Homogenization with fractional random fields

Johann Gehringer and Xue-Mei Li*
Imperial College London

27 November 2019

Abstract

We consider a system of differential equations in a fast long range dependent random environment and prove a homogenization theorem involving multiple scaling constants. The effective dynamics solves a rough differential equation, which is 'equivalent' to a stochastic equation driven by mixed Itô integrals and Young integrals with respect to Wiener processes and Hermite processes. Lacking other tools we use the rough path theory for proving the convergence, our main technical endeavour is on obtaining an enhanced scaling limit theorem for path integrals (Functional CLT and non-CLT's) in a strong topology, the rough path topology, which is given by a Hölder distance for stochastic processes and their lifts. In dimension one we also include the negatively correlated case, for the second order / kinetic fractional BM model we also bound the error.

Keywords: fractional Brownian motion, slow-fast systems, homogenization, Hermite processes, passive tracer, random environment, functional limit theorem, rough path

MSC Subject classification: 34F05, 60F05, 60F17, 60G18, 60G22, 60H05, 60H07, 60H10

Contents

*xue-mei.li@imperial.ac.uk

1 Introduction

In this article we prove a homogenization theorem to the following slow/fast system with long range dependent random environment and multiple scaling constants,

$$\dot{x}_t^\varepsilon = \sum_{k=1}^N \alpha_k(\varepsilon) f_k(x_t^\varepsilon) G_k(y_t^\varepsilon), \tag{1.1}$$

showing that the solutions converge. Here ε is a small positive parameter, y_t models a stationary long range dependent fast random environment, $G_k \in L^p \cap \mathcal{C}^0$ are centred (not necessarily in a finite chaos), and $\alpha_k(\varepsilon)$ are the scaling constants to be identified. The x_t^ε process models the position of a particle in a moving environment, y^ε is a stationary stochastic process moving at microscopic time scale. When f is divergence free, this is a popular model for passive tracers in a tubulent fluid. By homogenization we mean the following phenomenon: during a finite macroscopic period, the fast environment would have typically been everywhere, its effects can therefore be absorbed into one effective vector field. This way one obtains an autonomous equation whose solution approximate the position of the particle when the parameter ε is small.

Noise with long span of interdependence between their increments has attracted the attention of many mathematicians and physicists. In a study for loss in water storage, Hurst et al [HBS] observed long range time dependence in the time series data of water flows and found that the time dependence varies proportionally to t^H where $H \sim 0.73$. Economical data also exhibits cycles of varying lengths. By contrast, Brownian motions and stable processes have independent increments. Benoit Mandelbrot and John Van Ness introduced the use of fractional Brownian motions (fBM) in [MVN68] and found they are good models for the Hurst phenomenon, and the best among other models they compare with. Recall that a fBM is a continuous Gaussian process with $\mathbf{E}(B_t - B_s)^2 = |t-s|^{2H}$. (When $H = \frac{1}{2}$, this is the BM.) They are self-similar with similarity exponent H.

Self-similarity attracted attention also from Sinai, Dobrushin, and Jona-Lasinio for their relevance in mathematically rigorous description of critical phenomena and in the renormalisation theory. In [Sin76], for example, Sinai constructed non-Gaussian self-similar fields; while Dobrushin [Dob79] studied self-similar fields subordinated to self-similar Gaussian fields (multiple Itô integrals). Those self-similar stochastic processes with stationary increments are a particular interesting class. When normalized to begin at 0, to have mean 0 and variance 1, at $t = 1$, they necessarily have the covariance $\frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$. Those of Gaussian variety are fBMs. Hermite processes are non-Gaussian self-similar processes with the above mentioned covariance and stationary increments. They appear as scaling limits of functionals of long range dependent Gaussian processes. The first of these appeared in [Ros61], in which Rosenblatt constructed an example of a non-strong mixing sequence of random variables. He proved that the afore-mentioned sequence (with slow decaying auto-correlation) is not strong mixing by proving that the usual central limit theorem (CLT) fails and obtained a non-Gaussian scaling limit which is in fact a rank 2 Hermite process. Rank 2 processes are called Rosenblatt process. Jona-Lasinio was also concerned with the construction of a systematic theory of limit distributions for sums of 'strongly dependent' random variables for which the classical central limit theorems does not hold, [JL77], see also the book [EM02]. These processes also appear in our effective dynamics, in a mixed manner.

Despite of the evidence pointing to long range depend noise, the study of slow/fast systems has predominatedly focused on those with strongly mixing or Markovian properties. If the correlation of the random field decays sufficiently fast, see [Tay21, Gre51, Kub57], for small parameters the particle is expected to behave diffusively and can be approximated by a Markov process with covariance given by the integral, if finite, of the correlation functions of the vector field. If the correlation decays so slowly that the integral is infinite, the random field is said to have long range dependence.

We will take y_t^ε to be a fast fractional Ornstein-Uhlenbeck process (fOU). These are defined by the Langevin equation driven by fBM's, and also have long range dependence when the Hurst parameter $H > \frac{1}{2}$, see §3.2. These are fascinating processes. On one hand the solutions of the Langevin equation forgets its initial position exponentially fast. On the other hand its auto-correlation function, which measures how much the shifted process remembers, exhibits power law decay. The latter is not shared by all other functionals of fBMs. For example it was shown in [KKR12], that Donsker's Invariance Principle holds for fBM on the torus, in this case the correlation from the fBM is forgotten and lost in the wrapping. It is natural to expect the same loss of memory in fOU. After all, the linear contraction in the Langevin equation and the exponential convergence of the solutions would lead to the belief that it mixes as fast as the wrapped fBM, this is not so. Indeed, for $H > \frac{1}{2}$ it is not strong mixing, its auto correlation function is not integrable. The long range memory survives also in the second order model (the Kinetic model for fBMs).

For the equation $\dot{x}_t^\varepsilon = \varepsilon^{H-1} f(x_t^\varepsilon)\, y_t^\varepsilon$ on $\mathbf{R}$, it is easy to show

$$\left\| |x^\varepsilon - x|_{C^{\gamma'}([0,T])} \right\|_{L^p} \lesssim T^\gamma \varepsilon^{H-\gamma},$$

where $0 < \gamma' < \gamma < H$, and $H > \frac{1}{3}$ (this latter restriction is only needed for the error control).

The limit solves the Young differential equation: $\dot{x}_t = f(x_t)\, dX_t$, where X_t is a fBM, so the effective limit resembles, locally, a fBM. For passive tracers in homogeneous incompressible fluid, there are some studies [FK00, KNR12], in all of these eferences, the effective equation is driven by either a BM or by a fBM, the method is also different.

However, the effective dynamics for (1.1) will involve in general a mix of BM's, fBM, and the non-Gaussian Hermite processes. The appropriate scaling constants depend on the functions G_i. For example for $N = 1$ and m the Hermite rank of G_1, $m = \frac{1}{2(1-H)}$ is the critical value for the limit to be locally Gaussian. If m is small, the effective limit is locally the Hermite process of rank m.

We comment briefly on the effective equations. They are stochastic equations driven simultaneously by BM's, fBM's, Rosenblatt processes, and higher rank Hermite processes. For $H > \frac{1}{2}$ the integrals with respect to Hermite processes can be defined as Young integrals, those with respect to the Wiener components are Itô/ Stratonovich integrals. We show that the Wiener process part of the driving limit is independent of the Hermite processes, and the components of the Hermite processes can be written as multiple integrals with respect to the same Brownian motion W_t. This means if we fix a sample path $W_t(\omega)$, we can consider the limit equation as a mixed Wierner-Young integral equation. Mixed equations driven by BM's and fBM have been studied for example in [GN08, dSEE18, LH19]. Our convergence is actually in a strong topology, in C^γ and in the rough path topology. The limit equation is actually a rough differential equation, whose solution is in general not a semi-martingale and is defined for all chance variables, the driver is a stochastic process of Hölder regularity class, enhanced by iterated integrals.

Lacking other tools, we will use the solution theory for rough path differential equations to establish the required convergence, see [Lyo94, FH14], in the first p-variation norms are used. Here it is convenient to use the Hölder path formulation and so we follow the notation in [FH14]. However using the p-variation norms may help to improve the integrability conditions in the Functional limit theorem. Due to the length of the article we do not study that aspect. To use the continuity theorem of the Itô solution maps, we rewrite (1.1) as rough differential equations driven by stochastic processes with a parameter ε. It is then sufficient to prove the convergence of the drivers in the rough path topology, which is finer than the Hölder topology. We first prove the joint convergence of the drivers together with their iterated integrals in an appropriate Hölder space, in the finite dimensional distributions. Then we show that they converge also in the rough path topology. One of our key technical endeavours is therefore a vector valued functional 'Central' Limit Theorem in the rough path topology, this is accomplished in sections 4 and 5. The statements of the main results will be presented in §2. The preliminary computations are presented together with a simple example in §3. In §4.3, we treat the 1-dimensional case, its proof does not use the extensive estimations obtained later, nor the CLT in rough path topology (those in §4 are sufficient). The results for dimension 1 is of course stronger, including all Hurst parameters, however to single this case out, we hope to make transparent the proof for the multi-dimension and multi-scale case. The proof for the main theorem is finalised in §6. For reader's convenience, interpretation of the rough differential equation, studied of fOU processes, kernel convergence for multiple Wiener integrals, their asymptotic independence, and various preliminary estimates are presented in the appendix.

Notation

- $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $\|\bullet\|_p$ or $\|\bullet\|_{L^p}$ denotes the norm in $L^p(\Omega)$, also when we refer to the sapce L^p we mean $L^p(\Omega, \mathcal{F}, \mathbf{P})$.
- $W_t, t \in \mathbf{R}$ denotes a two-sided Wiener process.
- $\mathcal{F}_t$ denotes the filtration generated by the fractional Brownian Motion.
- H is the Hurst parameter of the fBM.
- $H^*(m) = m(H-1)+1$.
- m_k is the Hermit rank of G_k.
- Convention : $H^*(m_k) \le \frac{1}{2}$ for $k \le n$; otherwise $H^*(m_k) > \frac{1}{2}$,
- $\mu = N(0,1)$ is the standard Gaussian distribution.
- $L^p(\mu)$ denote the set of L^p integrable from $\mathbf{R}$ to $\mathbf{R}$ with respect to μ.
- $BC^r = \mathcal{C}_b^r$: bounded continuous functions with bounded continuous derivatives up to order r.
- $\mathcal{C}_a^r = C^r \cap BC^a$.
- $f \lesssim g$ means there exists a constant c such that $f \le cg$.
- $|x|_\alpha := \sup_{s\neq t} \frac{|x_t - x_s|}{|t-s|^\alpha}$ is the homogeneous Hölder semi-norm, $0 < \alpha < 1$.
- For a process x_t, set $x_{s,t} := x_t - x_s$.

2 Formulation of main results

We let $\{H_m, m \ge 0\}$ denote the orthogonal Hermite polynomial of degree m on $L^2(\mu)$, normalised to have leading coefficient 1 and $L^2(\mu)$ norm $\sqrt{m!}$. For any $H \in (0,1)$, we define the non-increasing transformation $m \to H^*(m)$,

$$H^*(m) = m(H-1)+1. \tag{2.1}$$

Definition 2.1 Let $G : \mathbf{R} \to \mathbf{R}$ be an $L_2(\mu)$ function with chaos expansion

$$G(x) = \sum_{k=m}^{\infty} c_k H_k(x), \qquad c_k = \frac{1}{k!}\langle G, H_k\rangle_{L^2(\mu)}. \tag{2.2}$$

(Observe that $\int G d\mu = 0$ if and only if $c_0 = 0$.)

1. The smallest m with $c_m \neq 0$ is called the Hermite rank of G.
2. If $H^*(m) \le \frac{1}{2}$ we say G has *high Hermite rank* (relative to H), otherwise it is said to have *low Hermite rank*.

For any $m \in \mathbf{N}$, we define a set of scaling constants as below, the intuition leading to this will be clear when we present the relevant central and non-central limit theroems.

$$\alpha(\varepsilon, H^*(m)) = \begin{cases} \frac{1}{\sqrt{\varepsilon}}, & \text{if } H^*(m) < \frac{1}{2}, \\ \frac{1}{\sqrt{\varepsilon |\ln(\varepsilon)|}}, & \text{if } H^*(m) = \frac{1}{2} \\ \varepsilon^{H^*(m)-1}, & \text{if } H^*(m) > \frac{1}{2}. \end{cases} \tag{2.3}$$

We fix a fractional Brownian motion B_t^H, with Hurst parameter $H \in (0,1) \setminus \{\frac{1}{2}\}$ (Homogenization for $H = \frac{1}{2}$ is classic, the result is independent of the Hermite rank and the scaling is given by $\alpha(\varepsilon) = \frac{1}{\sqrt{\varepsilon}}$).

Let y^ε be a fast stationary fractional Ornstein-Uhlenbeck process with standard Gaussian distribution. We also write $y = y^1$ for simpicity.

Convention 2.2 Given a collection of functions $(G_k, k \le N)$, we will label the high rank ones first, so the first n functions satisfy $H^*(m_k) \le \frac{1}{2}$, where $n \ge 0$, and the rest has $H^*(m_k) > \frac{1}{2}$.

2.1 Homogenization

Let $G_k : \mathbf{R} \to \mathbf{R}$ be centred function in $L_2(\mu)$ with Hermite rank m_k. Write

$$G_k = \sum_{l=m_k}^{\infty} c_{k,l} H_l, \qquad \alpha_k(\varepsilon) = \alpha(\varepsilon, H^*(m_k)).$$

We consider

$$\begin{cases} \dot{x}_t^\varepsilon = \displaystyle\sum_{k=1}^{N} \alpha_k(\varepsilon)\, f_k(x_t^\varepsilon)\, G_k(y_t^\varepsilon), \\ x_0^\varepsilon = x_0. \end{cases} \tag{2.4}$$

For the main theorem below, we assume that $G_k \in L_{p_k}$ for sufficiently high p_k, and G_k satisfies a fast chaos decay condition. *Both assumptions are automatically satisfied if G_k are polynomials,* in which case we take $p_k = \infty$ in the statement below, then the only extra condition is that the Hermite rank of G_k is not in the interval $[\frac{1}{1-H}, \frac{1}{2(1-H)}]$. The precise assumption will be detailed after the statement.

Theorem A (§6) *Given $H \in \left(\frac{1}{2}, 1\right)$, $f_k \in \mathcal{C}_b^3(\mathbf{R}^d; \mathbf{R}^d)$, and G_k's satisfying Assumption 2.5 below. Then, the solutions of (2.4) converge weakly in $\mathcal{C}^\gamma$, on any finite time interval, for any $\gamma \in (\frac{1}{3}, \frac{1}{2} - \frac{1}{\min_{k \le n} p_k})$, to the solution of the following stochastic differential equation*

$$dx_t = \sum_{k=1}^{n} f_k(x_t) \circ dX_t^k + \sum_{k=n+1}^{N} f_k(x_t) dX_t^k, \quad x_0 = x_0,$$

where X_t^k is a Wiener process for $k \le n$, a Hermite processes for $k > n$, and $\circ$ denotes Stratonovich integral, otherwise a Young integral.

