $\{(i,j); i \in \{1, ..., d\}, j \in \{1, ..., n\}\}$. We will also refer to $A^j$, the Subalphabet containing all pairs $\{(i,j); i \in \{1, ..., d\}\}$. Key to the following result is that the Subalphabets $A^j$ form a partition of the Alphabet $A$.

**Lemma 4.12.** Let $V$ be a vector space with finite Alphabet $A$ and suppose that $A$ can be partitioned into a finite number of Subalphabets denoted by $A^j$. Define

$$I^M(V) := \{ h \in T^M(V) : \langle h, e_I \rangle = 0, \forall I \text{ a word with letters in } A \text{ s.t. } \exists j \text{ where } I \text{ is a word of } A^j \}. $$

Then $I^M(V)$ is a closed ideal of the Lie Algebra $P^M(V)$.

**Proof.** We verify that for $h_1 \in I^M(V)$ and $h_2 \in P^M(V)$ that $[h_1, h_2] \otimes e_I \in I^M(V)$.

Let $I$ be a word that has the property that $\exists j$ such that $I$ is also a word of $A^j$. We denote

$$\Delta e_I = \sum_{I_1, I_2 = I} e_{I_1} \otimes e_{I_2}$$

using “Sweedler” notation and $I_1, I_2$ as being word concatenation. If $I$ is a word with letters in $A^j$ then any subword of $I$ is also a word with letters in $A^j$.

Therefore, for $h_1 \in I^M(V)$ and $h_2 \in P^M(V)$

$$\langle h_1 \otimes h_2, e_I \rangle = \langle h_1 \otimes h_2, \Delta e_I \rangle = \sum_{I_1, I_2 = I} \langle h_1, e_{I_1} \rangle \cdot \langle h_2, e_{I_2} \rangle = \sum_{I_1, I_2 = I} 0 \cdot \langle h_2, e_{I_2} \rangle = 0.$$

Similarly $\langle h_2 \otimes h_1, e_I \rangle = 0$, so naturally

$$\langle [h_1, h_2] \otimes e_I \rangle = 0.$$

Given an Ideal of a Lie Algebra, one can obtain a normal subgroup of the associated Lie Group by taking exponentials. Thus define

$$K^M(V) := \exp \left( I^M(V) \right),$$

and consider the quotient group $G^M(V)/K^M(V)$. There is a canonical isomorphism that maps this quotient group to $\oplus_j G^M(V^j)$ where $V^j$ is the vector space with Alphabet $A^j$.

In order to study the system of interacting particle equations for (4.1), we consider the following drift and diffusion terms. Before that, we introduce a notational convenience in order to distinguish between elements of $\mathbb{R}^d$ and $\mathbb{R}^{d \times n}$. Recall that for $i \in A$, $e_i$ is the unit vector in the vector space with Alphabet $A$. We denote $Y \in \mathbb{R}^{d \times n}$ and $\langle Y, e_{(\cdot, m)} \rangle \in \mathbb{R}^d$ to be the canonical projection of $Y$ where $m \in \{1, ..., n\}$.
Definition 4.13. Let \( b \) and \( \sigma \) satisfy Assumption 4.10. Let \( p = (p_k)_{k=1,...,n} \in \mathcal{P} \). Let \( B : \mathbb{R}^{d \times n} \to \mathbb{R}^{d \times n} \) and \( \Sigma : \mathbb{R}^{d \times n} \to L(\mathbb{R}^{d \times n}, \mathbb{R}^{d \times n}) \) be defined by

\[
B(X) := \bigoplus_{m=1}^{n} \left( b(\{X, e_{(c,m)}\}) \left( \sum_{k=1}^{n} p_k \delta(\{X, e_{(c,k)}\}) \right) \right),
\]

\[
\Sigma(X) := \text{Diag}_{m=1,...,n} \left( \sigma(\{X, e_{(c,m)}\}) \right).
\]

Let \( \tilde{W}_t^k \in G_{\alpha}(\mathbb{R}^d) \) for each \( k \in \{1,...,n\} \). Let \( \tilde{W} = \bigoplus_{k=1}^{n} \tilde{W}_t^k \) be the rough path taking values in the quotient group \( G^M((\mathbb{R}^d)^{\otimes n})/I^M((\mathbb{R}^d)^{\otimes n}) \). Let \( \mathbb{X}_t \) be the controlled rough that solves the Rough Differential Equation

\[
dX_t = B(X_t)dt + \Sigma(X_t)d\tilde{W}_t, \quad X_0 = \xi^x_{\mathbb{R}^n}, \tag{4.6}
\]

taking values in \( \mathbb{R}^{d \times n} \). By the properties of \( b \) and \( \sigma \) from Definition 4.13, we have that \( B \in \text{Lip}^\gamma((\mathbb{R}^{d \times n})^n) \) and \( \Sigma \in \text{Lip}^\delta((\mathbb{R}^{d \times n}, L(\mathbb{R}^{d \times n}, \mathbb{R}^{d \times n}))) \). Therefore, the existence of a solution to Equation (4.6) is standard.

Next we introduce a product on the space of vector fields from \( U \) into \( T^M(V, U) \) designed to simplify the representation of a controlled rough path.

Definition 4.14. Let \( V \) and \( U \) be vector spaces. Let \( i, j \in \mathbb{N} \). For differentiable Vector fields \( F : U \to L(V^{(i)}, U) \) and \( G : U \to L(V^{(j)}, U) \), we define the operation \( \ast \) such that \( F \ast G : U \to L(V^{(i+j)}, U) \) by

\[
F \ast G(u)[V_1, ..., V_j, V_{j+1}, ..., V_{j+i}] = \left( \lim_{\varepsilon \to 0} \frac{F(u + \varepsilon G(u)[V_1, ..., V_j]) - F(u)}{\varepsilon} \right)[V_{j+1}, ..., V_{j+i}]
\]

\[
= DF(u)(G(u)[V_1, ..., V_j])[V_{j+1}, ..., V_{j+i}] \tag{4.7}
\]

It is a natural observation to make that the controlled rough path \( \mathbb{X} \) that represents the solution to Equation (4.6) is equal to

\[
\mathbb{X}_s = \left( X_s, \Sigma(X_s), \Sigma \ast \Sigma(X_s), ..., \Sigma \ast (M-1)(X_s) \right), \quad s \in [0, T]. \tag{4.8}
\]

Lemma 4.15. Let \( V, U \) be vector spaces with alphabets \( A \) and \( \tilde{A} \). Suppose that \( V = \bigoplus_{j=1}^{n} V^j \) and \( U = \bigoplus_{j=1}^{n} U^j \) so that \( A \) and \( \tilde{A} \) can be partitioned into a collection of \( n \) subalphabets \( A^j \) and \( \tilde{A}^j \) for \( j = 1, ..., n \).

For \( k, l \in \mathbb{N} \), let \( F : U \to L(V^{\otimes k}, U) \) and \( G : U \to L(V^{\otimes l}, U) \), suppose that there exist \( f^j : U^j \to L\left(V^j \otimes k, U^j\right) \) and \( g^j : U^j \to L\left(V^j \otimes l, U^j\right) \) such that we have the representation

\[
F(u) = \text{Diag}_{j=1,...,n} \left( f^j(P_{U^j}[u]) \right), \quad G(u) = \text{Diag}_{j=1,...,n} \left( g^j(P_{U^j}[u]) \right).
\]

Suppose that \( F \) is differentiable. Then \( F \ast G \) has the representation

\[
F \ast G(u) = \text{Diag}_{j=1,...,n} \left( Df(P_{U^j}[u]) \times g(P_{U^j}[u]) \right). \tag{4.9}
\]

Proof. For fixed \( m \in \{1,...,n\} \), let \( u_m \in U^m \) and let \( I \) be a word of the subalphabet \( A^m \) such that \( I = (I_1, I_2) \) where \( |I_1| = k \) and \( |I_2| = l \).

Outside of this scenario, all derivatives will be 0 by construction. \(\Box\)
We know that by Theorem A.11, the controlled rough path $X$ can be lifted to a rough path. Our next result, the main result of this Section and similar to one found in [CL15], ensures the choice of lift does not affect the final solution to our equations.

**Theorem 4.16.** For $j = 1, \ldots, n$, let $W^j \in G^{\alpha}_\Omega(R^d')$ and define $W = \oplus_{j=1}^n W^j$. Let $\tilde{W}$ be the extension of $W$ to $G^{\alpha}_\Omega(R^d \times n)$. Let $B$ and $\Sigma$ be as defined in Definition 4.13 and let $X$ be the unique controlled rough path that solves the Rough Differential Equation (4.6).

Let $X \in G^{\alpha}_\Omega(R^d \times n)$ be the lift of $X$ as constructed in Equation (A.10). Then $X$ is dependent on $W$ but not $\tilde{W}$.

**Proof.** See Appendix B.

---

### 4.2.2 Existence and Uniqueness

For this section, we focus on the approach of [CL15]. Firstly, we introduce some of the notation and operators used in this paper to construct different elements for solving our McKean Vlasov equation. The methods and results of [BCD18] which are further explored in [BCD20] and [BCD19] are not used here.

**Definition 4.17.** Let $b$ and $\sigma$ satisfy Assumption 4.10. Let $L \in \mathcal{P}_1(G^{\alpha}_\Omega(R^d))$ and $\mu \in \mathcal{P}_1(G^{\alpha}_\Omega(R^d))$ be probability measures. Then define the map $\Psi_L : \mathcal{P}_2(G^{\alpha}_\Omega(R^d)) \to \mathcal{P}_2(G^{\alpha}_\Omega(R^d))$ by

$$\Psi_L(\mu) = L \circ \Theta_{b,\sigma}(\mu, \xi, \cdot)^{-1}. \quad (4.10)$$

The fixed point of the operator $\Psi_L$ will be the law of the solution to the McKean Vlasov Equation (4.1) where the law of the driving noise $W$ is given by $L$.

**Assumption 4.18.** Let $\varsigma > \frac{1}{\alpha} > 1$ and $\gamma > 1$. Suppose that

1. The measure $L^W \in \mathcal{P}_2(G^{\alpha}_\Omega(R^d))$ satisfies that for any $\theta \geq 0$

$$\int_{G^{\alpha}_\Omega(R^d)} \exp \left( \theta M_{1,[0,T]}(\omega_X) \right) dL^W(X) < \infty, \quad (4.11)$$

2. The functions $b$ and $\sigma$ satisfy Assumption 4.10

**Theorem 4.19 ([CL15]).** Suppose Assumption 4.18 holds. Then the operator $\Psi_L^W$ is a contraction operator with fixed point equal to the law of the solution to the McKean Vlasov Equation (4.1).

Hence there exists a unique solution to the Rough Differential Equation (4.1).

### 4.3 Propagation of Chaos and Quantization

The final result of [CL15] is to prove continuity of the map from the law of the driving noise to the law of the McKean Vlasov Equation. This is framed within the narrative of “Propagation of Chaos”. We exploit this result to show that the law of the associated particle systems of our quantizations converge to the true law of the McKean Vlasov Equation.
Definition 4.20. Let $K : (0, \infty) \to (0, \infty)$ be a monotone increasing real valued function. Define the collection of measures

\[ P_K \left( G\Omega_{\alpha}(\mathbb{R}^d) \right) := \left\{ \mathcal{L} \in P_1 \left( G\Omega_{\alpha}(\mathbb{R}^d) \right) : \forall \theta \in (0, \infty) \right. \]

\[ \int_{G\Omega_{\alpha}(\mathbb{R}^d)} \exp \left( \theta M_{1, [0, T]}(\omega X) \right) d\mathcal{L}(X) \leq K(\theta) \}

paired with the topology of weak convergence generated by the rough path Hölder norm.

A natural way to think about this collection of measures is the law of all rough paths such that the moment generating function of the Accumulated $1_{\alpha}$-variation is dominated by the function $K$.

Proposition 4.21. Suppose Assumption 4.18 is satisfied. Suppose additionally that there exists a monotone increasing function $K : (0, \infty) \to (0, \infty)$ that dominates Equation (4.11). Define the operator $\Xi : P_K \left( G\Omega_{\alpha}(\mathbb{R}^d) \right) \to P_2 \left( G\Omega_{\alpha}(\mathbb{R}^d) \right)$ by

\[ \Xi \left[ \mathcal{L}^{W} \right] = \mathcal{L}^{X}, \quad (4.12) \]

where $\mathcal{L}^{X}$ is the unique measure that is a fixed point of Equation (4.10) so that $\Psi_{\mathcal{L}^{W}}(\mathcal{L}^{X}) = \mathcal{L}^{X}$.

Then the operator is well defined and for $\mathcal{L}^{W_1}, \mathcal{L}^{W_2} \in P_K \left( G\Omega_{\alpha}(\mathbb{R}^d) \right)$ we have

\[ \mathcal{W}^{(2)}_{\rho_{\alpha-Höl,[0,T]}} \left( \Xi[\mathcal{L}^{W_1}], \Xi[\mathcal{L}^{W_2}] \right) \leq C \mathcal{W}^{(2)}_{\rho_{\alpha-Höl,[0,T]}} \left( \mathcal{L}^{W_1}, \mathcal{L}^{W_2} \right) \]

with a constant $C = C(\alpha, K, T, d, d')$.

Previously, this result was used to show that the empirical measure obtained by sampling paths of a Brownian motion could be used to obtain a particle system that would converge as the number of particles increased to the solution of a McKean Vlasov Equation. In the remarkable work [DFMS18], the authors study the rate of convergence of these empirical measures to the true law in probability.

Proof. Same as proof [CL15, Lemma 4.11].

4.4 Continuity with respect to the Occupation Measure path

In [CL15] Theorem 4.9, the goal was to establish the existence of a contraction operator whose fixed point would be the law of the McKean Vlasov Equation. In fact, computing the specific contraction operator is not simple. Here, we provide a more tangible operator that is (Lipschitz) continuous but not a contraction.

Proposition 4.22. Let $b$ and $\sigma$ satisfy Assumption 4.10 and let $\Theta_{b,\sigma}$ be the operator from Definition 4.11.

Then $\Theta_{b,\sigma}$ is Locally Lipschitz continuous in the measure component, that is $\forall \mu, \nu \in P_1 \left( G\Omega_{\alpha}(\mathbb{R}^d) \right)$ such that

\[ \int_{G\Omega_{\alpha}(\mathbb{R}^d)} \rho_{\alpha-Höl}(X, 1) d\mu(X), \int_{G\Omega_{\alpha}(\mathbb{R}^d)} \rho_{\alpha-Höl}(X, 1) d\nu(X) < C \]

and $\forall \xi \in \mathbb{R}^d$ and $\forall W \in G\Omega_{\alpha}(\mathbb{R}^d)$ such that $\rho_{\alpha-Höl,[0,T]}(W, 1) < C$, $\exists L_C > 0$ such that

\[ \rho_{\alpha-Höl} \left( \Theta_{b,\sigma}(\mu, \xi, W), \Theta_{b,\sigma}(\nu, \xi, W) \right) \leq L_C \mathcal{W}^{(2)}_{\rho_{\alpha-Höl,[0,T]}} \left( \mu, \nu \right) \]

(4.14)
Similarly, by choosing $C$ and $M = \lfloor \frac{1}{\alpha} \rfloor$, Denote the control $\omega(s, t) = \|W\|^p_{\text{p-\var};[s, t]} + \|\gamma^\mu\|_{1-\var;[s, t]} + \|\gamma^\nu\|_{1-\var;[s, t]}$. Then [CL15] Lemma 4.3 gives

$$\rho_{p-\omega};[0, T](\Theta_{b, \sigma}(\mu, \xi, W), \Theta_{b, \sigma}(\nu, \xi, W)) \leq C\rho_{1-\omega};[0, T](\gamma^\mu, \gamma^\nu) \exp(M_{\beta};[0, T](\omega)).$$

Indeed, we also have

$$\|\gamma^\mu_{s, t} - \gamma^\nu_{s, t}\|_{\text{Lip}^2(G^M(\mathbb{R}^d))} \leq \|t - s\|\mathbb{W}^{(2)}_{\rho_{\text{\alpha-\Hol}}};[0, t](\mu, \nu).$$

By assumption, the Wasserstein distance must be finite across the interval $[0, T]$, so we know the control can be dominated by $\omega(s, t) \leq |t - s|$. Thus $\rho_{p-\omega}$ will be equivalent to $\rho_{\alpha-\Hol}$ and we get

$$\rho_{\alpha-\Hol};[0, T](\Theta_{b, \sigma}(\mu, \xi, W), \Theta_{b, \sigma}(\nu, \xi, W)) \leq C\mathbb{W}^{(2)}_{\rho_{\text{\alpha-\Hol}}};[0, t](\mu, \nu) \cdot \exp(M_{\beta};[0, T](\omega)).$$

Next, we note that while the constant $C$ is uniform over the choice of $\mu$ and $\nu$, the control $\omega$ is dependent on them and so the Accumulated $\beta$-local $p$-variation is also dependent on their second moments.

With only Proposition 4.22, one can establish the distance between two paths driven by different occupation measure paths. Next we prove uniform continuity.

**Theorem 4.23.** Let $b$ and $\sigma$ satisfy Assumption 4.10 and let $\Theta_{b, \sigma}$ be the operator defined in Definition 4.11. Then the operator $\Theta_{b, \sigma}$ is jointly continuous over $\mathcal{P}_2\left(G\Omega_\alpha(\mathbb{R}^d)\right) \times \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d)$. In particular,

$$\lim_{(\mu_k, \xi_k, W_k) \to (\mu, \xi, W)} \Theta_{b, \sigma}(\mu_k, \xi_k, W_k) = \lim_{\mu_k \to \mu} \lim_{\xi_k \to \xi} \lim_{W_k \to W} \Theta_{b, \sigma}(\mu_k, \xi_k, W_k) = \Theta_{b, \sigma}(\mu, \xi, W).$$

**Proof.** Let $\xi, \chi \in \mathbb{R}^d$ and $p = \frac{1}{\alpha}$. For $W_1, W_2 \in G\Omega_\alpha(\mathbb{R}^d)$ and $\mu, \nu \in \mathcal{P}_2\left(G\Omega_\alpha(\mathbb{R}^d)\right)$, define the control

$$\omega(s, t) = \|W_1\|^p_{\text{p-\var};[s, t]} + \|W_2\|^p_{\text{p-\var};[s, t]} + \|\gamma^\mu\|_{1-\var;[s, t]} + \|\gamma^\nu\|_{1-\var;[s, t]}.$$

We have

$$\rho_{\text{\alpha-\Hol}};[0, T](\Theta_{b, \sigma}(\mu, \xi, W_1), \Theta_{b, \sigma}(\nu, \chi, W_2)) \leq C\left(\|\xi - \chi\| + \rho_{\text{\alpha-\Hol}};[0, T](W_1, W_2) + \rho_{1-\Hol};[0, T](\gamma^\mu, \gamma^\nu)\right) \exp(M_{\beta};[0, T](\omega)).$$

Proposition 4.22 shows continuity in measure pointwise for each Geometric rough path $W$. Therefore, to prove joint continuity via Moore-Osgood we verify the uniform continuity condition. Let $\mu_k, \mu \in \mathcal{P}_2\left(G\Omega_\alpha(\mathbb{R}^d)\right)$ and $\mathbb{W}^{(2)}_{\rho_{\text{\alpha-\Hol}}};[0, T](\mu_k, \mu) \to 0$. Then we also have

$$\lim_{k \to \infty} \|\gamma^\mu_k - \gamma^\mu\|_{1-\var;[0, T]} = 0.$$

Hence there must exist an $C' \in \mathbb{N}$ such that

$$\sup_{k > C'} \|\gamma^\mu_k\|_{1-\var;[0, T]} \leq \|\gamma^\mu\|_{1-\var;[0, T]} + 1.$$

Similarly, by choosing $C'$ large enough

$$\sup_{k > C'} \|W_k\|^p_{\text{p-\var};[0, T]} \leq \left(\|W\|^p_{\text{p-\var};[0, T]} + 1\right)^p.$$
Thus
\[
\sup_{k>C'} \rho_{\alpha-\text{Höyl};[0,T]} \left( \Theta_{b,\sigma}(\mu_k, \xi_k, W_1), \Theta_{b,\sigma}(\mu_k, \xi_k, W_2) \right) \\
\leq C \rho_{\alpha-\text{Höyl};[0,T]}(W_1, W_2) \\
\cdot \exp \left( \mathbf{M}_{\beta,p}(W_1) + \mathbf{M}_{\beta,p}(W_2) \right) \exp \left( \|\gamma_{\mu}\|_{1-\text{var};[0,T]} + 1 \right),
\]
\[
\sup_{k>C'} \rho_{\alpha-\text{Höyl};[0,T]} \left( \Theta_{b,\sigma}(\mu_k, \xi, W_k), \Theta_{b,\sigma}(\mu_k, \chi, W_k) \right) \\
\leq C |\xi - \chi| \cdot \exp \left( (\|W\|_{p-\text{var};[0,T]} + 1)^p \right) \exp \left( \|\gamma_{\mu}\|_{1-\text{var};[0,T]} + 1 \right),
\]
\[
\sup_{k>C'} \rho_{\alpha-\text{Höyl};[0,T]} \left( \Theta_{b,\sigma}(\mu, \xi_k, W_k), \Theta_{b,\sigma}(\nu, \xi_k, W_k) \right) \\
\leq C \rho_{1-\text{Höyl};[0,T]}(\gamma_{\mu}, \gamma_{\nu}) \\
\cdot \exp \left( (\|W\|_{p-\text{var};[0,T]} + 1)^p \right) \exp \left( \|\gamma_{\mu}\|_{1-\text{var};[0,T]} + \|\gamma_{\nu}\|_{1-\text{var};[0,T]} \right)
\]
which implies uniform continuity.