Definition 2.3 A function $G \in L^2(\mu)$, $G = \sum_{l=0}^\infty c_l H_l$, is said to satisfy the fast chaos decay condition with parameter $q \in \mathbf{N}$, if

$$\sum_{l=0}^{\infty} |c_l|\, \sqrt{l!}\, (2q-1)^{\frac{l}{2}} < \infty.$$

For Hermite polynomials and Gaussian measures we have the estimates:

$$\|H_k\|_{2q} \le (2q-1)^{\frac{k}{2}} \sqrt{\mathbf{E}(H_k)^2} = (2q-1)^{\frac{k}{2}} \sqrt{k!}.$$

Consequently, if G satisfies the fast chaos decay condition with parameter q, then

$$\|G\|_{2q} \le \sum_{l=0}^{\infty} |c_l| \|H_l\|_q < \infty,$$

and $G \in L_{2q}$. Observe that $\frac{1}{2} - \frac{1}{2q} > \frac{1}{3}$, a condition needed for the convergence in $\mathcal{C}^\gamma$, is equivalent to $q > 3$. Also, if G satisfies the decay condition with $q > 1$, then G is continuous. Indeed, we have

$$\left| e^{-\frac{x^2}{2}} H_k(x) \right|_\infty \leq 1.0865\sqrt{k!},$$

from [AS84, pp787]. (The polynomials in [AS84] are orthogonal with respect to $e^{x^2}dx$.) Thus the power series $e^{-\frac{x^2}{2}} \sum_{l=0}^\infty c_l H_l$ converges uniformly in x, the limit G is continuous.

Remark 2.4 If G satisfies the fast chaos decay condition with parameter $q > 1$, then G has a representation in $L^{2q} \cap \mathcal{C}$, with which we will work from here on.

Assumption 2.5 (Functional limit rough $\mathcal{C}^\gamma$- assumptions) *Each G_k belongs to $L^{p_k}(\mu)$, where $p_k > 2$, and has Hermite rank $m_k \geq 1$. Furthermore,*

(1) Each G_k satisfies the fast chaos decay condition with parameter $q \geq 4$.
(2) (Integrability condition) p_k is sufficiently large so the following holds:

$$\min_{k \leq n} \left(\frac{1}{2} - \frac{1}{p_k} \right) + \min_{n < k \leq N} \left(H^*(m_k) - \frac{1}{p_k} \right) > 1. \tag{2.5}$$

(3) If G_k has low Hermite rank, i.e. $H^*(m_k) > \frac{1}{2}$, assume $H^*(m_k) - \frac{1}{p_k} > \frac{1}{2}$.
(4) Either $H^*(m_k) < 0$ or $H^*(m_k) > \frac{1}{2}$.

Remark 2.6

1. The higher than usual moment assumptions arise from the necessity to obtain the convergence, not just in the space of continuous functions but also, in a rough path space $\mathscr{C}^\gamma$ for some $\gamma > \frac{1}{3}$ (which is naturally established by Kolmogorov type arguments) to be able to use the continuity of the solution maps in the rough path setting.
2. Condition (2) makes sure that the Hölder regularity of the terms which converge to a Wiener processes is at least η and the ones for the terms which converge to Hermite processes is at least τ with $\eta + \tau > 1$. This condition ensures that iterated integrals, in which one term converges to a Wiener and the other one to a Hermite process, can be interpreted as a Young integral.
3. In Condition (4) we have to assume $H^*(m_k) < 0$ instead of $H^*(m_k) \leq \frac{1}{2}$. This means we exclude functions with Hermite rank in $[\frac{1}{1-H}, \frac{1}{2(1-H)}]$. This restriction is due to Lemma 7.7, where we deal only with high Hermite rank functions, we only obtain the required integrability estimates for $H^*(m_k) < 0$.

The homogenisation problem for a passive tracer in a turbulent flow has been studied in [FK00, KNR12]. For a class of spatial homogeneous (incompressible) time stationary vector fields whose spectral density satisfies suitable conditions, they showed that the effective limit is either a Brownian motion or a fractional Brownian motion. A class of homogenization theorems was shown in [KM17], their integrability conditions were then lowered in [CFK+19] by using the p-variation rough path formulation instead of the Hölder one, a related work is also to be found in [BC17].

2.2 Lifted functional limit theorem

The static problem preluding the homogenization are functional limit theorems. Once appropriate limit theorems for the drivers are established, we may use the continuity theorem for rough differential equations.

For continuous processes this concerns the scaling limit of $\int_0^t G(y_{\frac{s}{\varepsilon}})ds$ where G is a a centred function. Let functions $(G_1, \ldots, G_N)$ be given. The pivot theorem is concerned with the scaling limit , as $\varepsilon \to 0$, for

$$X^\varepsilon := \big(X^{1,\varepsilon}, \ldots, X^{N,\varepsilon}\big), \qquad X^{k,\varepsilon} = \alpha_k(\varepsilon) \int_0^t G_k(y_s^\varepsilon) ds. \tag{2.6}$$

We further define the rough paths $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^{i,j,\varepsilon})$, where

$$\mathbb{X}_{u,t}^{i,j,\varepsilon} := \int_u^t (X_s^{i,\varepsilon} - X_u^{i,\varepsilon}) d\, X_s^{j,\varepsilon} = \alpha_i(\varepsilon)\alpha_j(\varepsilon) \int_u^t \int_u^s G_i(y_s^\varepsilon) G_j(y_r^\varepsilon)\, dr ds. \tag{2.7}$$

We call $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^{i,j,\varepsilon})$ the canonical lift of X^ε.

Such limit theorems are closely related to those for sums of sequence of random variables. For independent or strong mixing sequences, there is a central limit theorem (CLT), and the weak limit is always a Brownian motion. For interdependent long range dependent stationary sequences, this was pioneered by Rosenblatt, who constructed a not strong mixing stationary sequence, with a non-CLT limit, the limit is later known as the Rosenblatt process. For stationary continuous time strong mixing processes, the CLT states that $\sqrt{\varepsilon} \int_0^t G(y_{\frac{s}{\varepsilon}})ds$ converges weakly to a Markov process, this is classical and well understood. For stochastic processes whose auto-correlation functions do not decay sufficiently fast at infinity, there is no reason to have the $\sqrt{\varepsilon}$ scaling or to have a diffusive limit, see [BT13, Taq79, Ros61, BH02, DM79, BM83, BC09, HNX14, CNN18].

To state the main theorem clearly, we follow Convention 2.2 and label the first n functions such that $H^*(m_k) \le \frac{1}{2}$ for $k \le n$. Therefore, we will write

$$X^\varepsilon = (X^{W,\varepsilon}, X^{Z,\varepsilon}), \qquad X^{W,\varepsilon} \in \mathbf{R}^n,\ X^{Z,\varepsilon} \in \mathbf{R}^{N-n}. \tag{2.8}$$

Assumption 2.7 (Functional Limit $\mathcal{C}^\gamma$ assumptions) *Let $G_k \in L_{p_k}$ with Hermite rank $m_k \geq 1$.*

(1) If $H^*(m_k) \le \frac{1}{2}$ assume $\frac{1}{2} - \frac{1}{p_k} > \frac{1}{3}$, which is equivalent to $p_k > 6$.
(2) If $H^*(m_k) > \frac{1}{2}$ assume $H^*(m_k) - \frac{1}{p_k} > \frac{1}{2}$.

For the convergence in finite dimensional distributions we only assume that each $G_k \in L_2$, only for convergence in $\mathcal{C}^\gamma$ we need the above assumption. The following is extracted from section 5.

Theorem B *[CLT] Let $G_1, \ldots, G_N \in L^2(\mu)$ be given and let $H \in (0,1)$.*

(a) Then, there exists $X^W = (X^1, \ldots, X^n)$ and $X^Z = (X^{n+1}, \ldots, X^N)$, such that

$$(X^{W,\varepsilon}, X^{Z,\varepsilon}) \longrightarrow (X^W, X^Z),$$

in the sense of finite dimensional distributions on every finite interval. Furthermore, for any $t > 0$

$$\lim_{\varepsilon \to 0} \|X_t^{Z,\varepsilon} \to X_t^Z\|_{L^2} = 0.$$

(b) If furthermore, each G_k satisfies Assumption 2.7, one obtains convergence in $\mathcal{C}^\gamma$ for $\gamma < \frac{1}{2} - \frac{1}{\min_{k\le n} p_k}$.

(c) Suppose Assumption 2.5 holds and further assume that $H \in (\frac{1}{2}, 1)$. Then, on every finite interval and for every $\gamma \in (\frac{1}{3}, \frac{1}{2} - \frac{1}{\min_{k\le n} p_k})$,

$$\left(X_t^\varepsilon, \mathbb{X}_{s,t}^\varepsilon\right) \to (X_t, \mathbb{X}_{s,t} + (t-s)A)$$

weakly in the rough topology $\mathscr{C}^\gamma$, given in (3.11).

We now describe the limit. Set $X = (X^1, \dots, X^n, X^{n+1}, \dots, X^N)$.

(1) $X^W \in \mathbf{R}^n$ and $X^Z \in \mathbf{R}^{N-n}$ are independent.

(2) For $i, j \le n$, $\mathbf{E}\Big(X_t^i X_s^j\Big) = (t \wedge s)A^{i,j}$ where for $\varrho(r) = \mathbf{E}(y_r y_0)$,

$$A^{i,j} = \int_0^\infty \mathbf{E}(G_i(y_s)G_j(y_0))ds = \sum_{q=m_i\vee m_j}^\infty c_{i,q}\, c_{j,q}\, (k!) \int_0^\infty \varrho(r)^q\, dr.$$

In other words, $X^W = U\hat{W}_t$ where $\hat{W}_t$ is a standard Wiener process, U is a square root of A.

(3) Let $Z_t^{H^*(m_k),m_k}$ be the Hermite processes, represented by (3.1). Then,

$$X^Z = (c_{n+1,m_{n+1}} Z_t^{n+1}, \dots, c_{N,m_N} Z_t^N).$$

where

$$Z_t^k = \frac{m_k!}{K(H^*(m_k), m_k)} Z_t^{H^*(m_k),m_k}. \tag{2.9}$$

We emphasize that the Wiener process W_t defining the Hermite processes are the same, for every k, which is in addition independent of $\hat{W}_t$.

(4) For $s < t$ the limiting second order process is given by

$$\mathbb{X}_{s,t}^{i,j} = \int_s^t (X_r^i - X_s^i) dX_r^j, \qquad \begin{array}{ll} \textit{an Itô integral}, & \textit{for } i,j \le n, \\ \textit{a Young integral}, & \textit{otherwise.} \end{array}$$

$$A^{i,j} = \begin{cases} \textit{as in part 2}, & \textit{if } i,j \le n, \\ 0, & \textit{otherwise.} \end{cases}$$

For strong mixing processes, these are well known see [KV86], see also the relatively recent book [KLO12]. A different type of limit theorem for fractional Brownian motions is popular under the topic rough volatility where the parameter H is taken to be close to zero.

We now comment on the proof and give an overview on the results we used. At the level of the convergence of the base processes X^ε, there are a range of results. See for example [DM79, Taq79, BM83, BC09, PT00, BH02, CKM03, BT13, HNX14, NNZ16, CNN18]. Even at this level, the convergence has only been shown in finite dimensional distribution (occasionally in the continuous topology). Even for the convergence in finite dimensional distribution, the known results are fragmented, they are not consistent in the assumptions. They are often proved for a subclass of situations, such as finite chaos condition, moment determinant condition etc., while

Hermite processes are in general not determined by their moments. Some theorems are only proved for scalar processes, some are only at the level of sequences.

We first assemble the convergence of scalar processes, extending them to the same larger class of functions. To extend the convergence to the Hölder topology, we follow [CNN18] and use Malliavin calculus to obtain moment bounds. This section is quite short, only about 3 pages.

For the joint convergence in Hölder norm, we use [BT13, Taq79, BH02]. The joint convergence of a vector valued process with components including a Wiener process and a non semi-martingale is subtle. We use a reduction theorem, a normal convergence theorem from [NP12], and an extension of a limit theorem from [JS03]. We also use the fact that the low Hermite components of X^ε converge in L_2 (this is proven in the Appendix.)

For the functional limit theorems in the rough path topology we first show the joint convergence of the integrals and iterated integrals in finite dimensional distribution. For this we establish a martingale approximation and use ergodic theorems. Due to the fact that we have a non-strong mixing process, proving the L^2 boundedness of the martingale approximations is rather involved, this is where we had to exclude functions with Hermite rank satisfies $H^*(m) \in [0, \frac{1}{2}]$. For the L^2 boundedness we follow [Hai05a, LH19] and develop a locally independent decomposition for the fOU process and use this decomposition to compute the conditional moments. The final hurdle is the relatively compactness of the iterated integrals in the rough path topology, for which we relied on the Diagram formula and an upper bound, from [Taq77], on the number of eligible complete graphs of pairings.

2.3 Single scale model and examples

The following is extracted from section 5, where further detail is given. Restricting ourselves to the one dimensional case we can see how the methodology works without technicalities, here we extend to range of H to $(\frac{1}{3}, 1)$ and drop the exclusilons on m and we obtain the following theorem. Given a centred function $G \in L^2(\mu)$, with chaos expansion $G = \sum_{k=m}^\infty c_k H_k$ we set $c > 0$ by

$$c^2 = \begin{cases} (\frac{c_m m!}{K(H,m)})^2, & H^*(m) > \frac{1}{2} \\ 2\sum_{k=m}^\infty (c_k)^2 k! \int_0^\infty \varrho^k(s) ds, & H^*(m) < \frac{1}{2} \\ 2m!(c_m)^2, & H^*(m) = \frac{1}{2}. \end{cases} \tag{2.10}$$

Theorem C *Let $H \in (\frac{1}{3}, 1) \setminus \{\frac{1}{2}\}$, $f \in \mathcal{C}_b^3(\mathbf{R}; \mathbf{R})$, and G be a continuous function which satisfies Assumption 2.7. Consider*

$$\dot{x}_t^\varepsilon = \alpha(\varepsilon, H^*(m))\, f(x_t^\varepsilon)\, G(y_t^\varepsilon), \qquad x_0^\varepsilon = x_0. \tag{2.11}$$

1. *If $H^*(m) > \frac{1}{2}$, x_t^ε converges weakly in $\mathcal{C}^\gamma$ to the solution to the Young differential equation $d\bar{x}_t = cf(\bar{x}_t)\, dZ_t^{H^*(m),m}$ with initial value x_0 for $\gamma \in (0, H^*(m) - \frac{1}{p})$.*
2. *If $H^*(m) \le \frac{1}{2}$, x_t^ε converges weakly in $\mathcal{C}^\gamma$ to the solution of the Stratonovich stochastic differential equation $d\bar{x}_t = cf(\bar{x}_t) \circ dW_t$ with $\bar{x}_0 = x_0$, where $\gamma \in (0, \frac{1}{2} - \frac{1}{p})$.*

Remark 2.8

1. The constant c could be 0, for further details see Remark 4.4.

2. The condition $\frac{1}{2} - \frac{1}{p} > \frac{1}{3}$ is for the Hölder regularity of the solution paths to be at least $\frac{1}{3}$, so we could define the integral by an enhanced Riemann sum. See §3.3 for the precise meaning.
3. The condition $f \in \mathcal{C}_b^3$ is not optimal, it is only needed for applying the conclusions of Theorem 3.8, the continuity theorem which states that the solutions of a Young/ rough differential equations depends continuously on the driver. For part (1), the Young case, it is sufficient to assume that $f \in \mathcal{C}_b^2$.
4. By using the p-variation norm instead of the Hölder norms, one can reduce the integrabilty condition to $p = 1$ in (1) and $p > 1$ in (2) and obtain convergence in the respective p-variation spaces, see [CFK⁺19] for the use of such norms. The same is true for the other forthcoming theorems, by observing that there is no loss of regularity in the Besov p-variation embeddings compared to the Besov-Hölder ones, see Appendix A from the book [FV10].