\boxed{}

5 Support Theorem

Finally, we state and prove representations of the support of McKean Vlasov Equations in terms of the particle systems associated to the quantizations that we constructed in Section 3. We introduce a collection of sets of paths that to the best of our knowledge have not previously been described in another work. These sets are all subsets of $C^\alpha,0([0,T]; \mathbb{R}^d)$ and are defined solely with respect to the RKHS $\mathcal{H}$, the Hölder norm $\|\cdot\|_\alpha$ and the coefficients of the Rough Differential Equation (4.1).

In order to provide a clear exposition of the construction of the support, we briefly summarise the upcoming subsections: from the previous Section we have obtained a sequence of quantizations $q_n$ for the the law of the enhanced Brownian motion with codebooks $C_n$.

- For each quantization, we solve the system of interacting ODEs in Section 5.1.1 (see Equation (5.2)) by replacing the path of Brownian motion by the associated codebook path and replacing the law of the Brownian motion by the quantization.

- By associating to each of these ODEs the probability weight associated to the codebook element driving the equation, we obtain a finite support measure in Section 5.1.2 (see Equation (5.3)). We call this the quantization of the McKean Vlasov Equation. This sequence of finite support measures converges to the law of the McKean Vlasov Equation.

- In Section 5.2.1 for fixed $n$, we replace the law of the McKean Vlasov Equation inside the canonical skeleton process by the quantization of the McKean Vlasov Equation (see Definition 5.6). These paths will not generally be contained in the support of the McKean Vlasov Equation. However, a ball of large enough radius will have positive measure (see Lemma 5.7).

- In Section 5.2.2, we show that for $n$ chosen large enough, an $\varepsilon$ ball around this collection of paths will be a closed set of measure 1. By taking an intersection of these sets, we show the set of limit points has measure 1 (see Theorem 5.8).

28
Finally in Section 5.3 we extend our work to the case where the McKean Vlasov Equation has an initial law (see Theorem 5.13).

5.1 The Skeleton Process for a McKean Vlasov Equations

The law of a McKean Vlasov equation is deterministic; it is not dependent on the choice of driving noise. The Occupation Measure path is of bounded variation and does not interact with the noise. Thus when the Occupation Measure path is known, McKean Vlasov Equations can be thought of as Classical Rough Differential Equations with a drift term. Thus, we can define a Skeleton process in the following classical sense:

Definition 5.1. Let \( L \) be the law of an enhanced Brownian motion. Let \( b \) and \( \sigma \) satisfy Assumption 4.10. Let \( \xi \in \mathbb{R}^d \). Let \( L_X \) be the unique fixed point of the operator \( \Psi_L \). Then we define the True Skeleton Operator \( \Phi' : H \times \mathbb{R}^d \to G_{\Omega}(\mathbb{R}^d) \) to be the operator that maps the element of the RKHS to the solution of the ODE

\[
\begin{align*}
\frac{d\Phi'(h,\xi)}{dt} &= b(\Phi'(h,\xi))d\gamma_{L_X} + \sigma(\Phi'(h,\xi))dh, \\
\Phi(h,\xi)_0 &= \xi.
\end{align*}
\]

(5.1)

It is important to emphasise that the True Skeleton Operator (5.1) is dependent on the measure \( L_X \) and as such it cannot be solved without knowing the law exogenously. The main contribution of this Section is how one navigates around this issue.

5.1.1 Interacting Particle system derived from Quantization

We introduce a system of interacting Ordinary Differential Equations that model the dynamics of the McKean Vlasov Equation.

Definition 5.2. Let \( \xi \in \mathbb{R}^d \). Let \( L \in \mathcal{P}_c(G\Omega(\mathbb{R}^d)) \) be a finitely supported measure over the space of Geometric rough paths with the form \( L = \sum_{j=1}^n p_j \delta_{W_j} \) where \( (p_j)_{j=1,...,n} \) is a Probability vector. For a codebook \( C := \{ W_j : j = 1, ..., n \} \), let \( W := \oplus_{j=1}^n W_j \) and let \( \tilde{W} \) be the extension of \( W \) to \( G^M(\mathbb{R}^d \times \mathbb{N}) \) where \( M \) is the largest integer such that \( M \alpha < 1 \). Let \( b \) and \( \sigma \) satisfy Assumption 4.10. Let \( B \) and \( \Sigma \) be as in Definition 4.13.

Then we define the \( L \)-Interacting Particle System to be the solution to the Rough Differential Equation

\[
\begin{align*}
\frac{d\Phi(L)}{dt} &= B(\Phi(L))dt + \Sigma(\Phi(L))d\tilde{W}_t, \\
\Phi(0) &= \oplus_{j=1}^n \xi \in \mathbb{R}^d \times \mathbb{N}
\end{align*}
\]

(5.2)

taking values in \( G\Omega_{\alpha}(\mathbb{R}^d \times \mathbb{N}) \).

An important detail about this object is that this is a finite dimensional system of Rough Differential Equations. This system of interacting equations can be solved without having to consider any measures.

The existence and uniqueness of the ODE (5.2) is standard. In particular, by Theorem 4.16 the solution to Equation (5.2) is independent of the choice of \( W \) and only on \( W \).

5.1.2 Quantization of the McKean Vlasov

We use the interacting particle system (5.2) to obtain a law that approximates the law of the McKean Vlasov Equation (4.1).
Definition 5.3. Let \( L \in \mathcal{P}_c(G\Omega_{\alpha}(\mathbb{R}^d)) \) be a finite support measure over the space of geometric rough paths with the form 
\[
L = \sum_{m=1}^{n} p_m \delta_{\pi(m)}
\]
where \((p_m)_{m=1}^{n}\) is a Probability vector. Let \(b\) and \(\sigma\) satisfy Assumption 4.10.

Let \(\Phi(L)\) be the solution to Equation (5.2). Let \(\pi^{(m)} : G\Omega_{\alpha}(\mathbb{R}^{d\times n}) \to G\Omega_{\alpha}(\mathbb{R}^d)\) be the quotient operator obtained by extending the projection \((\cdot, e_{\cdot(m)})\). Then we define the Law of the \(L\)-Interacting Particle System to be the finite measure over \(G\Omega_{\alpha}(\mathbb{R}^d)\)
\[
\mathcal{L}^{\Phi(L)} := \sum_{m=1}^{n} p_m \delta_{\pi^{(m)}(\Phi(L))}.
\] (5.3)

Substituting a quantization of the Brownian motion into an Interacting Particle System and taking its law, we obtain a quantization for the McKean Vlasov Equation.

Proposition 5.4. Let \(L^W\) be the law of enhanced Brownian motion. Let \(L^W \circ q^{-1}_n\) be the sequence of quantizations of the enhanced Brownian motion from Definition 3.10.

Let \(\mathcal{L}^{\Phi(L^W \circ q^{-1}_n)}\) be the sequence of quantizations for the McKean Vlasov obtained from the sequence of finite support measures \(L^W \circ q^{-1}_n\). Then \(\mathbb{E} \left[ L^W \circ q^{-1}_n \right] = \mathcal{L}^{\Phi(L^W \circ q^{-1}_n)}\) so that
\[
\mathcal{W}^{(2)}_{\rho_{\alpha} - \text{Hölder}, [0, T]} \left( \mathcal{L}^{\Phi(L^W \circ q^{-1}_n)}, \mathcal{L}^X \right) \lesssim \left( \log(n) \right)^{\alpha^{-1}/2}.
\]

Proof. We have \(\mathbb{E} \left[ L^W \circ q^{-1}_n \right] = \mathcal{L}^{\Phi(L^W \circ q^{-1}_n)}\) and \(\mathbb{E} \left[ L^W \right] = \mathcal{L}^X\). By Proposition 4.21 we have
\[
\mathcal{W}^{(2)}_{\rho_{\alpha} - \text{Hölder}, [0, T]} \left( \mathbb{E} \left[ L^W \circ q^{-1}_n \right], \mathbb{E} \left[ L^W \right] \right) \lesssim \mathcal{W}^{(2)}_{\rho_{\alpha} - \text{Hölder}, [0, T]} \left( \mathcal{L}^W \circ q^{-1}_n, \mathcal{L}^W \right).
\]

Apply Theorem 3.12 for the rate of convergence. \(\Box\)

5.2 The Support of the McKean Vlasov Equation

The following result immediately holds from the methods laid out in [FV10b, Chapter 19].

Theorem 5.5. Let \(L^W\) be the law of an enhanced Brownian motion. Let \(\xi \in \mathbb{R}^d\). Let \(b\) and \(\sigma\) satisfy Assumption 4.10. Let \(L^X\) be the law of the McKean Vlasov Equation (4.1). Then the support of \(L^X\) can be characterised with respect to the rough path Hölder metric by
\[
\text{supp}(L^X) = \left\{ \Phi'(h, \xi) : h \in \mathcal{H} \right\}_{\rho_{\alpha} - \text{Hölder}, [0, T]}^{\text{Path}}
\] (5.4)

where \(\Phi'\) is the True Skeleton operator from Definition 5.1.

This is not a meaningful result as the True Skeleton Operator includes a priori knowledge of the law of the McKean Vlasov Equation. This measure can be proved to exist, but constructing it is another matter. We overcome this issue via functional quantization.

5.2.1 Quantized Skeleton of McKean Vlasov

We use the quantized McKean Vlasov to construct a Skeleton process that approximates the True Skeleton Process.
**Definition 5.6.** Let \( L^W \) be the law of an enhanced Brownian Motion. Let \( q_n \) be the sequence of quantizations of \( L^W \) constructed in Definition 3.10. Let \( \xi \in \mathbb{R}^d \) and let \( h \in \mathcal{H} \) and denote \( h = S_2[h] \). Let \( b \) and \( \sigma \) satisfy Assumption 4.10.

Fix \( \varepsilon > 0 \) and choose \( n \in \mathbb{N} \) such that
\[
\rho_{\alpha - \text{Hölder}[0,T]}(\Theta_{b,\sigma}(L^W \circ q^{-1}_n), \xi, h) \leq \varepsilon,
\]
and we define the sets \( A_\varepsilon(h) \) as
\[
A_\varepsilon(h) := \left\{ Y \in G_{\Omega_\alpha}(\mathbb{R}^d) : \rho_{\alpha - \text{Hölder}[0,T]}(Y, \Theta_{b,\sigma}(L^W \circ q^{-1}_n), \xi, h) < \varepsilon \right\}.
\]

We emphasise that the choice of \( n \) will not be uniform over all choices of \( h \in \mathcal{H} \). Also note that \( \Phi'(h, \xi) = \Theta_{b,\sigma}(L^X, \xi, h) \). The first goal is to show that each of these sets contains an element of the \( \text{supp}(L^W) \), regardless of \( \varepsilon \).

**Lemma 5.7.** Let \( h \in \mathcal{H} \) and \( h = S_2[h] \). Then \( \forall \varepsilon > 0 \), the open sets \( A_\varepsilon(h) \) of Definition 5.6 have positive measure with respect to \( L^X \),
\[
L^W[A_\varepsilon(h)] > 0.
\]

**Proof.** The condition for \( A_\varepsilon(h) \) in Equation (5.5) is the key. It ensures that for any choice of \( \varepsilon > 0 \), we have \( \Phi'(h, \xi) \in A_\varepsilon(h) \). By Theorem 5.5, we have that any open set \( B \subseteq G_{\Omega_\alpha}(\mathbb{R}^d) \) containing a path \( \Phi'(h, \xi) \) and for any choice of \( h \in \mathcal{H} \), we have
\[
L^X[B] > 0.
\]

\( \square \)

### 5.2.2 The Support Theorem without knowledge of the solution law

We now formulate our statement of the support theorem of McKean Vlasov Equations:

**Theorem 5.8.** Let \( L^W \) be the law of an enhanced Brownian motion. Let \( q_n \) be the sequence of quantizations obtained in Definition 3.10. Let \( L^\Phi(L^W \circ q^{-1}_n) \) be the law of the Interacting Particle System driven by the quantisation constructed in Definition 5.3. Let \( \xi \in \mathbb{R}^d \). Suppose that \( b \) and \( \sigma \) satisfy Assumption 4.10. Then the law of the solution to the McKean Vlasov Equation (4.1) satisfies
\[
\text{supp}(L^X) = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} \left\{ \Theta_{b,\sigma}(L^\Phi(L^W \circ q^{-1}_n), \xi, h) : h \in \mathcal{H}, h = S_2(h) \right\} \rho_{\alpha - \text{Hölder}[0,T]}.
\]

We emphasise that this expression of the support is only dependent on:

- The RKHS of Brownian motion \( \mathcal{H} \) and the initial condition \( \xi \in \mathbb{R}^d \)
- The coefficients \( b \) and \( \sigma \)
- The sequence of Systems of Interacting Particles \( \Phi(L^W \circ q^{-1}_n) \) which is in turn dependent on
  - The coefficients \( b \) and \( \sigma \)
  - The sequence of quantizations \( q_n \) which are only dependent on \( \mathcal{H} \) and \( \| \cdot \|_\alpha \).
We have not solved the law of the McKean Vlasov Equation or the Occupation measure path at any point of this approach.

**Proof.** For the simplicity of the proof, we rely on Theorem 5.5 for an expression of \( \text{supp}(L^X) \). By Proposition 5.4, we have that the law of the Interacting Particle System converges to the law of the McKean Vlasov Equation as \( n \to \infty \). Fix \( h \in \mathcal{H} \) and \( m \in \mathbb{N} \). Then \( \forall l \geq m \)

\[
\Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n_l}^{-1})}, \xi, h) \in \bigcup_{n \geq m} \left\{ \Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n}^{-1})}, \xi, h) : h \in \mathcal{H}, h = S_2(h) \right\}^{\rho_{\alpha - \text{Höl}}[0,T]}.
\]

Since this is closed, we have that the limit of these paths is also contained so

\[
\Phi'(h, \xi) = \Theta_{b,\sigma}(L^X, \xi, h) \in \bigcup_{n \geq m} \left\{ \Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n}^{-1})}, \xi, h) : h \in \mathcal{H}, h = S_2(h) \right\}^{\rho_{\alpha - \text{Höl}}[0,T]}.
\] (5.8)

Finally, Equation (5.8) holds for any choice of \( m \in \mathbb{N} \), so it must be contained in the intersection over all \( m \). This was true for any choice of \( h \in \mathcal{H} \), so it is also true for all \( h \in \mathcal{H} \). Thus

\[
\left\{ \Phi'(h, \xi) : h \in \mathcal{H} \right\} \subset \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} \left\{ \Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n}^{-1})}, \xi, h) : h \in \mathcal{H}, h = S_2(h) \right\}^{\rho_{\alpha - \text{Höl}}[0,T]}.
\]

Finally, as the right hand side is closed, we can take a closure on the left hand side to achieve the first implication.

Now we show the reverse implication. Suppose \( Y \in G\Omega_\alpha(\mathbb{R}^d) \) such that

\[
Y \in \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} \left\{ \Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n}^{-1})}, \xi, h) : h \in \mathcal{H} \right\}^{\rho_{\alpha - \text{Höl}}[0,T]}.
\]

Then there must exist a subsequence \( n_k \) and a sequence of \( h_k \in \mathcal{H} \) such that

\[
\lim_{k \to \infty} \rho_{\alpha - \text{Höl}}[0,T] \left( \Theta_{b,\sigma}(L^{\Phi(L^W \circ q_{n_k}^{-1})}, \xi, h_k), Y \right) = 0.
\]

Further, we know the sequence satisfies \( \lim_{k \to \infty} n_k = \infty \), since \( Y \) is in the intersect over all \( m \in \mathbb{N} \). Thus the weak limit of \( L^{\Phi(L^W \circ q_{n}^{-1})} \) must just be \( L^X \) as \( k \to \infty \).

By Theorem 4.23, we have Joint Continuity of \( \Theta_{b,\sigma} \). Therefore, taking the limit in the measure variable first, we get

\[
\lim_{k \to \infty} \rho_{\alpha - \text{Höl}}[0,T] \left( \Theta_{b,\sigma}(L^X, \xi, h_k), Y \right) = \lim_{k \to \infty} \rho_{\alpha - \text{Höl}}[0,T] \left( \Phi'(h_k, \xi), Y \right) = 0,
\]

which just means that \( Y \in \{ \Phi'(h, \xi) : h \in \mathcal{H} \}^{\rho_{\alpha - \text{Höl}}[0,T]} \). \( \square \)

### 5.3 Random Initial Conditions

An apparent limitation of the previous Section is that we restrict ourselves to McKean Vlasov Equations with constant initial conditions. However, there is an easy extension to the case where the initial condition is random.

We introduce a Theorem first proved in [CFN97] that allows for the consideration of random initial conditions.
Theorem 5.9 (CIFN97). Let \( F : \Omega \times \mathbb{R}^d \rightarrow E \) be a random variable taking values in a Banach space \( E \) such that \( x \mapsto F(\omega, x) \) is continuous for each \( \omega \). Suppose that \( G : H \times \mathbb{R}^d \rightarrow E \) is a uniform skeleton of \( F \). Suppose that \( \zeta \) is an \( d \)-dimensional random variable with skeleton \( \phi \). Then \( \tilde{G}(h) := G(h, \phi(h)) \) is a skeleton of \( \tilde{F}(\omega) := F(\omega, \zeta(\omega)) \).

We now turn to the McKean Vlasov Equation

\[
dX_t = \sigma(X_t)dW_t + b(X_t)d\gamma^{\mathcal{L}^n}, \quad X_0 \sim \mathcal{P}_r(\mathbb{R}^d)
\]

where \( r > 1 \).

Following in the footsteps of Definition 3.10, we construct a quantization for the law \( \xi \times \mathcal{L}^W \) over \( \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d) \).

Definition 5.10. Let \( r > 1 \). Let \( \mathcal{L}^W \) be the law of a Brownian motion over \( C^{\alpha, 0}([0, T]; \mathbb{R}^d) \). Let \( \xi \in \mathcal{P}_r(\mathbb{R}^d) \). Let \( m, n \in \mathbb{N} \).

1. By Theorem 2.10, there exists a codebook \( \mathcal{C}_m^{(1)} \subset \mathbb{R}^d \) that is an \( m \)-stationary set with Voronoi partition \( \mathcal{S}_m^{(1)} \). Let \( \mathcal{C}_n^{(2)} \) be the \( n \) element codebook constructed in Definition 3.10 with partition \( \mathcal{S}_n^{(2)} \).

2. Let \( \mathcal{C}_{m,n} := \mathcal{C}_m^{(1)} \times \mathcal{C}_n^{(2)} \) be a sequence of codebooks over \( \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d) \) and let \( \mathcal{S}_{m,n} := \mathcal{S}_m^{(1)} \times \mathcal{S}_n^{(2)} \) be a partition of \( \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d) \). Let \( q_{m,n} \) be the Quantization with codebook \( \mathcal{C}_{m,n} \) and partition \( \mathcal{S}_{m,n} \). Then \( |\mathcal{C}_{m,n}| = m \cdot n \).