We conclude this section with an example on $\mathbf{R}^d$ and a question. Take $H = \frac{8}{9}$ and consider

$$\dot{x}_t^\varepsilon = \varepsilon^{-\frac{2}{9}} f_1(x_t^\varepsilon) H_2(y_{\frac{t}{\varepsilon}}) + \varepsilon^{-\frac{4}{9}} f_2(x_t^\varepsilon) H_4(y_{\frac{t}{\varepsilon}}) + \frac{1}{\sqrt{\varepsilon}} f_3(x_t^\varepsilon) H_{10}(y_{\frac{t}{\varepsilon}}).$$

Their solutions converge to that of an equation driven simultaneously by a fBM, a Hermite process with similarity exponents $\frac{5}{9}$, and a Wiener process.

3 Preliminaries

A stochastic process (X_n) is strongly mixing if its auto correlation $\varrho(n) \to 0$ as $n \to \infty$ and

$$\sup_{A,B} |P(A \cap B) - P(A)P(B)| \le \varrho(n),$$

where the supremum is over $A \in \sigma\{X_k, k \le m\}$, $B \in \sigma\{X_k, k \ge n+m\}$.

Let us make the convention that $B_0^H = 0$, $\mathbf{E}(B_1^H)^2 = 1$. For simplicity we often omit the Hurst parameter H. A fBM B_t^H, with hurst parameter $H > \frac{1}{2}$ is not strong mixing, otherwise the Central Limit Theorem holds for $X_n = B_{n+1}^H - B_n^H$, but with suitable scaling, the limit yields a fractional Brownian motion.

The disjoint increments of B_t^H are dependent unless $H = \frac{1}{2}$:

$$\mathbf{E}(B_t - B_s)(B_u - B_v) = \frac{1}{2}\left(|t-v|^{2H} + |s-u|^{2H} - |t-u|^{2H} - |s-v|^{2H}\right).$$

The correlation function

$$\begin{aligned} \tilde{\varrho}(n) &= \mathbf{E}(B_{n+1} - B_n)(B_1 - B_0) \\ &\sim H(2H-1)n^{2H-2}, \qquad \text{at infinity.} \end{aligned}$$

Thus, $\sum_{n=1}^\infty \tilde{\varrho}(n) = \infty$ for $H > \frac{1}{2}$, and so long range correlation persists.

Common to a Brownian motion, a fBM has self-similarity with similarity exponent H and stationary increments. Since B_t^H has finite and non-trivial $\frac{1}{H}$-variation over $[0,T]$ with variation of the order $\mathbf{E}\left(|B_1^H|^p\right)T$, it has infinite total variation. It has zero quadratic variation for $H > \frac{1}{2}$ and infinite quadratic variation for $H < \frac{1}{2}$, and therefore B_t^H is not a semi-martingale unless $H = \frac{1}{2}$. We refer to [PT17, Sam06, CKM03] for further detail.

3.1 Hermite processes

We take the Hermite polynomials of degree m to be

$$H_m(x) = (-1)^m e^{\frac{x^2}{2}} \frac{d^m}{dx^m} e^{\frac{x^2}{2}}.$$

Thus, $H_0(x) = 1$, $H_1(x) = x$. Let $\hat{H}$ be the inverse of $H^*(m) = m(H-1)+1$:

$$\hat{H}(m) = \frac{1}{m}(H-1)+1.$$

Definition 3.1 Let $m \in \mathbf{N}$ with $\hat{H}(m) > \frac{1}{2}$. The class of *Hermite processes* of rank m are the following mean-zero process,

$$Z_t^{H,m} = \frac{K(H,m)}{m!} \int_0^t H_m\left(\int_{\mathbf{R}} (s-u)_+^{\hat{H}(m)-\frac{3}{2}} dW_u \right) ds, \tag{3.1}$$

where the constant $K(H,m)$ is chosen so their variances are 1 at $t=1$. The number H is called its Hurst parameter.

Since $\hat{H}(1) = H$, the rank 1 Hermite processes $Z^{H,1}$ are fractional BMs. Indeed (3.1) is exactly the Mandelbrot-Vanness representation for fBM. The Hermite processes have stationary increments and finite moments of all order with covariance

$$\mathbf{E}(Z_t^{H,m} Z_s^{H,m}) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}). \tag{3.2}$$

Therefore, using Kolmogorv's theorem one can show that the Hermite processes $Z_t^{H,m}$ have sample paths of Hölder regularity up to H. They are also self similar with exponent H

$$\lambda^H Z_{\frac{\cdot}{\lambda}}^{H,m} \sim Z_{\cdot}^{H,m}.$$

We recall another formulation, useful for proving convergence to Hermite processes. According to Itô [Itô51] and [Nua06, Thm1.1.2], if f is an L^2 function of norm 1, the multiple Itô-Wiener integral with kernel $\prod_i f(t_i)$ can be identified with the evaluation of H_m on a single Wiener integral:

$$\int_{\mathbf{R}} \cdots \int_{\mathbf{R}} f(t_1) \dots f(t_m)\, dW(t_1) \dots dW(t_m) = H_m\left(\int_{\mathbf{R}} f(s)\, dW_s \right),$$

So the Hermite processes can be defined by the multiple Itô-Wiener integrals:

$$Z_t^{H,m} = \frac{K(H,m)}{m!} \int_{\mathbf{R}^m} \left(\int_0^t \prod_{j=1}^m (s-t_j)_+^{-(\frac{1}{2}+\frac{1-H}{m})} ds \right) dW(t_1) \dots dW(t_m). \tag{3.3}$$

In particular, two Hermite processes $Z^{H,m}$ and $Z^{H',m'}$, defined by the same Wiener process, are uncorrelated if $m \neq m'$.

Remark 3.2 In some literature, e.g. [MT07], the Hermite processes are defined with a different exponent as below:

$$\tilde{Z}_t^{H,m} = \frac{K(H^*(m),m)}{m!} \int_0^t H_m\left(\int_{\mathbf{R}} (s-u)_+^{H-\frac{3}{2}} dW_u \right) ds.$$

They are related by

$$Z_t^{H^*(m),m} = \tilde{Z}_t^{H,m}, \qquad Z_t^{H,m} = \tilde{Z}_t^{\hat{H}(m),m}. \tag{3.4}$$

Further detail on Hermite processes can also be found in [MT07]. The limit processes in Theorem A are given by Hermite processes of the form $Z_t^{H^*(m),m} = \tilde{Z}_t^{H,m}$.

3.2 Fractional Ornstein-Uhlenbeck processes

We define the stationary fractional Ornstein-Uhlenbeck processes to be

$$y_t = \sigma \int_{-\infty}^t e^{-(t-s)} dB_s,$$

where σ is chosen such that y_t is distributed as $\mu = N(0,1)$ and B_t is a two-sided fractional BM. It is the solution of the following Langevin equation:

$$dy_t = -y_t dt + \sigma dB_t, \qquad y_0 = \sigma \int_{-\infty}^0 e^s dB_s.$$

We take y_t^ε, the fast fOU, to be the stationary solution of

$$dy_t^\varepsilon = -\frac{1}{\varepsilon} \lambda y_t^\varepsilon \, dt + \frac{\sigma}{\varepsilon^H} \, dB_t.$$

Observe that $y_\cdot^\varepsilon$ and $y_{\frac{\cdot}{\varepsilon}}$ have the same distributions, and

$$y_t^\varepsilon = \frac{\sigma}{\varepsilon^H} \int_{-\infty}^t e^{-\frac{1}{\varepsilon}(t-s)} dB_s. \tag{3.5}$$

Let us denote their auto-correlation function by ϱ and ϱ^ε:

$$\varrho(s,t) := \mathbf{E}(y_s y_t), \qquad \varrho^\varepsilon(s,t) := \mathbf{E}(y_s^\varepsilon y_t^\varepsilon)$$

Let $\varrho(s) = \mathbf{E}(y_0 y_s)$ for $s \geq 0$ and extended to $\mathbf{R}$ by symmetry, then $\varrho(s,t) = \varrho(t-s)$ and similarly for ϱ^ε. For $H > \frac{1}{2}$, the set of functions for which Wiener integrals are defined include L^2 functions and so ϱ has a nicer expression.

Indeed, since

$$\mathbf{E}(B_t B_s) = H(2H-1) \int_0^t \int_0^s |r_1 - r_2|^{2H-2} dr_1 dr_2,$$

we have

$$\frac{\partial^2}{\partial t \partial s} \mathbf{E}(B_t B_s) = H(2H-1)|t-s|^{2H-2},$$

which is integrable, and therefore we may use the Wiener isometry

$$\mathbf{E}(y_t y_s) = \sigma^2 H(2H-1) \int_{-\infty}^t \int_{-\infty}^s e^{-(s+t-r_1-r_2)} |r_1 - r_2|^{2H-2} dr_1 dr_2.$$

For $u > 0$,

$$\varrho(u) = \sigma^2 H(2H-1) \int_{-\infty}^u \int_{-\infty}^0 e^{-(u-r_1-r_2)} |r_1 - r_2|^{2H-2} dr_1 dr_2.$$

With this we observe the following correlation decay relation.

Lemma 3.3 *Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$. For any $t \neq s$,*

$$|\varrho(s,t)| \lesssim 1 \wedge |t-s|^{2H-2}. \tag{3.6}$$

The proof for this is elementary, for reader's convenience it is given in the appendix.

By Lemma 3.3, $\int_0^\infty \varrho^m(s) ds$ is finite if $H^*(m) < \frac{1}{2}$, or if $H = \frac{1}{2}$ and $m \in \mathbf{N}$, as in the latter the usual OU process admits exponential decay of correlations.

Lemma 3.4 *Let $H \in (0,1) \setminus \{\frac{1}{2}\}$, fix a finite time horizon T, then for $t \in [0,T]$ the following holds uniformly for $\varepsilon \in (0, \frac{1}{2}]$:*

$$\left(\int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} |\varrho(u,r)|^m \, dr \, du \right)^{\frac{1}{2}} \lesssim \begin{cases} \sqrt{\frac{t}{\varepsilon} \int_0^\infty \varrho^m(s) ds}, & \text{if} \quad H^*(m) < \frac{1}{2}, \\ \sqrt{(\frac{t}{\varepsilon}) |\ln\left(\frac{1}{\varepsilon}\right)|}, & \text{if} \quad H^*(m) = \frac{1}{2}, \\ \left(\frac{t}{\varepsilon}\right)^{H^*(m)}, & \text{if} \quad H^*(m) > \frac{1}{2}. \end{cases} \tag{3.7}$$

$$\left(\int_0^t \int_0^t |\varrho^\varepsilon(u,r)|^m \, dr \, du \right)^{\frac{1}{2}} \lesssim \begin{cases} \sqrt{t\varepsilon \int_0^\infty \varrho^m(s) ds}, & \text{if} \quad H^*(m) < \frac{1}{2}, \\ \sqrt{t\varepsilon |\ln\left(\frac{t}{\varepsilon}\right)|}, & \text{if} \quad H^*(m) = \frac{1}{2}, \\ \left(\frac{t}{\varepsilon}\right)^{H^*(m)-1}, & \text{if} \quad H^*(m) > \frac{1}{2}. \end{cases} \tag{3.8}$$

Note if $H = \frac{1}{2}$, for and any $m \in \mathbf{N}$, the bound is $\sqrt{\frac{t}{\varepsilon} \int_0^\infty \varrho^m(s) ds}$. The following is often used later,

$$t \int_0^t |\varrho^\varepsilon(s)|^m ds \lesssim \frac{t^{\left(H^*(m) \vee \frac{1}{2}\right)}}{\alpha(\varepsilon, H^*(m))}. \tag{3.9}$$

The following Hölder norm estimates will be used for proving Proposition 3.12.

Lemma 3.5 *The stationary fOU process is uniformly Hölder continuous of order γ over $[0,\infty)$ for any $\gamma \in (0,H)$. Furthermore, over $[0,\infty)$, the following estimates hold:*

$$\|y_s - y_r\|_{L^p} \lesssim 1 \wedge |s-r|^H, \qquad \mathbf{E} \sup_{s \neq t} \left(\frac{|y_s - y_t|}{|t-s|^\gamma} \right)^p \lesssim C(\gamma,p)^p M$$

for any $p > 1$, where $C(\gamma,p)$ is the universal constant in Garcia-Rodemich-Romsey-Kolmogorov inequality and

$$M = \int_0^\infty \int_0^\infty \frac{\mathbf{E}|y_s - y_r|^p}{|s-r|^{\gamma p + 2}} < \infty.$$

3.3 Some rough path theory

If X and Y are Hölder continuous functions on $[0,T]$ with exponent α and β respectively, such that $\alpha+\beta>1$, the Young integration theory enables us to define $\int_0^T Y dX$ via Riemann sums $\sum_{[u,v]\in\mathcal{P}} Y_u(X_v-X_u)$, where $\mathcal{P}$ denotes a partition of $[0,T]$. Furthermore $(X,Y)\mapsto \int_0^T Y dX$ is a continuous map. Thus, for $X\in\mathcal{C}^{\frac{1}{2}+}$, one can make sense of a solution Y to the Young integral equation $dY_s=f(Y_s)dX_s$. If $f\in\mathcal{C}_b^2$, the solution is continuous with respect to both the driver X and the initial data [You36]. In the case of X having Hölder continuity less or equal to $\frac{1}{2}$, this fails and one can not define a pathwise integration by the above Riemann sum anymore. Rough path theory provides us with a machinery to treat less regular functions by enhancing the process with a second order process, giving a better local approximation, which then can be used to enhance the Riemann sum and show it converges. If X_s is a Brownian motion, taking the dyadic approximation then the usual Riemann sum leads to convergent in probability to Itô integrals; but the enhanced Riemann sum provides better approximations and defines a pathwise integral agreeing with the Itô integral provided the integrand belongs to both domains of integration. Their domains of integration are quite different, the first uses an additional adaptedness condition and requires arguably less regularity than the second.

We restrict ourselves to the case where X_t is a continuous path over $[0,T]$, for now we assume it takes values in $\mathbf{R}^d$. A rough path of regularity $\alpha\in(\frac{1}{3},\frac{1}{2})$, is a pair of processes $\mathbf{X}=(X_t,\mathbb{X}_{s,t})$ where $(\mathbb{X}_{s,t})\in\mathbf{R}^{d\times d}$ is a two parameter stochastic processes satisfying the following algebraic conditions: for $0\le s<u<t\le T$,

$$\mathbb{X}_{s,t}-\mathbb{X}_{s,u}-\mathbb{X}_{u,t}=X_{s,u}\otimes X_{u,t}, \qquad \text{(Chen's relation)}$$

where $X_{s,t}=X_t-X_s$, and $(X_{s,u}\otimes X_{u,t})^{i,j}=X^i_{s,u}X^j_{u,t}$ as well as the following analytic conditions,

$$\|X_{s,t}\|\lesssim|t-s|^\alpha, \qquad \|\mathbb{X}_{s,t}\|\lesssim|t-s|^{2\alpha}. \tag{3.10}$$

The set of such paths will be denoted by $\mathscr{C}^\alpha([0,T];\mathbf{R}^d)$. The so called second order process $\mathbb{X}_{s,t}$ can be viewed as a possible candidate for the iterated integrals $\int_s^t X_{s,u}dX_u$.

Remark 3.6 Using Chen's relation for $s=0$ one obtains

$$\mathbb{X}_{u,t}=\mathbb{X}_{0,t}-\mathbb{X}_{0,u}-X_{0,u}\otimes X_{u,t},$$

thus one can reconstruct $\mathbb{X}$ by knowing the path $t\to(X_{0,t},\mathbb{X}_{0,t})$.