3. By combining Equation (2.7) and Theorem 3.12, the rate of convergence is

\[
\frac{1}{m^{1/d}} + \left( \log(n) \right)^{\alpha - 1/2} \approx \left( \int_{\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d)} d_\mathcal{H}((x, Y), q_{m,n}(x, Y)) \right)^{1/r} \mathcal{W}(n^{(1/2 - \alpha)d}(1/2 - \alpha)^d).
\]

4. By choosing \( m \approx [\log(n)]^{(1/2 - \alpha)d} \) and rescaling, we obtain the sequence of quantizations

\[
\left( \int_{\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d)} d_\mathcal{H}((x, Y), q_{m}(x, Y)) \right)^{1/r} \approx \mathcal{W}(n^{(1/2 - \alpha)d}(1/2 - \alpha)^d)\]

where, as in Proposition 3.8, \( \mathcal{W} \) is the Lambert \( W \) function.

Next, following Definition 5.3, we define a new interacting particle system.

Definition 5.11. Let \( \mathcal{L} \in \mathcal{P}_r(\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^d)) \) be a finite support measure of the form \( \mathcal{L} = \sum^n_{j=1} p_j \delta_{(x_j, W_j)} \) where \( (p_j)_{j=1,...,n} \) is a probability vector. For codebook \( \mathcal{C} := \{(x_j, W_j) : j = 1, \ldots, n\} \), let \( \mathcal{W} := \oplus^n_{j=1} W_j \) and \( X = \oplus^n_{j=1} x_j \in \mathbb{R}^{d \times n} \). Let \( \mathcal{W} \) be the lift of the path \( \mathcal{W} \) to a rough path. Let \( \mathcal{B} \) and \( \Sigma \) be as in Definition 2.13.

Then we define the \( \mathcal{L} \) Interacting Particle System with random initial condition to be the solution to the Rough Differential Equation

\[
d\Phi(\mathcal{L})_t = \mathcal{B}(\Phi(\mathcal{L})_t)dt + \Sigma(\Phi(\mathcal{L})_t)d\mathcal{W}_t, \quad \Phi(\mathcal{L})_0 = X.
\]
We also define the law of the \( \mathcal{L} \) Interacting Particle system in \( \mathcal{P}_{\mathcal{C}} \left( G \Omega_{\alpha} (\mathbb{R}^d) \right) \) to be

\[
\mathcal{L}^\Phi (\mathcal{L}) := \sum_{m=1}^{n} p_m \delta_{\frac{\left( \cdot, x \right)}{m}} \left( \phi (\mathcal{L}) \right).
\]

As with Theorem 4.16, the paths of this law are dependent only on \( \mathbf{W} \) and not of the lift of \( \tilde{\mathbf{W}} \). In this definition we do not limit ourselves to the case where many of the \( x_j \) values are repeated.

We use the quantization of the measure \( \xi \times \mathcal{L}^\mathbf{W} \) constructed in Definition 5.10 to solve the law of an Interacting Particle system that approximates the true law of the McKean Vlasov Equation

**Proposition 5.12.** Let \( \mathcal{L}^\mathbf{W} \) be the law of the enhanced Brownian motion. Let \( \left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n \) be the sequence of quantizations of the enhanced Brownian motion from Definition 5.10.

Let \( \mathcal{L}^\Phi (\left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n) \) be the sequence of quantizations for the McKean Vlasov obtained from the sequence of finite support measures \( \left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n \).

Then \( \mathbb{E} \left[ \left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n \right] = \mathcal{L}^\Phi (\left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n) \) so that

\[
\mathcal{W}^{(1)}_{\beta_{\alpha} - \text{Höf}} \left( \mathcal{L}^\Phi (\left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n), \xi \times \mathcal{L}^{\mathbf{X}} \right) \leq \log \left( \frac{n}{(1/2-\alpha)d(1/2-\alpha)d} \right) - \log \left( \mathcal{W} \left( \frac{n}{(1/2-\alpha)d} \right) \right) \left( (\alpha-1/2)d \right). \]

**Proof.** Same method as Proposition 5.4 with Equation (5.11).

### 5.3.1 Statement for the Support

Using classical tools, we combine the results of Theorem 5.5 with [CFN97] for this next Theorem:

**Theorem 5.13.** Let \( r > 1 \). Let \( \xi \in \mathcal{P}_r (\mathbb{R}^d) \). Let \( \mathcal{L}^\mathbf{W} \) be the law of an enhanced Brownian motion. Let \( q_n \) be the sequence of quantizations obtained in Definition 5.10. Let \( \mathcal{L}^\Phi (\left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n) \) be the law of the Interacting Particle System driven by the quantization constructed in Definition 5.11. Suppose that \( b \) and \( \sigma \) satisfy Assumption 4.10. Then the law of the solution to the McKean Vlasov Equation (4.1) satisfies

\[
\text{supp}(\mathcal{L}^\mathbf{X}) = \cap_{m=1}^{\infty} \cup_{n \geq m} \left\{ \Theta_{b,\sigma}(\mathcal{L}^\Phi (\left( \xi \times \mathcal{L}^\mathbf{W} \right) \circ q^{-1}_n), x, h) : h \in \mathcal{H}, x \in \text{supp}(\xi) \right\} \right]_{\beta_{\alpha} - \text{Höf} \in [0, T]}.
\]

**Proof.** See proof of Theorem 5.8 with Proposition 5.12 and Theorem 4.23.

### Acknowledgements

In no particular order, the authors thank Professor Sandy Davie, Professor François Delarue, Professor Peter Friz and Dr Mario Maurelli for the helpful discussions. We also thank the participants of the December 2018 10th Oxford-Berlin Young Researchers Meeting on Applied Stochastic Analysis for their feedback on the occasion of the first presentation of this work.
A Rough Path Primer

A.1 Algebraic Material

Let \( A \) be a finite alphabet, let \( V \) be the associated vector space and denote \( T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n} \) be the vector space of free monoids generated by \( A \) with the shuffle product \( \shuffle \). Let \( \Delta : T(V) \to T(V) \otimes T(V) \) be the deconcatenation coproduct. Thus \( (T(V), \shuffle, \Delta) \) is a commutative unital Hopf algebra with an antipode and canonical grading.

The characters (also known as Group-like elements) of \( T(V) \) to be the elements \( g \in G(V) \) such that \( \forall u, v \in T(V) \)

\[
\langle g, u \shuffle v \rangle = \langle g, u \rangle \langle g, v \rangle.
\]

\( G(V) \) forms a Lie group with Lie Algebra \( P(V) \). The diffeomorphic exponential map \( \exp : P(V) \to G(V) \) and its inverse the logarithm map \( \log : G(V) \to P(V) \) defined for \( g \in G(V) \) and \( h \in P(V) \) by

\[
\exp(h) = \sum_{i=0}^{\infty} \frac{h^{\otimes i}}{i!}, \quad \log(g) = \sum_{i=1}^{\infty} (-1)^{i-1} \frac{(g - 1)^{\otimes i}}{i}.
\]

Finally, we define \( T^M(V) \) to be the quotient space obtained from \( T(V) \) by quotienting against the ideal \( \oplus_{n=M+1}^{\infty} \). The Lie algebra \( P^M(V) \) is graded, so can be expressed as

\[
P^M(V) = \oplus_{i=1}^{M} V_i
\]

where \( V_i+1 = [V, V_i]_\otimes \) and \( V_1 = V \). We define the dilation on \( P^M(V) \) to be the linear map \( \delta_t : P^M(V) \to P^M(V) \) such that

\[
\delta_t[h_1 + ... + h_M] = th_1 + ... + t^M h_M.
\]

Similarly, the dilation can be extended to the Lie Group \( \delta_t : G^M(V) \to G^M(V) \) for \( g = \exp(h_1 + ... + h_M) \) by

\[
\delta_t g = \exp(\delta_t h_1 + ... + t^M h_M).
\]

A homogeneous norm on a Carnot group is a function \( \| \cdot \|_G : G \to \mathbb{R}^+ \) such that for any \( g \in G \), \( \|g\|_G = 0 \) if and only if \( g = 1 \) the unit of \( \otimes \) and \( \|\delta_t g\|_G = |t| \cdot \|g\|_G \).

As the Lie Algebra \( P^M(V) \) is finite dimensional, all homogeneous norms on \( G^M \) are equivalent. By considering the collection of homogeneous norms, one can induce a left invariant metric over \( G^M(\mathbb{R}^d) \). This is traditionally called the Carnot-Carathéodory metric which we denote by \( d_{cc} \). Further, the Carnot-Carathéodory norm satisfies the additional properties for any \( g, g_1, g_2 \in G^m(\mathbb{R}^d) \)

\[
\|g_1 \otimes g_2\|_{cc} \leq \|g_1\|_{cc} + \|g_2\|_{cc} \quad \text{and} \quad \|g\|_{cc} = \|g^{-1}\|_{cc}.
\]

\( A^M \) be all the words generated by the Alphabet \( A \) such that \( |A| \leq M \). One example of a homogeneous norm that we work with is

\[
\|g\|_{G^M} = \sum_{A \in A^M} |\langle \log_{\otimes}(g), e_A \rangle|^{1/|A|}.
\]

A.2 Rough Paths

**Definition A.1.** Let \( V \) be a vector space. For a path \( x \in C^{1-\text{var}}([0,T]; V) \), the iterated integrals of \( x \) are canonically defined using Young integration. The collection of iterated integrals of the path \( x \) is called the truncated Signature of \( x \) and is defined as

\[
S_M(x)_{s,t} := 1 + \sum_{n=1}^{M} \int_{s \leq u_1 \leq ... \leq u_n \leq t} dx_{u_1} \otimes ... \otimes dx_{u_n} \in T^M(V) = \bigoplus_{n=0}^{M} V^{\otimes n}.
\]
It is well known that $S_M(x)$ takes values in $G^M(V)$.

**Definition A.2.** For $\alpha \in (0, 1)$ and let $M$ be the largest integer such that $M\alpha < 1$. A path $X: [0, T] \to G^M(V)$ is called an $\alpha$-Hölder continuous geometric rough paths if

$$
\langle X_{s,t}, e_A \rangle \langle X_{s,t}, e_B \rangle = \langle X_{s,t}, e_A \cup e_B \rangle, \quad \langle X_{s,u}, e_A \rangle = \langle X_{s,t} \otimes X_{t,u}, \Delta[e_A] \rangle
$$

and

$$
\sup_{A \in A_M} \sup_{s,t \in [0, T]} \|X_{s,t}, e_A\|_1 < \infty. \tag{A.2}
$$

**Definition A.3.** Denote $p = \frac{1}{\alpha}$. We define the $\alpha$-Hölder rough path metric

$$
d_{\alpha}(X, Y) = \|X^{-1} \otimes Y\|_\alpha = \sup_{s,t \in [0, T]} \left\|X^{-1}_{s,t} \otimes Y_{s,t}\right\|_{cc}^{\frac{1}{\alpha}}. \tag{A.3}
$$

By quotienting with respect to $X_0$, one can make this a norm. We use the convention that $\|X\|_{\|p\text{-var};[0,T]\}} = \|1^{-1} \otimes X\|_{\|p\text{-var};[0,T]\}}$ and $\|X\|_\alpha = \|1^{-1} \otimes X\|_\alpha$. We denote the metric space of $\alpha$-Hölder continuous geometric rough paths to be $G_{\Omega_\alpha}(\mathbb{R}^d)$.

Similarly, we define the homogeneous $p$-variation metric $d_{\|p\text{-var}\}}$ by

$$
d_{\|p\text{-var}\};[0,T]}(X, Y) := \|X^{-1} \otimes Y\|_{\|p\text{-var};[0,T]\}} := \left( \sup_{D=t_i} \sum_{i \in D} \left\|X^{-1}_{t_i, t_{i+1}} \otimes Y_{t_i, t_{i+1}}\right\|_{cc}^{p} \right)^{\frac{1}{p}}. \tag{A.4}
$$

When studying rough paths, one can either work with $p$-variation or $\alpha$-Hölder norms. For the most part, authors choose one and stick with it for the entirety of their work. While $p$-variation is slightly more general, $\alpha$-Hölder allows for a wavelet representation in the Banach space which is more favourable for this work.

It is important to understand that for this paper, we work with both norms. The Hölder norm, being more restrictive, is assumed to be the bound on regularity. However, we are required to work with the $p$-variation in order to establish an integrability condition.

**Definition A.4.** Let $\Delta_T = \{(s,t) : 0 \leq s \leq t \leq T\}$ denote the two-dimensional simplex. The map $\omega: \Delta_T \to \mathbb{R}^+$ is a Control if it is a continuous, non negative, super-additive function which vanishes on the diagonal.

**Example A.5.** Suppose that $X$ is a geometric rough path with finite $p$-variation, so that Equation (A.4) is finite. Then $\omega(x, t) := \|X\|_{\|p\text{-var};[s,t]\}}$ is a control.

The Carnot-Carathéodory metric as already described takes its structure from the Group $G^M(\mathbb{R}^d)$ and so is homogeneous with respect to the group dilation $\delta_\lambda$. However, there is another metric that takes its structure from the vector space $T^M(\mathbb{R}^d)$.

For two elements $g_1, g_2 \in T^M(\mathbb{R}^d)$ and $i \in \{1, \ldots, M\}$ we have the collection of pseudo-metrics

$$
\rho_i(g_1, g_2) = \sum_{A \in A_M} \sum_{|A| = i} \left| \langle g_1, e_A \rangle - \langle g_2, e_A \rangle \right|. \tag{A.5}
$$

We also have the inhomogeneous Tensor metric

$$
\rho(g_1, g_2) = \max_{i=1,\ldots,M} \rho_i(g_1, g_2).
$$
**Definition A.6.** Let $p = \frac{1}{\alpha} > 2$. For a fixed control $\omega$, we define the inhomogeneous $\omega$-modulus metric to be

$$\rho_{p-\omega;[0,T]}(X, Y) := |X_0 - Y_0|_{T(p)(\mathbb{R}^d)} + \max_{i=1,\ldots,[p]} \sup_{s,t \in [0,T]} \frac{\rho_i(X_{s,t}, Y_{s,t})}{\omega(s,t)^{1/p}}. \tag{A.6}$$

When we additionally have that $\omega(s,t) \leq C|t-s|$ where $C$ is a constant independent of $s,t$, we also have the inhomogeneous $\alpha$-Hölder metric to be

$$\rho_{\alpha;[0,T]}(X, Y) := |X_0 - Y_0|_{T(p)(\mathbb{R}^d)} + \max_{i=1,\ldots,[p]} \sup_{s,t \in [0,T]} \frac{\rho_i(X_{s,t}, Y_{s,t})}{|t-s|^\alpha}. \tag{A.7}$$

The inhomogeneous rough path metrics satisfy the simple relation

$$\rho_{p-\text{var};[0,T]}(X, Y) \leq \left( 1 \vee \max_{i=1,\ldots,[p]} \omega(0, T)^{1/p} \right) \rho_{p-\omega;[0,T]}(X, Y) \leq \rho_{p;[0,T]}(X, Y)$$

by simple manipulation of the standard relation between $p$-variation and $\frac{1}{p}$-Hölder regularity, see [FV10b].

**Definition A.7.** Let $E$ and $F$ be normed spaces. A map $f : E \rightarrow F$ is called $\gamma$-Lipschitz (in the sense of Stein) if $f$ is $[\gamma]$ continuously differentiable (in the sense of Fréchet) and such that there exists a constant $M < \infty$ such that the supremum norm of the $k^{th}$ derivative for $k = 1, \ldots, [\gamma]$ and the $[\gamma]$-Hölder norm of its $[\gamma]^{th}$ derivative are bounded by $M$. The smallest $M \geq 0$ satisfying this condition is the $\gamma$-Lipschitz norm of $f$, denoted $\|f\|_{\text{Lip}^\gamma}$. The space of all such functions is denoted $\text{Lip}^\gamma(E, F)$.

We also emphasise the distinction between $\text{Lip}^1_\gamma(E, F)$, the space of functions $f : E \rightarrow F$ that are Lipschitz.

**Theorem A.8 ([LV07]).** Let $V = \bigoplus V^j$ be a vector space.

Let $\alpha < 1/2$ such that $\frac{1}{\alpha} \notin \mathbb{N}$ and $M = [\frac{1}{\alpha}]$. Suppose that $X^j_i$ are $\alpha$-Hölder continuous paths taking values in $G^M(V^j)$. Then $\bigoplus V^j_i$ can be thought of as an $\alpha$-Hölder continuous path taking values in $\bigoplus V^j_i$ and there exists an extension $X_t$ taking values in $G^M(V)$ that is $\alpha$-Hölder continuous with respect to the Carnot norm on $G^M(V)$.

### A.3 Controlled Rough Path

A controlled rough path, first introduced in [Gub04], provides a path that is known to be adequately regular enough to be integrable with respect to a rough path.

Let $V$ and $U$ be vector spaces and denote by $L(V, U)$ the space of Linear operators from $V$ to $U$. We define $T(V^*, U) := \bigoplus_{n=0}^\infty L((V^*)^{\otimes n}, U)$ and use the convention that $L((V^*)^{\otimes 0}, U) = U$. As earlier, we are interested in the case where $V = \mathbb{R}^d$ and $U = \mathbb{R}^d$.

Given an element $X \in T(V^*)$ and $Y \in T(V^*, U)$, we naturally obtain $YX \in U$. Also, in practice we work in the truncated tensor algebra $T^M(V, U) := \bigoplus_{n=0}^M L((V^*)^{\otimes n}, U)$ obtained by quotienting with respect to the ideal $\bigoplus_{n=M+1}^\infty L((V^*)^{\otimes n}, U)$.

**Definition A.9.** Let $\alpha \in (0, 1/2)$, let $M$ be the smallest integer such that $M\alpha < 1$ and let $X \in G\Omega_\alpha(V)$. Let $A_v$ be the alphabet of $V$.

A $X$-controlled rough path $Y : [0, T] \rightarrow T^{M-1}(V, U)$ and a remainder term $R : \Delta_T \rightarrow T^{M-1}(V, U)$ is any path such that for any word $A$ of the alphabet for $A_v$

$$\langle Y_t, e_A \rangle - \langle Y_s, X_{s,t} \otimes e_A \rangle = \langle R_{s,t}, e_A \rangle,$$
where
\[
\sup_{t,s \in [0,T]} \frac{|\langle R_{s,t}, e_A \rangle|}{|t-s|^{(M-|A|)\alpha}} < \infty.
\]

The space of \(X\)-controlled rough paths, denoted \(D^{M,\alpha}_X([0,T]; U)\) is the vector space of all \(X\)-Controlled paths with the norm
\[
\|Y\|_{X,M,\alpha} = \sum_{A \in \mathcal{A}_M \setminus \{\emptyset\}} \left( \left\| \langle Y, e_A \rangle \right\|_{|A|-\text{H"{o}l};[0,T]} + \left\| \langle R, e_A \rangle \right\|_{M,\text{\H"{o}l};[0,T]} \right).
\]

Given an \(X\)-controlled rough path \(Y\) taking values on \(L(V,U)\), we define the integral
\[
\int_0^T Y_t dX_t = \lim_{|D| \to 0} \sum_{i,t_i \in D} \langle Y_{t_i}, X_{t_i,t_{i+1}} \rangle_{T^{M-1}(V,L(V,U)),T^M(V)}
\]
taking values in \(U\).

\textbf{Definition A.10.} The Shuffle product over \(T(V)\) can be represented as two Left and Right Half-shuffle products \(e_A \sqcup e_B = e_A < e_B + e_A \succ e_B\) that satisfy the identities
\[
\begin{align*}
(e_A < e_B) < e_C &= e_A < (e_B \sqcup e_C), \\
(e_A \succ e_B) < e_C &= e_A \succ (e_B < e_C),
\end{align*}
\]

\[
\begin{align*}
(e_A \sqcup e_B) > e_C &= e_A > (e_B > e_C).
\end{align*}
\]

Using the additional identity \(e_A < e_B = e_B > e_A\), equivalent to commutivity of \(\sqcup\), we observe that the Left and Right Half-shuffles satisfy a Left and Right Zinbiel identity. Thus > and < are sometimes referred to as Paraproduts. For any geometric rough path and any two words \(A\) and \(B\) we have
\[
\int_s^t \langle X_{s,r}, e_A \rangle d\langle X_{s,r}, e_B \rangle = \langle X_{s,t}, e_A \succ e_B \rangle = \langle \delta_{\succ}[X_{s,t}], e_A \otimes e_B \rangle.
\]  
(A.9)

where \(\delta_{\succ}\) is the Right Half-Unshuffle. Using the Right Half-Unshuffle, we are able to “stitch” two controlled rough paths together to obtain an object that will satisfy the Sewing Lemma, providing us with a meaningful way to integrate a Controlled Rough Path with respect to another controlled rough path.