Given a path X, which is regular enough to define its iterated integral, for example $X\in\mathcal{C}^1([0,T];\mathbf{R}^d)$, we define its natural rough path lift to be given by

$$\mathbb{X}_{s,t}:=\int_s^t X_{s,u}dX_u.$$

It is now an easy exercise to verify that $\mathbf{X}=(X,\mathbb{X})$ satisfies the algebraic and analytic conditions (depending on the regularity of X), by which we mean Chen's relation and (3.10). Note that given any function $F\in\mathcal{C}^{2\alpha}(\mathbf{R}^{d\times d})$, setting $\tilde{\mathbb{X}}_{s,t}=\mathbb{X}_{s,t}+F_t-F_s$, $\tilde{\mathbb{X}}$ would also be a possible choice for the rough path lift. Given two rough paths $\mathbf{X}$ and $\mathbf{Y}$ we define their distance to be

$$\varrho_\alpha(\mathbf{X},\mathbf{Y})=\sup_{s\neq t}\frac{\|X_{s,t}-Y_{s,t}\|}{|t-s|^\alpha}+\sup_{s\neq t}\frac{\|\mathbb{X}_{s,t}-\mathbb{Y}_{s,t}\|}{|t-s|^{2\alpha}} \tag{3.11}$$

This defines a complete metric on $\mathscr{C}^\alpha([0,T];\mathbf{R}^d)$, this is called the inhomogenous α-Hölder rough path metric. We are also going to make use of the norm like object

$$\|\mathbf{X}\|_\alpha = \sup_{s\neq t\in[0,T]} \frac{\|X_{s,t}\|}{|t-s|^\alpha} + \sup_{s\neq t\in[0,T]} \frac{\|\mathbb{X}_{s,t}\|^{\frac{1}{2}}}{|t-s|^\alpha} \tag{3.12}$$

We also denote for any two parameter process $\mathbb{X}$ a semi-norm:

$$\|\mathbb{X}\|_{2\alpha} := \sup_{s\neq t\in[0,T]} \frac{\|\mathbb{X}_{s,t}\|}{|t-s|^{2\alpha}}.$$

Given a path X, as the second order process $\mathbb{X}$ takes the role of an iterated integral, another sensible conditions to impose is the chain rule (or integration by parts formulae) leading to the following definition.

Definition 3.7 A rough path $\mathbf{X}$ satisfying the following condition,

$$Sym(\mathbb{X}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t} \tag{3.13}$$

is called a geometric rough path. The space of all of geometric rough paths of regularity α is denoted by $\mathscr{C}_g^\alpha([0,T];\mathbf{R}^d)$ and forms a closed subspace of $\mathscr{C}^\alpha([0,T];\mathbf{R}^d)$.

Furthermore, one can show that if a sequence of $\mathcal{C}^1([0,T],\mathbf{R}^d)$ paths X_n converges in the rough path metric to $\mathbf{X}$, then $\mathbf{X}$ is a geometric rough path. To obtain a geometric rough path from a Wiener process, as $\int_0^t W_s \circ dW_s = \frac{W_t^2}{2}$, one has to enhance it with its Stratonovich integral, $\mathbb{W}_{s,t} = \int_s^t (W_r - W_s) \circ dW_r$.

Given a rough path $\mathbf{X} \in \mathscr{C}^\alpha([0,T];\mathbf{R}^d)$, we may define the integral $\int_0^T Y\, d\mathbf{X}$ for suitable paths $Y \in \mathcal{C}^\alpha([0,T],\mathbb{L}(\mathbf{R}^d,\mathbf{R}^m))$, which admit a Gubinelli derivative $Y' \in \mathcal{C}^\alpha([0,T],\mathbb{L}(\mathbf{R}^{d\times d},\mathbf{R}^m))$ with respect to $\mathbf{X}$, meaning

$$Y_{s,t} = Y'_s X_{s,t} + R_{s,t},$$

and the two parameter function R, satisfies $\|R\|_{2\alpha} < \infty$. The pair $\mathbf{Y} := (Y, Y')$ is said to be a controlled rough path, their collection is denoted by $\mathcal{D}_X^{2\alpha}$. The remainder term for the case $Y = f(X)$ with f smooth is the remainder term in the Taylor expansion. This is done by showing that the enhanced Riemann sums

$$\sum_{[s,t]\in\mathcal{P}} Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t},$$

converge as the partition size is going to zero, and the limit is defined to be $\int \mathbf{Y}\, d\mathbf{X}$. With this theory of integration one can study the equation,

$$dY = f(Y) d\mathbf{X}.$$

Unlike in the theory of stochastic differential equations one now has continuous dependence on the noise $\mathbf{X}$. This can be interpreted as the following, once one has chosen a candidate for the iterated integral, the second order process, solving rough differential equations is a continuous operation, but obtaining the second order process is not, see the introduction to [FH14].

Now given $Y \in \mathcal{D}_X^{2\alpha}$, then $(\int \mathbf{Y}\, d\mathbf{X}, Y) \in \mathcal{D}_X^{2\alpha}$, and the map $(\mathbf{X}, \mathbf{Y}) \mapsto (\int \mathbf{Y}\, d\mathbf{X}, Y)$ is continuous with respect to $\mathbf{X} \in \mathscr{C}^\alpha$ and $Y \in \mathcal{D}_X^{2\alpha}$. The domain is not a product space, the continuity is best formulated with the appropriate distances, however, we do not need the precise formulation, for details we refer to [FH14]. We now state the precise theorem for our application, see also [Lyo94].

Theorem 3.8 *[FH14] Let $Y_0 \in \mathbf{R}^m$, $\beta \in (\frac{1}{3}, 1)$, $f \in \mathcal{C}_b^3(\mathbf{R}^m; \mathbb{L}(\mathbf{R}^d; \mathbf{R}^m))$ and $\mathbf{X} \in \mathscr{C}^\beta([0,T], \mathbf{R}^d)$. Then the differential equation*

$$Y_t = Y_0 + \int_0^t f(Y_s) d\mathrm{X}_s \tag{3.14}$$

has a unique solution which belongs to $\mathcal{C}^\beta$. Furthermore, the solution map $\Phi_f : \mathbf{R}^d \times \mathscr{C}^\beta([0,T], \mathbf{R}^d) \to \mathcal{D}_X^{2\beta}([0,T], \mathbf{R}^m)$, where the first component is the initial condition and the second component the driver, is continuous.

As continuous maps preserve weak convergence to show weak convergence of solutions to rough differential equations

$$dY^\varepsilon = f(Y^\varepsilon) d\mathbf{X}^\varepsilon,$$

it is enough to establish weak convergence of the rough paths $\mathbf{X}^\varepsilon$ in the topology defined by the rough metric. Obtaining convergence in this topology follows the convergence of the finite dimensional distributions of the rough paths $\mathbf{X}^\varepsilon$ plus tightness in the space of rough paths with respect to that topology. To apply this theory, we will enhance our stochastic processes, c.f. Proposition 3.12, Proposition 4.8 and the following proof of Theorem C and Section 5 to bring it to this framework.

3.3.1 Tightness of rough paths

We now show that moment bounds on increments lead to tightness in the rough path topologies. The following lemma is similar to the compact embedding theorems between Hölder spaces and can be obtained via an Arzela-Ascoli argument.

Lemma 3.9 *Let 0 denote the rough path obtained from the 0 function enhanced with a 0 function, then for $\gamma > \gamma'$, the sets $\{\mathbf{X} \in \mathscr{C}^{\gamma'} : \varrho_\gamma(\mathrm{X}, 0) < R, \mathbf{X}(0) = 0\}$ are compact in $\mathscr{C}^{\gamma'}$.*

The next lemma relates uniform moment bounds on the increments of the the stochastic processes and their secondary process, to uniform moment bounds on the rough path norm.

Lemma 3.10 *Let $\theta \in (0,1)$, $\gamma \in (0, \theta - \frac{1}{p})$ and $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon)$ such that*

$$\|X_{s,t}^\varepsilon\|_{L^p} \lesssim |t-s|^\theta, \qquad \|\mathbb{X}_{s,t}^\varepsilon\|_{L^{\frac{p}{2}}} \lesssim |t-s|^{2\theta},$$

then

$$\sup_{\varepsilon \in (0,1]} \mathbf{E}(\|\mathbf{X}^\varepsilon\|_\gamma)^p < \infty$$

Proof. The proof is based on a Besov-Hölder embedding, for details we refer to [FV10, CFK⁺19]. □

Lemma 3.11 *Let $\mathbf{X}^\varepsilon$ be a sequence of rough paths and $\gamma \in (\frac{1}{3}, \frac{1}{2} - \frac{1}{p})$, such that $\mathbf{X}(0) = 0$, and*

$$\sup_{\varepsilon\in(0,1]} \mathbf{E}(\|\mathbf{X}^\varepsilon\|_\gamma)^p < \infty,$$

then $\mathbf{X}^\varepsilon$ is tight in $\mathscr{C}^\gamma$.

Proof. Choose $\gamma' \in (\frac{1}{3}, \gamma)$, as $\varrho_{\gamma'}(\mathbf{X}, 0) \le \|\mathbf{X}\|_{\gamma'} + \|\mathbf{X}\|^2_{\gamma'}$ we obtain

$$\mathbf{P}\big(\varrho_{\gamma'}(\mathbf{X}^\varepsilon, 0) > R\big) \le \frac{\mathbf{E}\big(\varrho_{\gamma'}(\mathbf{X}^\varepsilon, 0)\big)^{\frac{p}{2}}}{R^{\frac{p}{2}}} \le \frac{\mathbf{E}\Big(\|\mathbf{X}\|_{\gamma'} + \|\mathbf{X}\|^2_{\gamma'}\Big)^{\frac{p}{2}}}{R^{\frac{p}{2}}} \lesssim \frac{C}{R^{\frac{p}{2}}}.$$

This proves the claim by Lemma 3.9. □

3.4 Example: linear driver and kinetic fBM

Here we consider the toy model on **R**,

$$\begin{cases} \dot{x}^\varepsilon_t = \varepsilon^{H-1} f(x^\varepsilon_t)\, y^\varepsilon_t, \\ x^\varepsilon_0 = x_0. \end{cases} \tag{3.15}$$

We study this without using any of the complicated estimates obtained later, nevertheless we are already able to explain our methodology.

Proposition 3.12

(a) *Let $H \in (0,1)$, $\gamma \in (0, H)$, $p > 1$ and fix a finite time T. Let $X^\varepsilon_t = \varepsilon^{H-1} \int_0^t y^\varepsilon_s ds$, then, for $s, t \in [0, T]$,*

$$\sup_{s,t\in[0,T]} \big\|X^\varepsilon_{s,t} - B_{s,t}\big\|_{L^p} \lesssim \varepsilon^H, \qquad \Big\| |X^\varepsilon - B|_{\mathcal{C}^{\gamma'}([0,t]}\Big\|_{L^p} \lesssim t^\gamma \varepsilon^{H-\gamma},$$

for any $\gamma' < \gamma < H$.

(b) *Let $H \in (\frac{1}{3}, 1)$ and $f \in \mathcal{C}^3_b$. Then for any $\gamma \in (0, H)$, x^ε_t converges in L^p in $\mathcal{C}^{\gamma'}([0,T]); \mathbf{R}^d)$ to the solution of the rough differential equation:*

$$\dot{x}_t = f(x_t)\, dB_t, \tag{3.16}$$

furthermore, for $t \in [0, T]$,

$$\Big\| |x^\varepsilon - x|_{C^{\gamma'}([0,t])} \Big\|_{L^p} \lesssim t^\gamma \varepsilon^{H-\gamma}.$$

Proof. (a) Set $v^\varepsilon_t = \varepsilon^{H-1} y^\varepsilon_t$, then v^ε_t solves the following equation

$$dv^\varepsilon_t = -\frac{1}{\varepsilon} v^\varepsilon_t dt + \frac{1}{\varepsilon} dB_t$$

Using the equation for v_t^ε we have

$$X_{s,t}^\varepsilon = \varepsilon^{H-1}\int_s^t y_r^\varepsilon dr = \int_s^t v_r^\varepsilon dr = \varepsilon(v_s^\varepsilon - v_t^\varepsilon) + B_{s,t}.$$

Therefore, for any $p > 1$,

$$\begin{aligned}\sup_{s,t\in[0,T]} \left\|X_{s,t}^\varepsilon - B_{s,t}\right\|_{L^p} &= \sup_{s,t\in[0,T]} \|\varepsilon(v_t^\varepsilon - v_s^\varepsilon)\|_{L^p} \\ &= \varepsilon^H \sup_{s,t\in[0,T]} \|y_t^\varepsilon - y_s^\varepsilon\|_{L^p} \lesssim \varepsilon^H.\end{aligned}$$

In the last step we used the stationarity of y_t^ε, which follows from that of y_t. By Lemma 3.5, we have the following estimates on their Hölder norms for any $t \in [0, T]$:

$$\| |X^\varepsilon - B|_{\mathcal{C}^{\gamma'}([0,t])} \|_{L^p} \lesssim \varepsilon^H \left(\frac{t}{\varepsilon}\right)^\gamma,$$

this holds for any $p > 1$ and any $\gamma' < \gamma$, and part (a) follows.

(b) The system of equations is clearly well posed and has global solutions. The idea is to consider the equation as a differential equation driven by rough paths $\mathbf{X}^\varepsilon$ as below:

$$\dot{x}_t^\varepsilon = f(x_t^\varepsilon) d\mathbf{X}_t^\varepsilon$$

For $H \in (\frac{1}{2}, 1)$, the integral is simply the Young integral, and Young's continuity theorem states that x^ε converges provided $\mathbf{X}^\varepsilon$ converges. For $H \in (\frac{1}{3}, \frac{1}{2})$ it is only left to deal with the geometric rough path lift of x^ε, which in dimension one has only symmetric part and so the convergence in the $\mathcal{C}^\gamma$ topology is the same as convergence in the rough path topology.

For the L^p convergence, we start with the L^p convergence of the drivers X^ε. Since X_t^ε is in $\mathcal{C}^\gamma$ for some $\gamma > \frac{1}{3}$, The solutions map Φ for this equation is Lipschitz continuous: for any $\gamma' < \gamma$,

$$|\Phi(X^\varepsilon) - \Phi(B)|_{\gamma'} \lesssim |X^\varepsilon - B|_\gamma.$$

This shows the convergence in L^p of the solutions. □

Remark 3.13 From part (a) of the theorem we deduce that the solution to the equation

$$\dot{z}_t^\varepsilon = v_t^\varepsilon, \quad dv_t^\varepsilon = -\frac{1}{\varepsilon} v_t^\varepsilon dt + \frac{1}{\varepsilon} dB_t^H, \quad z_0^\varepsilon \to z_0, \quad v_0^\varepsilon = \varepsilon^{H-1} y_0, \tag{3.17}$$

converges in $\mathcal{C}^\gamma$ weakly to a fractional Brownian motion $z_0 + \tilde{B}_t$. Krammer-Smoluchowski limits /Kinetic fBM's are studied in [BT05, Zha08, ATH12]. See also [FK00, FGL15, FH14]. Here y_0 is distributed as μ. But our initial condition is not optimal. Since the solution depends on the initial condition affinely, stronger scaling and y_0 should also work.