\textbf{Theorem A.11.} Let \(Y\) and \(Z\) be \(X\)-controlled rough paths. Then by exploiting Equation (A.9) we obtain
\[
\int_0^T Y_t \otimes dZ_t = \lim_{|D| \to 0} \sum_{i,t_i \in D} Y_{t_i} \otimes Z_{t_i,t_{i+1}} + \langle \left( Y_{t_i} - Y_{t_i} \right) \otimes \left( Z_{t_i} - Z_{t_i} \right), \delta_{\succ}[X_{t_i,t_{i+1}}] \rangle.
\]

In a similar fashion, we obtain
\[
\int_s^t Y_{s,r} \otimes dZ_r = \lim_{|D| \to 0} \sum_{i,t_i \in D} \langle \left( Y_{t_i} - Y_{t_i} \right) \otimes \left( Z_{t_i} - Z_{t_i} \right), \delta_{\succ}[X_{t_i,t_{i+1}}] \rangle.
\]
Given an $X$-controlled rough path $\mathcal{Y}$, one can extend it to a rough path $\mathcal{Y}$ taking values in $G^M(U)$. Define the path $\mathcal{Y} : [0, T] \rightarrow G^M(U)$ by

$$Y_{s,t} = 1 + \sum_{k=1}^{M} \lim_{|D| \to 0} \sum_{i_t \in D} \left( (Y_{t_i} - Y_{t_{i+1}})^{\otimes k}, (\delta_\mathcal{Y})^k [X_{t_{i+1}}, t_{i+1}]. \right) \tag{A.10}$$

where the iterated coproduct $(\delta_\mathcal{Y})^k : T^M(V^*) \rightarrow T^M(V^*)^{\otimes k}$ is defined inductively by

$$ (\delta_\mathcal{Y})^2 = ((\delta_\mathcal{Y}) \otimes I)\delta_\mathcal{Y}, \quad (\delta_\mathcal{Y})^{k+1} = ((\delta_\mathcal{Y}) \otimes I^{\otimes k})\delta_\mathcal{Y}.$$ 

Proof of Theorem A.11. The ideas behind this proof are well understood, although to the best of the authors knowledge have not been written using the language of Zinbiel algebras before.

Firstly,

$$\int_s^t Y_r \otimes dZ_r = Y_s \otimes Z_{s,t} + \int_s^t Y_{s,r} \otimes dZ_r$$

and from the definition of controlled rough paths we have

$$Y_s \otimes Z_{s,t} = Y_s \otimes (Z_s, X_{s,t} - 1) + Y_s \otimes (R_{s,t}^Z, e_\varepsilon),$$

$$Y_{s,r} = (Y_s, X_{s,r} - 1) + (R_{s,r}^Y, e_\varepsilon).$$

Thus

$$\int_s^t Y_{s,r} \otimes dZ_r$$

$$= \left( Y_s - Y_s, \int_s^t X_{s,r} \right) + \left( \int_s^t R_{s,r}^Y, e_\varepsilon \right) \otimes \left( \int_s^t Z_s - Z_s, dX_{s,r} \right) + \left( \int_s^t dR_{s,r}^Z, e_\varepsilon \right)$$

$$= \left( Y_s - Y_s \otimes (Z_s - Z_s), \int_s^t X_{s,r} dX_{s,r} \right) + o(|t - s|)$$

as $|t - s| \to 0$ where we use the identity from Equation (A.9) and the regularity of Definition A.9.

Similarly

$$\int_s^t Y_r \otimes dZ_r = Y_s \otimes Z_{s,t} + \left( Y_s - Y_s \otimes (Z_s - Z_s), \delta_\mathcal{Y}[X_{s,t}] \right) + o(|t - s|).$$

Motivated by this, we verify the conditions of the Sewing Lemma (see [FH14, Lemma 4.2]) with

$$\Xi_{s,t} := Y_s \otimes Z_{s,t} + \left( Y_s - Y_s \otimes (Z_s - Z_s), \delta_\mathcal{Y}[X_{s,t}] \right).$$

Thus for $s < t < u \in [0, T]$,

$$\delta \Xi_{s,t,u} = \Xi_{s,u} - \Xi_{s,t} - \Xi_{t,u}$$

$$= -Y_{s,t} \otimes Z_{t,u} + \left( (Y_s - Y_s) \otimes (Z_s - Z_s), \delta_\mathcal{Y}[X_{s,t}] \right)$$

$$+ \sum_{A,B} \left( (Y_s - Y_s) \otimes (Z_s - Z_s) \right) [e_A \otimes e_B] \left( X_{s,t} \otimes X_{t,u}, \mathcal{Z}[e_A \geq e_B] \right)$$

39
where $\overline{\Delta}$ is the reduced Coproduct. Next, we substitute in for the increments using the identities

$$\langle Y_{t}, e_{A} \rangle - \langle Y_{s}, X_{s,t} \otimes e_{A} \rangle = \langle R_{s,t}, e_{A} \rangle,$$

$$\langle Z_{t}, e_{B} \rangle - \langle Z_{s}, X_{s,t} \otimes e_{B} \rangle = \langle R_{s,t}, e_{B} \rangle.$$  

Next, we use Sweedler notation to represent the identity

$$\overline{\Delta}[e_{A} \triangleright e_{B}] = \sum_{A',A''} \sum_{B',B''} e_{A' \coprod B'} \otimes e_{A'' \triangleright B''}.$$  

Therefore

$$\sup_{s,t,u \in [0,T]} \frac{\delta \Xi_{s,t,u}}{|u-s|^{1/\alpha}} = o \left( |u-s| \right).$$  

The ideas behind this proof are well understood (see [LCL07, p.74]) where $Y$ is the solution to a linear rough differential equation, although to the best of the authors' knowledge they have not been written before using the language of Zinbiel algebras and for general controlled rough paths. We refer the reader to the forthcoming preprint [CDFL20], where a proof is given of this result.

\[\square\]

### B Appendix

**Proof of Lemma 2.14** From Definition 2.8, we have

$$\mathfrak{E}_{n,r}(\mathcal{L}) = \min_{h_{1},\ldots,h_{n} \in E} \left( \int_{E} \min_{i=1,\ldots,n} \|x - h_{i}\|_{E}^{r} d\mathcal{L}(x) \right)^{1/r} \leq \min_{h_{1},\ldots,h_{n} \in P_{U}[E]} \left( \int_{E} \min_{i=1,\ldots,n} \|x - h_{i}\|_{E}^{r} d\mathcal{L}(x) \right)^{1/r},$$

$$\leq 2^{(r-1)/r} \min_{h_{1},\ldots,h_{n} \in P_{U}[E]} \left( \int_{E} \int_{P_{U}[E]} \min_{i=1,\ldots,n} \|P_{U}[x] - h_{i}\|_{E}^{r} d\mathcal{L}(P_{U}[x]) d\mathcal{L} \left( (I - P_{U})[x] \right) \right)$$

$$+ \int_{E} \left\| (I - P_{U})[x] \right\|_{E}^{r} d\mathcal{L}(x),$$

since by the assumption that $P_{U}$ is a projection on $\mathcal{H}$ (rather than $E$), the two laws $\mathcal{L} \circ (P_{U})^{-1}$ and $\mathcal{L} \circ (I - P_{U})^{-1}$ are independent with respect to the joint law $\mathcal{L}$. Exploiting this, we get

$$\mathfrak{E}_{n,r}(\mathcal{L}) \leq 2^{(r-1)/r} \left( \int_{(I - P_{U})[E]} \mathfrak{E}_{n,r}(\mathcal{L} \circ (P_{U})^{-1}) d\mathcal{L}( (I - P_{U})[x] ) \right)$$

$$+ \int_{E} \left\| x - P_{U}[x] \right\|_{E}^{r} d\mathcal{L}(x)^{1/r},$$

\[\square\]

**Proof of Proposition 3.1** Define the Covariance Kernel $\mathcal{S} : C^{\alpha,0}([0,T];\mathbb{R}^{d}) \rightarrow C^{\alpha,0}([0,T];\mathbb{R}^{d})$ by

$$\mathcal{S}[f]_{t} = \mathbb{E} \left[ f(W) \cdot W_{t} \right], \quad \mathcal{S} = ii^{*}$$

\[40\]
where \(i^*\) is the Spectral representation of \(\mathcal{S}\). For a Hilbert space \(\mathcal{H}\) and a Banach space \(E\), we define the Operator \(l\)-topology, on the collection of bounded linear operators \(i : \mathcal{H} \to E\) to be

\[
i(i) := \mathbb{E}\left[ \left\| \sum_{k \in \mathbb{N}} \int_{\mathcal{H}} \xi_k f_k \right\|^2_E \right]^{1/2}
\]

where \((h_k)_{k \in \mathbb{N}}\) is an orthonormal basis of \(\mathcal{H}\) and \(\xi_k\) are i.i.d normal random variables. It is well known, see for example \([FTJ79]\), that the closure in the \(l\)-topology of the finite rank operators is the compact operators. We wish to find the finite dimensional operator that best approximates the Spectral representation \(i\) of the Covariance Kernel \(\mathcal{S}\) of Brownian motion in the \(l\)-topology.

We follow the methods of \([BC19]\). Using Theorem 2.3, we can equivalently think of \(\mathcal{L}^W\) as a law over the Banach spaces of sequences \((W_{pm})_{(p,m) \in \Delta}\) that satisfy

\[
\sup_{(p,m) \in \Delta} 2^{p(\alpha-1/2)}|W_{pm}| < \infty, \quad \lim_{p \to \infty} 2^{p(\alpha-1/2)} \sup_{m=1, \ldots, 2^p} |W_{pm}| = 0.
\]

Equivalently, we think of elements of the dual space \(C^{\alpha,0}([0,T]; \mathbb{R}^d)^*\) as being sequences over \(\Delta\) that satisfy

\[
f = (f_{pm})_{(p,m) \in \Delta}, \quad \|f\|_{\alpha-\text{Hölder}} = \sum_{(p,m) \in \Delta} 2^{p(1/2-\alpha)}|f_{pm}| < \infty.
\]

It is well known that we work with the operator

\[
i(f) = \sum_{(p,m) \in \Delta} f_{pm} W_{pm},
\]

where \(W_{pm} = \int_0^T W_{pm}(s) H_{pm}(s) ds\) are independent normally distributed random variables with mean 0 and variance 1.