Remark 3.14 We explain this with the previous example. In one dimension the canonical lift of a process is a function of the process itself:

$$\int_0^t X_s^\varepsilon dX_s^\varepsilon = \varepsilon^{2H-2} \int_0^t \int_0^s y_s^\varepsilon y_r^\varepsilon dr ds = \frac{1}{2}(X_t^\varepsilon)^2 \to \frac{1}{2}(B_t)^2.$$

This is because taking squares is a continuous operation in $\mathcal{C}^\alpha$, the convergence of $(X_t^\varepsilon, \int_0^t X_s^\varepsilon dX_s^\varepsilon)$ follows. By part (1), X_t^ε is uniformly bounded in $\mathscr{C}^\gamma$ for any $\gamma < H$. Since we can choose $\gamma > \frac{1}{3}$, to show convergence in the rough path topologies we only need to establish the moment bounds of the iterated integrals. In our case, it follows from Proposition 3.12:

$$\begin{aligned}
\left\| \int_v^t (X_s^\varepsilon - X_v^\varepsilon) dX_s^\varepsilon \right\|_{L^p} &= \frac{1}{2} \left\| \varepsilon^{2H-2} \int_v^t \int_v^s y_s^\varepsilon y_r^\varepsilon dr ds \right\|_{L^p} \\
&\le \frac{1}{2} \left(\left\| \varepsilon^{H-1} \int_v^t y_s^\varepsilon ds \right\|_{L^{2p}} \right)^{\frac{1}{p}} \\
&\lesssim |t-v|^{2H}.
\end{aligned}$$

The general G case will be discussed in section 4.3, before that we will make use of Malliavin calculus to obtain the L^p estimates and a suitable central/functional limit theorem. We then have to explore different scaling constants, since the limits constitutes of components not necessarily simultaneously in the same universality classes. If the equation involves two functions G_i with different decay rates, to which different methods were used for the convergence, we work harder for their joint convergence and relatively compactness in a topology suitable for all the limiting classes.

This means that X^ε, with its canonical lift to the step-2 rough path space, is tight in the rough path topology $\mathscr{C}^\gamma$ for any $\gamma < H$.

4 Enhanced functional limit theorem in 1-d

We first prove a central limit theorem for functionals of the fractional Ornstein-Uhlenbck processes, with convergence in finite dimensional distributions. Then we go ahead and establish moment bounds on the increments of our process to conclude convergence in a suitable space of rough paths. This enables us to use the continuity of the solution maps to Young rough differential equations to prove our single scale homogenization result.

4.1 Convergence in finite dimensional distributions

The scaling limit for path integrals of functionals of the fOU can be either in the Gaussian or in the non-Gaussian universality classes. Traditionally the first ones are called CLT's and the latter non-CLT's. We would call both CLT's.

4.1.1 CLT and non CLT for sequences

The intuition for scalar valued CLT's comes from its counter part for sequences which we explain in the next paragraph. If Y_n is a mean zero, stationary, and strong mixing sequence, such that

$$\sigma_n^2 = \mathbf{E}(\sum_{i=1}^n Y_i)^2 \to \infty, \qquad \mathbf{E}(\sum_{i=1}^n Y_i)^4 = O(\sigma_n^4),$$

then the CLT holds:

$$\frac{1}{\sigma_n} \sum_{i=1}^n Y_i \longrightarrow N(0,1).$$

If Y_n is not strong mixing, the CLT may fail. An example of which are Gaussian sequences with slow decaying correlation functions. A guiding principle, for Gaussian sequences, can be found in [BM83] for short range correlations, and in [Taq79, BT13] for long range correlations. A simple version is as follows.

Let X_n be a sequence of stationary mean zero variance 1 Gaussian random variables with auto-correlation $n^{-\gamma}$ with $\gamma \in (0,1)$. Let G be a function with Hermite rank $m \geq 1$ and $A(n)$ a sequence such that

$$\lim_{n\to\infty} \mathrm{var}\left(\frac{1}{A(n)} \sum_{k=1}^{n} G(X_k)\right) = 1.$$

1. If $\gamma \in (\frac{1}{m}, 1)$, then the theorems following holds in finite dimensional distributions,

$$\frac{1}{A(n)} \sum_{k=1}^{[nt]} G(X_n) \to W(t).$$

2. If $\gamma \in (0, \frac{1}{m})$, then the scaling limit is a Hermite process in the m-th chaos.
3. If $\gamma = \frac{1}{m}$, then, the scaling limit is also a Wiener process.

The scaling constant is of the order $n^{1-\frac{1}{2}\gamma m}$ in the second case, of order $\sqrt{n}$ for the first case, and of order $\sqrt{n \ln n}$ for (3). From this the continuous version CLT for $\gamma \in (0, \frac{1}{m})$ was obtained in [BH02]. The borderline case $\gamma = \frac{1}{m}$ for the continuous version was analysed in [BC09].

4.1.2 Functional CLT in finite dimensional distributions

If y_t is any stationary Gaussian process with correlation function ϱ, then for $m \geq 1$,

$$\mathbf{E}(H_n(y_t) H_m(y_s)) = \delta_{n,m}(\varrho(s,t))^m.$$

Since

$$\sqrt{\mathbf{E}\left(\int_0^t H_m(y_s^\varepsilon) ds\right)^2} = \left(\int_0^t \int_0^t \varrho^\varepsilon(|s-r|)^m dr ds\right)^{\frac{1}{2}},$$

composing this with Lemma 3.4, we therefore expect that the correct scaling to be $\frac{1}{\sqrt{\varepsilon}}$ for the case $H^*(m) < \frac{1}{2}$; $\sqrt{\frac{1}{\varepsilon |\ln \varepsilon|}}$ for the case $H^*(m) = \frac{1}{2}$; and $\varepsilon^{H^*(m)-1}$ otherwise. Observing that

$$\alpha(\varepsilon, H^*(m)) \int_0^t H_m(y_s^\varepsilon) ds \sim \varepsilon \alpha(\varepsilon, H^*(m)) \int_0^{\frac{t}{\varepsilon}} H_m(y_s) ds.$$

This suggest that the self-similarity of the limiting process are determined by $\alpha(\varepsilon, H^*(m))$. In the first two cases the limit will be a Wiener process, and in the later one the limit Z_t should have the scaling property:

$$\varepsilon^{H^*(m)} Z_{\frac{t}{\varepsilon}} \sim Z_t.$$

These limits turn out indeed to be the Hermite processes.

We first consider G with low Hermite rank: $H^*(m) > \frac{1}{2}$. Since $m \geq 1$ this restricts to the case $H > \frac{1}{2}$.

Lemma 4.1 *Let $G = \sum_{k=m}^{\infty} c_k H_k$ where $m > 0$ be in $L^2(\mu)$. Let $H \in (\frac{1}{2}, 1)$. Then the following statements hold for the stationary fOU y_t^ε. If $H^*(m) > \frac{1}{2}$ then*

$$\left\| \varepsilon^{H^*(m)-1} \int_0^t G(y_s^\varepsilon) ds - \frac{c_m m!}{K(H^*(m), m)} Z_t^{H^*(m),m} \right\|_{L_2(\Omega)} \to 0.$$

Proof. This result looks slightly mysterious which can be explained easily by kernel convergence, since $H^*(m)$ decreases with m, it is sufficient to work with H_m for m the Hermite rank of G, c.f. Equation (4.3). The key idea is to write a Wiener integral representations for these integrals beginning with

$$\begin{aligned} y_t^\varepsilon &= \varepsilon^{-H} \int_{-\infty}^t e^{-\frac{t-r}{\varepsilon}} dB_r = \int_{\mathbf{R}} h_\varepsilon(t,s) dW_s, \quad \text{where} \\ h_\varepsilon(t,s) &= \varepsilon^{-\frac{1}{2}} \frac{1}{c_1(H)} e^{-\frac{t-s}{\varepsilon}} \int_0^{\frac{t-s}{\varepsilon}} e^v v_+^{H-\frac{3}{2}} \, dv, \end{aligned} \tag{4.1}$$

and $c_1(H) = \sqrt{\int_{-\infty}^0 \left((1-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right)^2 ds + \frac{1}{2H}}$. This leads to, using properties of Hermite polynomials, the following multiple Wiener integral representation:

$$\begin{aligned} &\varepsilon^{H^*(m)-1} \int_0^t H_m(y_s^\varepsilon) ds \\ &= \frac{\varepsilon^{H^*(m)-1}}{m!} \int_{\mathbf{R}^m} \left(\int_0^t \prod_{i=1}^m h_\varepsilon(s, u_i) ds \right) dW_{u_1} \dots dW_{u_m}. \end{aligned} \tag{4.2}$$

The L^2 convergence then follows from the following lemma.

Lemma 4.2 *As $\varepsilon \to 0$, $\varepsilon^{H^*(m)-1} \int_0^t H_m(y_s^\varepsilon) ds$ converges to $\frac{m!}{K(H,m)} Z^{H^*(m),m}$ in L^2. Equivalently,*

$$\left\| \int_0^t \prod_{i=1}^m h_\varepsilon(s, u_i) ds - \int_0^t \prod_{i=1}^m (s - u_i)_+^{H-\frac{3}{2}} ds \right\|_{L^2(\mathbf{R}^m)} \to 0.$$

This is shown in Appendix 7.4 by applying [Taq79, Theorem 4.7], where weak convergence is obtained, and making a small modification.

We proceed with the L^2 convergence. The Wiener integral representation for y_t^ε, (4.1) can be obtained by applying the integral representation for fBM's:

$$B_t^H = \int_{-\infty}^t g(t,s) dW_s, \qquad \text{where} \quad g(t,s) = \frac{1}{c_1(H)} \int_{-\infty}^t (r-s)_+^{H-\frac{3}{2}} dr,$$

and by repeated applications of integration by parts (to the Young integrals):

$$\begin{aligned}\sigma\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}}dB^H_s &= \sigma B^H_t - \frac{\sigma}{\varepsilon}\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}}B^H_s ds\\ &= \sigma B^H_t - \frac{\sigma}{\varepsilon}\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}}\left(\int_{\mathbf{R}} g(s,r)dW_r\right)ds\\ &= \sigma B^H_t - \frac{\sigma}{\varepsilon}\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}} g(s,r)dsdW_r\\ &= \sigma\int_{\mathbf{R}}\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}}\partial_s g(s,r)dsdW_r\\ &= \frac{\sigma}{c_1(H)}\int_{\mathbf{R}}\int_{-\infty}^t e^{-\frac{t-s}{\varepsilon}}(s-r)_+^{H-\frac{3}{2}}dsdW_r.\end{aligned}$$

Alternatively, one may use the following for Wiener integrals [PT00], taking $f \in L^1 \cap L^2$:

$$\int_{\mathbf{R}} f(u)dB^H_u = \frac{1}{c_1(H)}\int_{\mathbf{R}}\int_{\mathbf{R}} f(u)(u-s)_+^{H-\frac{3}{2}}\,du\,dW_s.$$

To return to the case of a general G, we apply Lemma 4.7 to $G - c_m H_m$.

$$\begin{aligned}&\left\|\varepsilon^{H^*(m)-1}\int_0^t (G-c_mH_m)(y^\varepsilon_s)ds\right\|_{L^p}\\ &\lesssim \|G-c_mH_m\|_{L^p}\frac{t^{H^*(m+1)}}{\varepsilon^{H^*(m+1)-H^*(m)}} \to 0\end{aligned} \tag{4.3}$$

as H^* decreases with m. This shows, in particular that in the high Hermite rank case only the first non-zero term in the Hermite expansion of G contributes to the limit. This finishes the proof □

Lemma 4.3 *Let $G = \sum_{k=m}^\infty c_k H_k$, where $m > 0$, be in $L^2(\mu)$. Then the following statements hold for the fast stationary fractional Ornstein-Uhlenbeck process y^ε_t for every parameter $H \in (0,1)$. Let $T > 0$ and c be given as in Equation (2.10),*

(a) If $H^*(m) < \frac{1}{2}$ (and for all $m \in \mathbf{N}$ in case $H = \frac{1}{2}$), then,

$$\frac{1}{\sqrt{\varepsilon}}\int_0^t G(y^\varepsilon_s)ds \to c\,\hat{W}_t,$$

in the sense of finite dimensional distributions.

(b) If $H^*(m) = \frac{1}{2}$, then,

$$\sqrt{\frac{1}{\varepsilon|\ln\varepsilon|}}\int_0^t G(y^\varepsilon_s)ds \to c\,\hat{W}_t,$$

in the sense of finite dimensional distributions.

(c) Finally if $H^*(m) > \frac{1}{2}$, then,

$$\varepsilon^{H^*(m)-1}\int_0^t G(y^\varepsilon_s)ds \to cZ_t^{H^*(m),m},$$

in the sense of finite dimensional distributions.

Proof. As mentioned above, convergence in finite dimensional distributions in case (a) was shown in [BH02] and for case (b) in [BC09]. For case (c) Lemma 4.1 proves L^2 convergence , thus , in particular convergence in finite dimensional distributions. This finishes the proof in all cases. □

Remark 4.4

1. For $H < \frac{1}{2}$ and $m = 1$, Proposition 4.3 appears to contradict with Proposition 3.12: in the first we claim the limit is a Brownian motion and in the second we claim it is a fraction Brownian motion. Both results are correct and can be easily explained. It lies in the fact that $\int_{\mathbf{R}} \varrho(s)\, ds$ vanishes if $H < \frac{1}{2}$, and so the Brownian motion limit is degenerate. Since according to [CKM03],

$$\varrho(s) = \sigma^2 \frac{\Gamma(2H+1)\sin(\pi H)}{2\pi} \int_{\mathbf{R}} e^{isx} \frac{|x|^{1-2H}}{1+x^2} dx, \tag{4.4}$$

   and by the decay estimate from (3.6), ϱ is integrable, $s(\lambda)$ is the value at zero of the inverse Fourier transform of $\varrho(s)$, which is up to a multiplicative constant $\frac{|\lambda|^{1-2H}}{1+\lambda^2}$. This is also the spectral density of y_t and has value 0 at 0. This means we have scaled too much and the correct scaling is to multiply the integral $\int_0^{\frac{t}{\varepsilon}} y_s ds$ by ε^H in which case we have B_t^H as a limit.
2. For $m > 1$ and $H < \frac{1}{2}$, the Gaussian limit is not trivial. Indeed,

$$\begin{aligned}\int_{\mathbf{R}} \varrho(s)^m ds &= C \int_{\mathbf{R}} \overbrace{\int_{\mathbf{R}} \cdots \int_{\mathbf{R}}}^{m} \prod_{k=1}^{m} e^{isx_k} \frac{|x_k|^{1-2H}}{1+|x_k|^2} dx_1 \dots dx_m\, ds \\ &= C \overbrace{\int_{\mathbf{R}} \cdots \int_{\mathbf{R}}}^{m} \frac{|x_2+\cdots+x_m|^{1-2H}}{1+|x_2+\cdots+x_m|^2} \prod_{k=2}^{m} \frac{|x_k|^{1-2H}}{1+|x_k|^2} \neq 0.\end{aligned}$$

3. For $H = \frac{1}{2}$ and for any $m \in \mathbf{N}$, the CLT is included in part (1), as the Ornstein-Uhlenbeck process driven by a Wiener process has exponentially decaying correlations.

In the next section we bound the L^p norm of the random variable

$$X^\varepsilon := \alpha(\varepsilon, H^*(m)) \int_0^t G(y_r^\varepsilon) dr$$

where y_t^ε is the rescaled stationary fractional Ornstein-Uhlenbeck process and G an L^p function of Hermite rank at least one. Since $\mathbf{E}(H_m(y_r)H_m(y_s)) \sim (\mathbf{E}(y_r y_s))^m$, these are trivial to obtain for functions in the finite chaos expansion. We show that the upper bounded of its L^p norm is of order $\frac{1}{\alpha(\varepsilon,H^*(m))}$. Hence it is of order $\frac{1}{\sqrt{\varepsilon}}$ if and only if $H^*(m) < \frac{1}{2}$; otherwise it is one of the higher orders : $\frac{|\ln \varepsilon|}{\sqrt{\varepsilon}}$ or $\varepsilon^{-H^*(m)+1}$.

4.2 Moment bounds

We will use some results from Malliavin Calculus. Let x_s be a stationary Gaussian process with $\alpha(s) = \mathbf{E}(x_s x_0)$, such that $\alpha(0) = 1$. As a real separable Hilbert space we use $\mathscr{H} = L^2(\mathbf{R}_+, \nu)$

where for a Borel-set A we have $\nu(A) = \int_{\mathbf{R}_+} \mathbf{1}_A(s) d\alpha_s$. We can replace $\mathbf{R}_+$ by $\mathbf{R}$ or by $[0,1]$. Let $\mathscr{H}^{\otimes q}$ denote the q-th tensor product of $\mathscr{H}$. For $h \in \mathscr{H}$, we may define the Wiener integrals $W(h) = \int_0^\infty h_s dx_s$ by $W([a,b]) = x(b) - x(a)$ (where $a, b \geq 0$), linearity and the Wiener isometry ($\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle = \alpha(t-s)$). Iterated Wiener integrals are defined similarly and by its values on indictor functions: $I_m(\mathbf{1}_{A_1 \times \cdots \times A_m}) = \prod_{i=1}^m W(A_i)$ where A_i are pairwise disjoint Borel subsets of $\mathbf{R}_+$. If $\mathcal{F}$ denotes the σ-field generated by x, then any $\mathcal{F}$-measurable L^2 function F has the chaos expansion: $F = \mathbf{E}F + \sum_{m=1}^\infty I_m(f_m)$ where $f_m \in L^2(\mathbf{R}_+^m)$, the latter space is with respect to the product measures. This is due to the fact that $L^2(\Omega) = \bigoplus_{m=0}^\infty \mathscr{H}_m$ where $\mathscr{H}_m$ is the closed linear space generated by $\{H_m(W(h)) : |h|_{L^2} = 1\}$, H_m are the m-th Hermite polynomials, and that $\mathscr{H}_m = I_m(L^2_{\text{sym}}(T^m))$. The last fact is due to $H_m(W(h)) = I_m(h^{\otimes m})$. In the following $\mathbb{D}^{k,p}(\mathscr{H}^{\otimes m})$ denotes the closure of Malliavin smooth random variables under the following norm $\|u\|_{\mathbb{D}^{k,p}(\mathscr{H}^{\otimes m})} = \Big(\sum_{j=0}^k \mathbf{E}\big(\|D^j u\|^p_{\mathscr{H}^{\otimes m}}\big)\Big)^{\frac{1}{p}}$.

Lemma 4.5 (Meyer's inequality) *[NP12] Let δ denote the divergence operator (one can think of δ^m as an m times iterated Wiener-Itô-integral), then for $u \in \mathbb{D}^{k,p}(\mathscr{H}^{\otimes m})$,*

$$\|\delta^m(u)\|_{L^p} \lesssim \sum_{k=0}^m \|u\|_{\mathbb{D}^{k,p}(\mathscr{H}^{\otimes m})}. \tag{4.5}$$

Lemma 4.6 *[CNN18] If $G : \mathbf{R} \to \mathbf{R}$ is a function of Hermite rank m, then G has the following multiple Wiener-Itô-integral representation:*

$$G(x_s) = \delta^m\Big(G_m(x_s)\mathbf{1}_{[0,s]}^{\otimes m}\Big), \tag{4.6}$$

where G_m has the following properties:

(1) $\|G_m(x_1)\|_{L^p} \lesssim \|G(x_1)\|_{L^p}$,

(2) $G_m(x_1)$ is m times Malliavin differentiable and its k^{th} derivative, denoted by $G_m^{(k)}(x_1)\mathbf{1}_{[0,1]}^{\otimes k}$, satisfies $\|G_m^{(k)}(x_1)\|_{L^p} \lesssim \|G(x_1)\|_{L^p}$.

In the lemma below we estimate the moments of $\int_0^t G(x_{\frac{r}{\varepsilon}}) dr$, where we need the multiple Wiener-Itô-integral representation above to transfer the correlation function to L^2 norms of indicator functions. We use an idea from [CNN18] for the estimates below.

Lemma 4.7 *Let $x_t = W([0,t])$ be a stationary Gaussian process with correlation $\alpha(t) = \mathbf{E}(x_t x_0)$ and $\mathscr{H}$ the L^2 space over $\mathbf{R}_+$ with measure $\alpha(r)dr$. If G is a function of Hermite rank m and $G \in L^p(\mu)$, then*

$$\left\|\frac{1}{\varepsilon}\int_0^t G(x_{\frac{r}{\varepsilon}}) dr\right\|_{L^p} \lesssim \|G\|_{L^p(\mu)} \left(\int_0^{\frac{t}{\varepsilon}}\int_0^{\frac{t}{\varepsilon}} \alpha(|u-r|)^m dr du\right)^{\frac{1}{2}}. \tag{4.7}$$

For the fast fractional OU process y_t^ε, we have

$$\left\|\frac{1}{\varepsilon}\int_0^t G(y_r^\varepsilon) dr\right\|_{L^p} \lesssim \begin{cases} \|G\|_{L^p(\mu)} \sqrt{\frac{t}{\varepsilon}\int_0^\infty \varrho^m(s) ds}, & \text{if} \quad H^*(m) < \frac{1}{2}, \\ \|G\|_{L^p(\mu)} \sqrt{\frac{t}{\varepsilon} \ln|\frac{1}{\varepsilon}|}, & \text{if} \quad H^*(m) = \frac{1}{2}, \\ \|G\|_{L^p(\mu)} \left(\frac{t}{\varepsilon}\right)^{H^*(m)}, & \text{otherwise}. \end{cases}, \tag{4.8}$$

in particular,

$$\left\|\int_0^t G(y_r^\varepsilon)dr\right\|_{L^p} \lesssim \frac{\|G\|_{L^p(\mu)} t^{H^*(m)\vee\frac{1}{2}}}{\alpha(\varepsilon, H^*(m))}. \tag{4.9}$$

Proof. We first use Lemma 4.6 and then apply Meyer's inequality from Lemma 4.5 to obtain

$$\begin{aligned}
\left\|\frac{1}{\varepsilon}\int_0^t G(x_{\frac{r}{\varepsilon}})dr\right\|_{L^p} &= \left\|\int_0^{\frac{t}{\varepsilon}} G(x_r)dr\right\|_{L^p} \\
&= \left\|\int_0^{\frac{t}{\varepsilon}} \delta^m\Big(G_m(x_r)\mathbf{1}_{[0,r]}^{\otimes m}\Big)\, dr\right\|_{L^p} \\
&\lesssim \sum_{k=0}^m \left\|\int_0^{\frac{t}{\varepsilon}} D^k\Big(G_m(x_r)\mathbf{1}_{[0,r]}^{\otimes m}\Big)dr\right\|_{L^p(\Omega,\mathscr{H}^{\otimes m+k})} \\
&= \sum_{k=0}^m \left\|\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)\mathbf{1}_{[0,r]}^{\otimes m+k}dr\right\|_{L^p(\Omega,\mathscr{H}^{\otimes m+k})}.
\end{aligned}$$

We estimate the individual terms, Using linearity of the inner product, and the isometry $\langle \mathbf{1}_{[0,r]}, \mathbf{1}_{[0,s]}\rangle_{\mathscr{H}} = \mathbf{E}(x_r x_s) = \alpha(r-s)$,

$$\begin{aligned}
&\left(\left\|\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)\mathbf{1}_{[0,r]}^{\otimes m+k}dr\right\|_{\mathscr{H}^{\otimes m+k}}\right)^2 \\
&= \left\langle \int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)\mathbf{1}_{[0,r]}^{\otimes m+k}dr, \int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_u)\mathbf{1}_{[0,r]}^{\otimes m+k}du \right\rangle_{\mathscr{H}^{\otimes m+k}} \\
&= \int_0^{\frac{t}{\varepsilon}}\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)G_m^{(k)}(x_u)\langle \mathbf{1}_{[0,r]}^{\otimes m+k}, \mathbf{1}_{[0,u]}^{\otimes m+k}\rangle_{\mathscr{H}^{\otimes m+k}}\, drdu \\
&= \int_0^{\frac{t}{\varepsilon}}\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)G_m^{(k)}(x_u)\alpha(r-u)^{m+k}drdu.
\end{aligned}$$

Using Minkowski's inequality we obtain

$$\begin{aligned}
&\sum_{k=0}^m \left\|\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)\mathbf{1}_{[0,r]}^{\otimes m+k}dr\right\|_{L^p(\Omega,\mathscr{H}^{\otimes m+k})} \\
&\le \sum_{k=0}^m \left(\left\|\int_0^{\frac{t}{\varepsilon}}\int_0^{\frac{t}{\varepsilon}} G_m^{(k)}(x_r)G_m^{(k)}(x_u)\alpha(r-u)^{m+k}drdu\right\|_{L^{\frac{p}{2}}(\Omega)}\right)^{\frac{1}{2}} \\
&\le \sum_{k=0}^m \left(\int_0^{\frac{t}{\varepsilon}}\int_0^{\frac{t}{\varepsilon}} \left\|G_m^{(k)}(x_r)G_m^{(k)}(x_u)\right\|_{L^{\frac{p}{2}}(\Omega)}\alpha(r-u)^{m+k}drdu\right)^{\frac{1}{2}}.
\end{aligned}$$

We then estimate $\mathbf{E}|G_m^{(k)}(x_r)G_m^{(k)}(x_u)|^{\frac{p}{2}}$ by Hölder's inequality and the fact that x_t is stationary.

The right hand side is then controlled by

$$RHS \le \sum_{k=0}^{m} \|G_m^{(k)}(x_1)\|_{L^p} \left(\int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} \alpha(|u-r|)^{m+k} dr du \right)^{\frac{1}{2}}$$
$$\lesssim \|G\|_{L^p(\mu)} \left(\int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} \alpha(|u-r|)^{m} dr du \right)^{\frac{1}{2}},$$

concluding (4.7). We finally apply Lemma 3.4 to conclude (4.8). □

4.3 Limit theorems in 1-d

Proposition 4.8 (Enhanced limit theorem 1-d) *Let $H \in (\frac{1}{3}, 1)$ and $G = \sum_{k=m}^{\infty} c_k H_k$ in $L^p(\mu)$ where $m > 0$ and $p > 2$. Set*

$$X_t^\varepsilon = \alpha(\varepsilon, H^*(m)) \int_0^t G(y_s^\varepsilon) ds, \qquad \mathbb{X}_{s,t}^\varepsilon = \int_0^t (X_r^\varepsilon - X_s^\varepsilon) dX_r^\varepsilon.$$

Let $T > 0$, then for c as in Equation (2.10),

(a) If $H^*(m) \le \frac{1}{2}$ (and for all $m \in \mathbf{N}$ in case $H = \frac{1}{2}$), then for any $\gamma \in (0, \frac{1}{2} - \frac{1}{p})$,

$$\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon) \to c\, \hat{\mathbf{W}},$$

weakly in $\mathscr{C}^\gamma([0,T])$, where $\hat{\mathbf{W}}$ denotes a Wiener process enhanced with its Stratonovich integral.

(b) If $H^*(m) > \frac{1}{2}$ and $\frac{1}{p} < H^*(m) - \frac{1}{2}$, then for any $\gamma \in (0, H^*(m) - \frac{1}{p})$,

$$X_t^\varepsilon \to c Z_t^{H^*(m), m},$$

weakly in $\mathcal{C}^\gamma([0,T])$.

Proof. Convergence in finite dimensional distributions for both cases was shown in Lemma 4.3, using the moment bounds obtained in Lemma 4.7 together with an application of Kolmogorv's Lemma we obtain that X^ε is tight in $\mathcal{C}^\gamma$ for the stated γ. Now $\mathbb{X}_{s,t}^\varepsilon = \frac{1}{2}(X_t^\varepsilon - X_s^\varepsilon)^2$, and so the joint convergence in the finite dimensional distribution follows. The convergence in the rough path topology is concluded by Lemmas 3.10 and 3.11. See also Remark 3.14. □

4.4 Homogenization/proof of Thm C

We have all the ingredients at hand for proving Theorem C. Consider

$$dx_t^\varepsilon = \alpha(\varepsilon) f(x_t^\varepsilon) G(y_t^\varepsilon), \quad G = \sum_{k=m}^{\infty} c_k H_k \tag{4.10}$$

where $\alpha(\varepsilon) = \alpha(\varepsilon, H^*(m))$. We show $x_t^\varepsilon \to x$ where x is the solution to $\dot{x}_t = c\ f(x_t)\ dX_t$ where X_t is either a Wiener process or a Hermite processes, depending on $H^*(m)$, and the constant c is given as in Equation 2.10.

Proof. For the first case we rewrite our equation $dx_t^\varepsilon = \alpha(\varepsilon)f(x_t^\varepsilon)G(y_t^\varepsilon)$ into the Young differential equation $dx_t^\varepsilon = f(x_t^\varepsilon)dX_t^\varepsilon$. By Proposition 4.8 X_t^ε converges weakly in $\mathcal{C}^\gamma$ to a Hermite process $Z^{H^*(m),m}$. Using the continuity of the solution map of Young differential equation with respect to its driver we obtain the first result. Concerning the second result, we rewrite our equation $dx_t^\varepsilon = \alpha(\varepsilon)f(x_t^\varepsilon)G(y_t^\varepsilon)$ into the rough differential equation $dx_t^\varepsilon = f(x_t^\varepsilon)d\mathbf{X}_t^\varepsilon$. Now, by Proposition 4.8, $\mathbf{X}^\varepsilon$ converges weakly in $\mathscr{C}^\gamma$ for $\gamma \in (\frac{1}{3}, \frac{1}{2} - \frac{1}{p})$. Therefore, using the continuity Theorem 3.8 finishes the proof. □

5 Enhanced functional limit theorems

The following convention is in place throughout this section unless otherwise stated.

Convention 5.1 For $k = 1, \dots, N$, each $G_k : \mathbf{R} \to \mathbf{R}$ is an L_2 function with Hermite ranks $m_k \geq 1$. Set

$$X_t^\varepsilon = \left(X_t^{1,\varepsilon}, \dots X_t^{N,\varepsilon}\right),$$

where

$$X_t^{k,\varepsilon} = \alpha_k(\varepsilon)\int_0^t G_k(y_s^\varepsilon)ds, \qquad \alpha_k(\varepsilon) = \alpha(\varepsilon, H^*(m_k)). \tag{5.1}$$

Furthermore, we remind the reader that $n \geq 0$ is a number such that for $k \leq n$ we have $H^*(m_k) \leq \frac{1}{2}$ and for $k > n$, $H^*(m_k) > \frac{1}{2}$. We also set

$$X_t^{W,\varepsilon} = \left(X_t^{1,\varepsilon}, \dots X_t^{n,\varepsilon}\right)$$
$$X_t^{Z,\varepsilon} = \left(X_t^{n+1,\varepsilon}, \dots X_t^{N,\varepsilon}\right),$$

so that $X_t^\varepsilon = (X_t^{W,\varepsilon}, X_t^{Z,\varepsilon})$.

5.1 Convergence of the vector valued processes of mixed type

The convergence of each component has already been proved earlier, so it is only left to show that they converge jointly. We must specify the correlations between the limiting components. If they converge jointly and if the limiting distribution is independent, then the covariance has to converge to zero also, this we do not expect to hold in general. For example if all G_i are equal then all the components of the limiting driver are the same. If we have H_i and H_j where $i \neq j$, we may expect non-trivial correlations. On the other hand we know different scales $y_{t/\varepsilon}$ and y_{t/ε^α} where $\alpha \neq 1$ are 'expected' to have uncorrelated scaling limits, this is reflected in the different scaling constants. First we will establish joint convergence under the assumption that each component converges to a Wiener process. We then show the joint convergence for the case each component limit is a Hermite process, and then for the case the component limits can be either a Brownian motion or a Hermite process. Due to a reduction lemma (Lemma 5.3 below), the joint convergence can be reduced to G_i being a finite sum of Hermite polynomials.