Thus

\[
\mathcal{S}[(f_{pm})_{(p,m) \in \Delta}](t) = \sum_{(p,m) \in \Delta} f_{pm} G_{pm}(t), \quad f\left(\mathcal{S}[f]\right) = \sum_{(p,m) \in \Delta} |f_{pm}|^2.
\]

We wish to maximise this Quadratic form subject to the linear condition

\[
\|f\| = \sum_{(p,m) \in \Delta} 2^{p(1/2-\alpha)}|f_{pm}| = 1.
\]

By a simple convexity argument, the functionals that attain this maximisation problem will be wavelet evaluation functionals and hence

\[
\lambda^{(001)} = \sup_{\|f\|_{\alpha-\text{Hölder}} = 1} f\left(\mathcal{S}[f]\right) = 1 = f^{(001)}\left(\mathcal{S}[f^{(001)}]\right),
\]

where \(f^{(001)} = (f^{(001)}_{pm})_{(p,m) \in \Delta}\) satisfies \(f^{(001)}_{00} = e_1\) and \(f^{(001)}_{pm} = 0\) else. We label \(\mathcal{S}[f^{(001)}](t) = x^{(001)}(t)\). We define \(\mathcal{S}_{\mu^{001}}[f] := \lambda^{(001)} f(x^{(001)}) x^{(001)}\) and \(\mathcal{S}_{001}[f] = \mathcal{S}[f] - \mathcal{S}_{\mu^{001}}[f]\).

By construction, we have that the operator \(\mathcal{S}_{\mu^{001}}\) is the Covariance Kernel of the 1-dimensional Gaussian measure that best approximates \(\mathcal{L}^W\) in mean square. Equivalently, the Spectral representation \(\mathcal{S}_{\mu^{001}} = i_{\mu^{001}} \mu_{\mu^{001}}^*\) yields that \(i_{\mu^{001}}\) is the 1-dimensional operator that best approximates \(i\) in the \(l\)-topology.
By repeating this method, we obtain a sequence of so-called “Rayleigh coefficients” and “Rayleigh Functionals” parametrised by \((q, n, i) \in \Delta \times \{1, \ldots, d\}\) as
\[
\lambda^{(qni)} = 2^{(2\alpha-1)}, \quad f^{(qni)} = (f^{(qni)}_{pm})_{(p,m) \in \Delta}, \quad f^{(qni)}_{pm} = \delta_{p,q} \delta_{m,n} e_i,
\]
and elements \(G_{qni}\) that are orthonormal in \(\mathcal{H}\).

For fixed \(N \in \mathbb{N}\), we obtain the first \(2^N\) elements of these sequences. We construct the projection operator \(P_N : C^{\alpha,0}([0, T], \mathbb{R}^d) \to C^{\alpha,0}([0, T], \mathbb{R}^d)\) defined by
\[
P_N[x](t) = \sum_{(q, n) \in \Delta_N} \sum_{i=1}^{d'} f^{(qni)}(x) G_{qni}(t)e_i.
\]

Next, we decompose the law \(L^W = \mu_N \ast L^W_N\) where \(\mu_N = L^W \circ P_N^{-1}\) and \(L^W_N = L^W \circ (I - P_N)^{-1}\). \(\mu_N\) is a \(2^N\)-dimensional multivariate Gaussian distribution. \(L^W_N\) is a Gaussian measure over \(C^{\alpha,0}([0, T], \mathbb{R})\) with Kernel \(S_N\) that satisfies
\[
\sup_{f \in C^{\alpha,0}([0, T], \mathbb{R})} f\left(S_N[f]\right) \leq 2^{(N+1)(\alpha-1/2)}.
\]

In particular, for a random variable \(W\) with law \(L^W\) we have that random variable
\[
P_N[W](t) = \sum_{(p,m) \in \Delta_N} W_{pm} G_{pm}(t)
\]
has law \(\mu_N\) and
\[
\sup_{f \in C^{\alpha,0}([0, T], \mathbb{R})} \mathbb{E}\left[f\left(W - P_N[W]\right)\right] = 2^{(N+1)(2\alpha-1)}.
\]

**Proof of Theorem 4.16** Let \(V = \mathbb{R}^{d' \times n}\) and \(U = \mathbb{R}^{d' \times n}\) with alphabets \(\mathcal{A}\) and \(\hat{\mathcal{A}}\) both with \(n\) subalphabets
\[
\mathcal{A}^j = \{(i, j) : i \in \{1, \ldots, d'\}\}, \quad \hat{\mathcal{A}}^j = \{(i, j) : i \in \{1, \ldots, d\}\}.
\]
where \(j \in \{1, \ldots, n\}\). Thus all the vector spaces \(V^j\) are isomorphic to \(\mathbb{R}^{d'}\) and \(U^j\) are isomorphic to \(\mathbb{R}^d\) but each \(V^j\) and \(U^j\) is distinct and identifiable. As with the normal subgroup constructed in Equation (4.5), we know the normal subgroup that generates the cosets for the quotient group is
\[
I^M(\mathbb{R}^{d' \times n}) = \{h \in T^M(\mathbb{R}^{d' \times n}) : \langle h, e_I \rangle = 0, \forall I \text{ s.t. } \exists j \in \{1, \ldots, n\} \text{ with } I \in \mathcal{A}^j\},
\]
\[
K^M(\mathbb{R}^{d' \times n}) = \exp(\mathcal{I}^M(\mathbb{R}^{d' \times n})).
\]
For \(s, t \in [0, T]\), let
\[
\pi_{G^M(\mathbb{R}^{d' \times n})/K^M(\mathbb{R}^{d' \times n})}[\tilde{W}_{s,t}] = \tilde{W}_{s,t} \otimes K^M(\mathbb{R}^{d' \times n}) = W_{s,t}.
\]
By Theorem A.11, we know this is equal to
\[
X_{s,t} := 1 + \lim_{|D| \to 0} \sum_{t_1 \in D} \sum_{k=1}^{M} \left( \sum_{i=1}^{d} \left( X_{t_i} - X_{t_i} \right)^{\otimes k}, (\delta_{e_i})^k[\tilde{W}_{s,t},t_{i+1}] \right) + B(X_{t_i})(t_{i+1} - t_i), \quad (B.1)
\]
42
and $X$ is defined as in Equation (4.8). It is important to realise that the drift term, the only term that contains the “measure like” contributions, is only included in the first level of the signature. Measure dependencies are generally smoother than path dependencies and their higher regularity means they are $o(D^{1+})$.

Next for $t_i \in D$, we have

$$(X_{t_i} - X_{t_i}) \otimes^k = \left( \Sigma(X_{t_i}) + \Sigma \Sigma(X_{t_i}) + \ldots + \Sigma^{(M-1)}(X_{t_i}) \right) \otimes^k$$

$$= \sum_{l_1, \ldots, l_k=1}^{M-1} \otimes^k \Sigma^{l_m}(X_{t_i}).$$

Using Definition 4.13 and Lemma 4.15 we have that there exists $f_j : U^j \to L\left((V^j)^{\otimes l_m}, U^j\right)$ such that

$$\Sigma^{l_m}(X_s) = \text{Diag}_{j=1, \ldots, n} f_j(\langle X_s, e_{(\cdot, j)} \rangle).$$

Similarly, there exist functions $g_j : U^j \to L\left((V^j)^{\otimes (l_1+\ldots+l_k)}, U^j\right)$ such that

$$\otimes^k \Sigma^{l_m}(X_s) = \text{Diag}_{j=1, \ldots, n} g_j(\langle X_s, e_{(\cdot, j)} \rangle),$$

which is an operator restricted to the subgroup $\oplus j = 1^n G^M(V^j)$. Thus Equation (B.1) is dependent on the tensor of rough paths $W$ and not on the Extension $\tilde{W}$.

References

[ADP90] J.-P. Aubin and G. Da Prato, Stochastic viability and invariance, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 17 (1990), no. 4, 595–613. MR1093711

[BAGL94] G. Ben Arous, M. Grădinaru, and M. Ledoux, Hölder norms and the support theorem for diffusions, Ann. Inst. H. Poincaré Probab. Statist. 30 (1994), no. 3, 415–436. MR1288358

[BC19] X. Bay and J.-C. Croix, Karhunen-Loève decomposition of Gaussian measures on Banach spaces, Probability and Mathematical Statistics 39 (2019), no. 2.

[BCCP98] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti, A non-Maxwellian steady distribution for one-dimensional granular media, J. Statist. Phys. 91 (1998), no. 5-6, 979–990. MR1637274

[BCM07] M. Burger, V. Capasso, and D. Morale, On an aggregation model with long and short range interactions, Nonlinear Anal. Real World Appl. 8 (2007), no. 3, 939–958. MR2307761

[BG13] F. Bolley, I. Gentil, and A. Guillin, Uniform convergence to equilibrium for granular media, Arch. Ration. Mech. Anal. 208 (2013), no. 2, 429–445. MR305983

[Bit13] J. Bittencourt, Fundamentals of plasma physics, Springer Science & Business Media, 2013.

[BL16] J. M. Borwein and S. B. Lindstrom, Meetings with Lambert $W$ and other special functions in optimization and analysis, Pure Appl. Funct. Anal. 1 (2016), no. 3, 361–396. MR3622396
[McK66] H. P. McKean Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Nat. Acad. Sci. U.S.A. **56** (1966), 1907–1911. MR221595

[MSS94] A. Millet and M. Sanz-Solé, *A simple proof of the support theorem for diffusion processes* **1583** (1994), 36–48. MR1329099

[Pis89] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Mathematics, vol. 94, Cambridge University Press, Cambridge, 1989. MR1036275

[Pop01] S. Pope, *Turbulent flows*, IOP Publishing, 2001.

[PS11] G. Pagès and A. Sellami, *Convergence of multi-dimensional quantized SDE’s*, Séminaire de Probabilités XLIII, 2011, pp. 269–307. MR2790377

[Rie17] S. Riedel, *Transportation-cost inequalities for diffusions driven by Gaussian processes*, Electron. J. Probab. **22** (2017), Paper No. 24, 26. MR3622894

[SV72] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, 1972, pp. 333–359. MR0400425

[Zab00] J. Zabczyk, *Stochastic invariance and consistency of financial models*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **11** (2000), no. 2, 67–80. MR1797512