Lemma 5.2 (CLT-Gaussian) *Fix $H \in (0,1) \setminus \{\frac{1}{2}\}$. Here we consider the fist n components of X^ε, which are denoted by $X^{W,\varepsilon}$, so $G_k \in L^2(\mu)$, $k \leq n$, with Hermite rank $m_k > 0$ and $H^*(m_k) \leq \frac{1}{2}$.*

1. *Then, as* $\varepsilon \to 0$*, the following converges in finite dimensional distribution:*
$$X^{W,\varepsilon} \to (X^1, X^2, \dots X^n) = X^W.$$
2. *The limiting distribution is Gaussian with covariance between the* ith *and the* jth *components given by*
$$\mathbf{E}[X^i(t)X^j(s)] = 2(s \wedge t) \int_0^\infty \mathbf{E}(G_i(y_r)G_j(y_0))dr.$$
3. *If, in addition,* $G_k \in L^{p_k}(\mu)$ *for* $p_k > 2$*, then the convergence is weakly in* $\mathcal{C}^\gamma$ *where* $\gamma \in (0, \min_{k=1\dots n} \frac{1}{2} - \frac{1}{p_k})$.

Proof. First we define the truncated functions $G_{k,M} = \sum_{j=m_k}^M c_{k,j} H_j$ and set
$$X_M^{k,\varepsilon} = \alpha_k(\varepsilon) \int_0^t G_{k,M}(y_s^\varepsilon)ds.$$
Then, by Lemma 5.3 below, it is sufficient to show the convergence of $(X_M^{1,\varepsilon}, \dots, X_M^{n,\varepsilon})$ for every M. By earlier considerations each $X_M^{k,\varepsilon}$ converges to a Wiener process X_M^k. As each $X_M^{k,\varepsilon}$ belongs to a finite chaos we can make use of the normal approximation theorem from [NP12, Theorem 6.2.3]: if each component of a family of mean zero vector valued stochastic processes, with components of the form $I_{q_i}(f_{i,n})$ where $f_{i,n}$ are symmetric L^2 functions in q_i variables, converges in law to a Gaussian process, then they converge jointly in law to a vector valued Gaussian process, provided that their correlation functions converge. Furthermore, the correlation functions of the limit distribution are: $\lim_{\varepsilon \to 0} \mathbf{E}[X^{i,\varepsilon}(t)X^{j,\varepsilon}(s)]$. Let $m = \min(m_i, m_j)$ we use
$$\mathbf{E}(H_n(y_t)H_m(y_s)) = \delta_{n,m}(\mathbf{E}(y_s^\varepsilon y_t^\varepsilon))^m$$
to obtain, for $s \le t$,
$$\begin{aligned}
&\mathbf{E}\left[\alpha_i(\varepsilon)\alpha_j(\varepsilon) \int_0^t G_{i,M}(y_u^\varepsilon)du \int_0^s G_{j,M}(y_r^\varepsilon)dr\right] \\
&= \sum_{k=m}^M \alpha_i(\varepsilon)\alpha_j(\varepsilon)c_{i,k}c_{j,k}(k!)^2 \int_0^t \int_0^s (\mathbf{E}(y_r^\varepsilon y_u^\varepsilon))^k drdu \\
&= \sum_{k=m}^M \alpha_i(\varepsilon)\alpha_j(\varepsilon)c_{i,k}c_{j,k}(k!)^2 \left(\int_0^s \int_0^s \varrho^\varepsilon(u-r)^k drdu + \int_s^t \int_0^s \varrho^\varepsilon(u-r)^k drdu\right)
\end{aligned}$$
By Lemma 3.4 we obtain, for $\varepsilon \to 0$,
$$\alpha_i(\varepsilon)\alpha_j(\varepsilon) \int_s^t \int_0^s \varrho^\varepsilon(u-r)^k drdu \to 0.$$
Hence,
$$\begin{aligned}
\lim_{\varepsilon \to 0} RHS &= 2 \sum_{k=m}^M c_{i,k}c_{j,k}(k!)^2 \lim_{\varepsilon \to 0} \left(\alpha_i(\varepsilon)\alpha_j(\varepsilon) s \int_0^{\frac{s}{\varepsilon}} (\varrho(v))^k dv\right) \\
&= 2(s \wedge t) \sum_{k=m}^M c_{i,k}c_{j,k}(k!)^2 \int_0^\infty \varrho(u)^k du \\
&= 2(s \wedge t) \int_0^\infty \mathbf{E}(G_{i,M}(y_s)G_{j,M}(y_0))ds,
\end{aligned}$$

proving the finite chaos case. We now prove that the correlations of the limit converge as $M \to \infty$. Indeed,

$$\lim_{M\to\infty} 2(s\wedge t)\sum_{k=m}^{M} c_{i,k}c_{j,k}(k!)^2\int_0^\infty \varrho(u)^k du = 2(s\wedge t)\sum_{k=m}^{\infty} c_{i,k}c_{j,k}(k!)^2\int_0^\infty \varrho(u)^k du$$
$$= 2(s\wedge t)\int_0^\infty \mathbf{E}(G_i(y_s)G_j(y_0))ds.$$

As $G_{i,M} \to G_i$ in L^2, and similarly for j, this proves the first two claims. The convergence in Hölder spaces follows from Lemma 4.7, which states these processes are tight in $\mathcal{C}^\gamma$, c.f. Proposition 4.8. This concludes the proof. □

Now we prove the reduction lemma for the high Hermite rank case.

Lemma 5.3 (Reduction Lemma) *Fix $H \in (0,1)\setminus\{\frac{1}{2}\}$. For $M \in \mathbf{N}$, define*

$$X_M^{k,\varepsilon}(t) = \alpha_k(\varepsilon)\int_0^t G_{k,M}(y_s^\varepsilon)ds.$$

If for every $M \in \mathbf{N}$,

$$(X_M^{1,\varepsilon},\dots,X_M^{N,\varepsilon}) \longrightarrow (X_M^1,\dots,X_M^N)$$

in finite dimensional distributions, then,

$$(X^{1,\varepsilon},\dots,X^{N,\varepsilon}) \longrightarrow (X^1,\dots,X^N),$$

in finite dimensional distributions.

Proof. Firstly we show, for any sequence of positive numbers $\{t_{\gamma_{k,l}}, k\le N, l\le A\}$, $\sum_{k=1}^N\sum_{l=1}^A \gamma_{k,l}X_M^{k,\varepsilon}(t_l)$ converges as $M\to\infty$. By the triangle inequality we can reduce

$$\left\|\sum_{k,l}\gamma_{k,l}\Big(X^{k,\varepsilon}(t_l) - X_M^{k,\varepsilon}(t_l)\Big)\right\|_{L^2} \to 0.$$

to $\|X^{k,\varepsilon}(t) - X_M^{k,\varepsilon}(t)\|_{L^2} \to 0$. Now,

$$X^{k,\varepsilon}(t) - X_M^{k,\varepsilon}(t) = \alpha_k(\varepsilon)\int_0^t \Big(G_k(y_s^\varepsilon) - G_{k,M}(y_s^\varepsilon)\Big)ds$$
$$= \alpha_k(\varepsilon)\int_0^t \sum_{j=M+1}^\infty c_{k,j}H_j(y_s^\varepsilon)ds.$$

Using properties of the Hermite polynomials we obtain

$$\begin{aligned}
&\mathbf{E}\left(\alpha_k(\varepsilon)\int_0^t \sum_{j=M+1}^{\infty} c_{k,j}H_j(y_s^\varepsilon)ds\right)^2 \\
&= \alpha_k(\varepsilon)^2 \int_0^t\int_0^t \sum_{j=M+1}^{\infty} (c_{k,j})^2\mathbf{E}(H_j(y_s^\varepsilon)H_j(y_r^\varepsilon))drds \\
&= \alpha_k(\varepsilon)^2 \sum_{j=M+1}^{\infty} (c_{k,j})^2 j! \int_0^t\int_0^t \varrho^\varepsilon(|s-r|)^j drds \\
&\lesssim t\int_0^\infty \varrho(u)^{M+1}du \sum_{j=M+1}^{\infty} (c_{k,j})^2 j!
\end{aligned}$$

As $\sum_{j=m}^{\infty}(c_{k,j})^2 j! < \infty$ we obtain $\sum_{j=M+1}^{\infty}(c_{k,j})^2 j! \to 0$ as $M \to \infty$. Then,

$$\lim_{M\to\infty}\lim_{\varepsilon\to 0}\mathbf{E}\left(\alpha_k(\varepsilon)\int_0^t G_k(y_s^\varepsilon)ds - \alpha_k(\varepsilon)\int_0^t G_{k,M}(y_s^\varepsilon)ds\right)^2 \to 0, \tag{5.2}$$

and finally using theorem 3.2 in [Bil99] we conclude the proof. □

Now, we go ahead and deal with the low Hermite rank case, so focus on the vector component whose entries satisfy $H^*(m_k) > \frac{1}{2}$. Recall, this implies $H > \frac{1}{2}$.

Lemma 5.4 (CLT- Hermite) *Fix $H \in (0,1)\setminus\{\frac{1}{2}\}$. Write $G_k = \sum_{j=m_k}^{N} c_{k,j}H_j$. Suppose that $m_k \geq 1$ and $H^*(m_k) > \frac{1}{2}$.*

1. *Then, $(X^{n+1,\varepsilon}, X^{2,\varepsilon}\dots, X^{N,\varepsilon})$ converges in finite dimensional distribution and for every $t \in [0,T]$*

$$\left\|(X_t^{n+1,\varepsilon}, X_t^{n+2,\varepsilon}\dots, X_t^{N,\varepsilon}) - (X_t^{n+1}, X_t^{n+2}, \dots, X_t^{N})\right\|_{L_2(\Omega)} \to 0.$$

2. *The marginals of the limit are the following Hermite processes, each given by the representation 3.1 with a common Wiener process W*

$$X^k = \frac{c_{k,m_k} m_k!}{K(H^*(m_k), m_k)} Z^{H^*(m_k),m_k}.$$

3. *If in addition $G_k \in L^{p_k}$ for $p_k > 2$, then the convergence is weakly in $\mathcal{C}^\gamma$ on any finite time interval where $\gamma \in (0, \min_{k=n+1,\dots,N} H^*(m_k) - \frac{1}{p_k})$.*

Proof. By Lemma 4.1 each component convergences in L^2, hence the above converges as well in L^2 yielding convergence in finite dimensional distributions by an application of the Cramer-Wold theorem. The convergence in Hölder spaces follows from Lemma 4.7, which states these processes are tight in $\mathcal{C}^\gamma$, c.f. Proposition 4.8. □

Proposition 5.5 (CLT-mixed) *For each k, write*

$$G_k = \sum_{j=m_k}^{\infty} c_{k,j} H_j.$$

Then, for X^k given in Lemmas 5.2 and 5.4,

$$X^\varepsilon = (X^{1,\varepsilon}, X^{2,\varepsilon} \dots, X^{N,\varepsilon}) \longrightarrow (X^1, X^2 \dots, X^N) = (X^W, X^Z)$$

in finite dimensional distributions. Furthermore,

1. *X^W and X^Z are independent.*
2. *If $i, j \le n$, so X^i and X^j are Gaussian, their correlation is*

$$2(s \wedge t) \int_0^\infty \int_0^\infty \mathbf{E}(G_i(y_s) G_j(y_0))\, ds.$$

3. *If $i, j \ge n+1$, both X^i and X^j are Hermite processes, then their correlation is given by a common Wiener process W_t. Specifically,*

$$\begin{aligned}&\mathrm{Cov}(X^i, X^j)\\ =&\delta_{m_i,m_j} c_{i,m_i} c_{j,m_j} \int_0^t \int_0^t \left(\int_{\mathbf{R}} (s-\xi)_+^{\hat{H}(m)-\frac{3}{2}} (r-\xi)_+^{\hat{H}(m)-\frac{3}{2}} d\xi \right)^{m_i} dr ds.\end{aligned}$$

4. *If in addition $G_k \in L^{p_k}$ for $p_k > 2$, then the convergence is weakly in $\mathcal{C}^\gamma$ for every $\gamma \in (0, \min_{k=1,\dots n} \frac{1}{2} - \frac{1}{p_k} \wedge \min_{k=n+1,\dots N} H^*(m_k) - \frac{1}{p_k})$.*

Proof. Using Lemma 5.2 and 5.4, $X^{W,\varepsilon} \to X^W$ and $X^{Z,\varepsilon} \to X^Z$ in finite dimensional distributions. By Lemma 5.3 and Equation (4.3) , we may reduce the problem to

$$G_i = \sum_{k=m_i}^{M} c_{i,k} H_k, \quad G_j = c_{j,m_j} H_{m_j}, \qquad 1 \le i \le n,\ j > n.$$

Now, we can rewrite $H_m(y_s^\varepsilon) = I_m(f_s^{m,\varepsilon})$, where I_m denotes a m-fold Wiener-Ito integral and a function $f_s^{m,\varepsilon} \in L^2(\mathbf{R}^m, \mu)$. Now, for $1 \le i \le n$ we obtain,

$$\begin{aligned}\alpha_i(\varepsilon) \int_0^t G_i(y_s^\varepsilon) ds &= \alpha_i(\varepsilon) \int_0^t \sum_{k=m_i}^{M} c_{i,k} H_k(y_s^\varepsilon) ds\\ &= \alpha_i(\varepsilon) \int_0^t \sum_{k=m_i}^{M} c_{i,k} I_k(f_s^{k,\varepsilon}) ds = \sum_{k=m_i}^{M} c_{i,k} I_k(\hat{f}_t^{k,\varepsilon}),\end{aligned}$$

where $\hat{f}^{k,\varepsilon} = \int_0^t f_s^{k,\varepsilon} ds$. Similarly for $j > n$,

$$\int_0^t G_j(y_s^\varepsilon) ds = \int_0^t c_{j,m_j} H_{m_j}(y_s^\varepsilon) ds = c_{j,m_j} I_{m_j}(\hat{f}_t^{m_j,\varepsilon}).$$

Hence, we only need to show that the collection of stochastic processes of the form $I_{m_k}(\hat{f}^{m_k,\varepsilon})$ converge jointly in finite dimensional distribution. It is then sufficient to show for every finite collection of times $t_l \in [0,T]$, that the vector

$$\Big\{ I_{m_k}(\hat{f}^{k,\varepsilon}_{t_l}), k = m, \dots M \Big\},$$

converges jointly, where $m = \min_{k=1,\dots,N} m_k$. Let n_0 denote the smallest natural number such that $H^*(n_0) \le \frac{1}{2}$. For $k \le n_0$, the collection $I_k(\hat{f}^{k,\varepsilon}_{t_l})$ converges to a normal distribution and therefore, by the fourth moment theorem [NP05, Theorem 1],

$$\|\hat{f}^{k,\varepsilon}_{t_l} \otimes_r \hat{f}^{k,\varepsilon}_{t_l}\|_{\mathscr{H}^{2k-2r}} \to 0, \qquad r = 1, \dots, k-1.$$

By Cauchy-Schwartz we obtain for $r = 1, \dots, k_1$,

$$\begin{aligned} &\Big\| \hat{f}^{k_1,\varepsilon}_{t_{l_1}} \otimes_r \hat{f}^{k_2,\varepsilon}_{t_{l_2}} \Big\|_{\mathscr{H}^{k_1+k_2-2r}} \\ &\le \Big\| \hat{f}^{k_1,\varepsilon}_{t_{l_1}} \otimes_r \hat{f}^{k_1,\varepsilon}_{t_{l_1}} \Big\|_{\mathscr{H}^{p-r}} \Big\| \hat{f}^{k_2,\varepsilon}_{t_{l_2}} \otimes_r \hat{f}^{k_2,\varepsilon}_{t_{l_2}} \Big\|_{\mathscr{H}^{q-r}} \to 0, \end{aligned}$$

for all $t_{l_1}, t_{l_2} \in \mathbf{R}$, $1 \le k_1 < n_0 \le k_2 \le M$. We can now apply the asymptotic independent theorem (see Proposition 7.14 in the Appendix), to conclude the joint convergence in finite dimensional distributions of X^ε to (X^W, X^Z). Furthermore X^W is independent of X^Z.

The correlations between X^i_t and $X^j_{t'}$, where $i, j > n$, are 0 if $m_i \ne m_j$, otherwise given by the L^2 norm of their integrands. They follow from the Itô isometries:

$$\begin{aligned} &c_{i,m_i} c_{j,m_j} \mathbf{E} \int_0^t \int_0^{t'} H_{m_i}\Big(\int_{\mathbf{R}} (s-u)_+^{\hat{H}(m_i)-\frac{3}{2}} dW_u \Big) H_{m_i}\Big(\int_{\mathbf{R}} (r-u)_+^{\hat{H}(m_i)-\frac{3}{2}} dW_u \Big) dr ds \\ &= c_{i,m_i} c_{j,m_j} \int_0^t \int_0^{t'} \int_{\mathbf{R}^{m_i}} \prod_{i=1}^{m_i} (s-\xi_i)_+^{\hat{H}(m)-\frac{3}{2}} \prod_{i=1}^{m_i} (r-\xi_i)_+^{\hat{H}(m)-\frac{3}{2}} d\xi_1 \dots \xi_{m_i} dr ds. \end{aligned}$$

The convergence in Hölder spaces follows from Lemma 4.7, which states these processes are tight in $\mathcal{C}^\gamma$, c.f. Proposition 4.8, completing the proof.

□

5.2 Convergence in finite dimensional distributions of the rough paths

We study the canonical lifts of X^ε to a rough path. We denote by $\mathbb{X}^\varepsilon$ the canonical/geometric lift. Its components are

$$\mathbb{X}^{i,j,\varepsilon}_{0,t} = \alpha_i(\varepsilon)\alpha_j(\varepsilon) \int_0^t \int_0^s G_i(y^\varepsilon_s) G_j(y^\varepsilon_r) dr ds.$$

From here on we assume Assumption 2.5, and Convention 5.1 so the high rank functions are in the first n components, for which $X^{k,\varepsilon}$ converges to a Wiener process. We first work on the case where one of the components of the iterated integral corresponds to a low Hermite rank.

Lemma 5.6 *[Young integral case] Assume Assumption 2.5. Below $i, j \in \{1, \dots N, i \vee j > n\}$. Then,*

$$(X^\varepsilon, \ \mathbb{X}^{i,j,\varepsilon}), \tag{5.3}$$

converges in finite dimensional distributions to $(X, \mathbb{X}^{i,j})$ where $\mathbb{X}^{i,j} = \int_0^t X^i dX^j$, and these integrals are well defined as Young integrals.

Proof. By Assumption 2.5, the functions G_k posses enough integrability such that each component of X^ε converges in a Hölder space. Furthermore, by Assumption 2.5 (2) there exist numbers η and τ, with $\eta + \tau > 1$, such that the Hölder regularity of the limits corresponding to a Wiener processes, are bounded below by η, and the ones corresponding to a Hermite process bounded from below by τ. Therefore, taking the integrals

$$\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s G_j(y_s^\varepsilon)G_i(y_r^\varepsilon)drds = \int_0^t X_s^{i,\varepsilon}dX_s^{j,\varepsilon}$$

is a continuous and well-defined operation from $C^\eta \times C^\tau \to C^\tau$ or $C^\tau \times C^\eta \to C^\eta$, thus weak convergence in C^η follows, and in particular convergence in finite dimensional distributions. □

Remark 5.7 Note that thee proof shows convergence in Hölder space C^η, however η here could be a very small number, below $\frac{1}{2}$, we would work harder in a later section on the tightness in the rough path space topology. For the convergence in rough topology, we want this to work in $\mathcal{C}^{2\alpha}$ for a $\alpha > \frac{1}{3}$. We would finally prove tightness of the iterated integrals in higher Hölder spaces.

Now it is only left to deal with the parts of the natural rough path lift involving two Wiener scaling terms, this is carried out in the next section.

5.2.1 Approximations of iterated integrals: Itô integral case

We proceed to establish convergence of the iterated integrals where both components correspond to the high Hermit rank case, appearing in (5.3).

Remark 5.8 We further assume $H^*(m_k) < 0$ for each k which gives rise to a Wiener scaling, we do not obtain Logarithmic terms and therefore work with the $\frac{1}{\sqrt{\varepsilon}}$ scaling from here on. Furthermore, in this case $\alpha(\varepsilon)\int_0^t G(y_s^\varepsilon)ds$ equals $\sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}} G(y_s)ds$ in law and for simplicity we will work with the latter in this chapter.

In this section, from here onwards we assume that both G_i and G_j give rise to a Wiener process in the homogenization process, so $i, j \leq n$ and

$$X_t^{k,\varepsilon} = \sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}} G(y_s)ds,$$

with the corresponding iterated integrals.

By Lemma 5.2, we know that $(X^{i,\varepsilon}, X^{j,\varepsilon}) \to (W^i, W^j)$, we will see their integral $\int_0^t X^{i,\varepsilon}dX^{j,\varepsilon}$, a double integral, can be discretised and decomposed into integrals over strips of two significant regions, the integral on the region away from the diagonal is of the form

$$\sum_{k=1}^{[\frac{t}{\varepsilon}]}\int_{k-1}^k G_i(y_s)ds\int_0^{k-1} G_j(y_r)dr,$$

which resembles a Riemann sum for an integral which we might hope to be the stochastic integral $\int_0^t W_s^j dW_s^i$. This is not quite true, however its martingale approximation does converge to the stochastic integral. We want to show that

$$\int_0^{\frac{t}{\varepsilon}} X_s^{j,\varepsilon} dX_s^{i,\varepsilon} = \varepsilon \int_0^{\frac{t}{\varepsilon}} \int_0^s G_i(y_s) G_j(y_r) dr ds$$
$$= I_1(\varepsilon) + I_2(\varepsilon),$$

where $I_1(\varepsilon) \to \int_0^t W_s^j dW_s^i$, where the integral is understood in the Itô-sense, weakly and $I_2(\varepsilon) \to tA^{i,j}$ in probability. For this we want to use the continuity property of stochastic integrals with respect to martingales and should approximate $X^{i,\varepsilon}$ with a martingale that is predictably uniform-tight, c.f. Lemma 5.19. We begin to describe this approximation.

For any L^2 functions U, V we introduce the stationary process:

$$\Phi_U(t) = \int_t^\infty U(y_r) dr.$$

which unfortunately does not have good integrability properties. We would explore a local independent decomposition of the FOU. It turns out that for every t there exists a decomposition,

$$y_t = \overline{y}_t^k + \tilde{y}_t^k,$$

where the first term $\overline{y}_t^k$ is $\mathcal{F}_k$ measurable and $\tilde{y}_t^k$ is independent of $\mathcal{F}_k$, where $\mathcal{F}_k$ is the filtration generated by the driving fractional Brownian motion. We will show later, in Proposition 7.8, for $H > \frac{1}{2}$ and $H^*(m) < 0$,

$$\sup_k \sup_{q \ge m} \int_{k-1}^\infty \int_{k-1}^\infty \mathbf{E}\Big(\overline{y}_s^k \overline{y}_t^k\Big)^q \, dt \, ds < \infty, \tag{5.4}$$

We therefore define

$$\begin{aligned} \hat{U}(k) &:= \int_{k-1}^\infty \mathbf{E}(U(y_r) \,|\, \mathcal{F}_k) \, dr, \\ \hat{V}(k) &:= \int_{k-1}^\infty \mathbf{E}(V(y_r)) \,|\, \mathcal{F}_k) \, dr. \end{aligned} \tag{5.5}$$

Note both $\hat{U}$ and $\hat{V}$ are shift covariant, i.e.$(\hat{U} \circ \tau)(k) = \hat{U}(k+1)$ where τ is the shift operator. To proceed further we need a couple of lemmas.

Lemma 5.9 *For $x, y, a, b \in \mathbf{R}$ such that $a^2 + b^2 = 1$,*

$$H_m(ax + by) = \sum_{j=0}^m \binom{m}{j} a^j b^{m-j} H_j(x) H_{m-j}(y). \tag{5.6}$$

Lemma 5.10 *Let $H > \frac{1}{2}$. Set $a_t = \|\overline{y}_t^k\|_{L^2}$. Then*

$$\mathbf{E}[H_m(y_t) | \mathcal{F}_k] = (a_t)^m H_m\left(\frac{\overline{y}_t^k}{a_t}\right).$$

Proof. Set $b_t = \|\tilde{y}_t^k\|_{L^2}$. By the independence of $\overline{y}_t^k$ and $\tilde{y}_t^k$ we obtain

$$1 = \|y_k\|_{L^2}^2 = \|\overline{y}_t^k\|_{L^2}^2 + \|\tilde{y}_t^k\|_{L^2}^2 = (a_t)^2 + (b_t)^2.$$

Now we decompose $H_m(y_t)$ using the above identity and obtain,

$$\begin{aligned} H_m(y_t) = H_m\Big(\overline{y}_t^k + \tilde{y}_t^k\Big) &= H_m\left(a_t\left(\frac{\overline{y}_t^k}{a_t}\right) + b_t\left(\frac{\tilde{y}_t^k}{b_t}\right)\right) \\ &= \sum_{j=0}^m \binom{m}{j} a_t^j b_t^{m-j} H_j\left(\frac{\overline{y}_t^k}{a_t}\right) H_{m-j}\left(\frac{\tilde{y}_t^k}{b_t}\right). \end{aligned}$$

Note that by construction $\frac{\overline{y}_t^k}{a_t}$ and $\frac{\tilde{y}_t^k}{b_t}$ are standard Gaussian random variables. Therefore, by the independence $\tilde{y}_t^k$ of $\mathcal{F}_t$,

$$\begin{aligned} \mathbf{E}[H_m(y_t)|\mathcal{F}_k] &= \sum_{j=0}^m \binom{m}{j} (a_t)^j (b_t)^{m-j} H_j\left(\frac{\overline{y}_t^k}{a_t}\right) \mathbf{E}\left[H_{m-j}\left(\frac{\tilde{y}_t^k}{b_t}\right)|\mathcal{F}_k\right] \\ &= (a_t)^m H_m\left(\frac{\overline{y}_t^k}{a_t}\right). \end{aligned}$$

We have used the fact that $\mathbf{E}H_j\left(\frac{\tilde{y}_t^k}{b_t}\right)$ vanishes for any $j \geq 1$ and $H_0 = 1$. □

Proposition 5.11 *[See §7.3] Let $H > \frac{1}{2}$ and $U \in L^2(\mathbf{R}, \mu)$ with Hermite rank $m > \frac{1}{1-H}$. Then the process $\hat{U}(j)$ is bounded in L^2 (provided (5.4) holds.)*

Proof. The proof is less straightforward due to the lack of the strong mixing property. Here we rely on (5.4), whose proof is lengthy and independent of the error estimates here and is therefore postponed to §7.3.

We compute the L^2 norm, using the definition of $\hat{U}$ and the Hermite expansion $U = \sum_{q=m}^\infty c_q H_q$,

$$\begin{aligned} \|\hat{U}(k)\|_{L^2} &= \int_{k-1}^\infty \int_{k-1}^\infty \mathbf{E}(\mathbf{E}[U(y_s)|\mathcal{F}_k]\mathbf{E}[U(y_r)|\mathcal{F}_k])\, dr ds \\ &= \int_{k-1}^\infty \int_{k-1}^\infty \sum_{q=m}^\infty \sum_{j=m}^\infty c_q c_j \mathbf{E}\Big(\mathbf{E}[H_q(y_s)|\mathcal{F}_k]\, \mathbf{E}[H_j(y_r)|\mathcal{F}_k]\Big) dr\, ds \\ &= \int_{k-1}^\infty \int_{k-1}^\infty \sum_{q=m}^\infty (c_q)^2 \mathbf{E}\left((a_s)^q (a_r)^q H_q\left(\frac{\overline{y}_s^k}{a_s}\right) H_q\left(\frac{\overline{y}_r^k}{a_r}\right)\right) dr\, ds \\ &= \int_{k-1}^\infty \int_{k-1}^\infty \sum_{q=m}^\infty (c_q)^2\, q!\, (a_s)^q (a_r)^q \mathbf{E}\left(\frac{\overline{y}_s^k}{a_s}\frac{\overline{y}_r^k}{a_r}\right)^q dr\, ds \\ &= \int_{k-1}^\infty \int_{k-1}^\infty \sum_{q=m}^\infty (c_q)^2\, q!\, \mathbf{E}\Big(\overline{y}_s^k \overline{y}_r^k\Big)^q dr\, ds. \end{aligned}$$

By summability of $(c_q)^2 q!$, following from $U \in L^2$. The proof follows from the assumption that $\sup_{q\geq m,k} \int_{k-1}^\infty \int_{k-1}^\infty \mathbf{E}\big(\overline{y}_s^k \overline{y}_r^k\big)^q dr ds$ is finite. This concludes the lemma. □

As a corollary of Proposition 5.11 we have

Corollary 5.12 *The process* $(M_k, k \geq 1)$*, where*

$$M_k = \sum_{j=1}^{k} \Big(\hat{U}(j) - \mathbf{E}\Big(\hat{U}(j)|\mathcal{F}_{j-1}\Big)\Big),$$

is an $\mathcal{F}_k$*-adapted* L^2 *martingale with shift covariant martingale difference. Similarly,*

$$N_k = \sum_{j=1}^{k} \Big(\hat{V}(j) - \mathbf{E}(\hat{V}(j)|\mathcal{F}_{j-1})\Big),$$

defined also an $\mathcal{F}_k$*-adapted* L^2 *martingale.*

Proposition 5.13 *There exists a function* $\mathbf{Er}(\varepsilon)$ *converging to zero in probability as* $\varepsilon \to 0$ *such that*

$$\varepsilon \int_0^{\frac{t}{\varepsilon}} \int_0^s U(y_s)V(y_r)drds \quad = \varepsilon \sum_{k=1}^{[\frac{t}{\varepsilon}]} (M_{k+1} - M_k)N_k + (s \wedge t)\gamma + \mathbf{Er}_1(\varepsilon) \tag{5.7}$$

where

$$\gamma = \int_0^\infty \mathbf{E}(U(y_s)V(y_0))ds.$$

The proof for this is given in the rest of the section. Note that the Itô integral approximations work well while the processes involved have independent increments or satisfy strong mixing properties. To tackle the lack of these properties, we use a locally independent decompositions of the fOU. We also use Birkhoff's ergodic theorem. After proving the Proposition, in the next section we show that $\varepsilon \sum_{k=1}^{[\frac{t}{\varepsilon}]} (M_{k+1} - M_k)N_k$ converges to the relevant Itô integrals of the limits of $\sqrt{\varepsilon} \int_0^{[\frac{t}{\varepsilon}]} U(y_r)dr$ and $\sqrt{\varepsilon} \int_0^{[\frac{t}{\varepsilon}]} V(y_r)dr$.

Lemma 5.14 *The stationary Ornstein-Uhlenbeck process is ergodic.*

A stationary Gaussian process is ergodic if its spectral measure has no atom, [CFS82, Sam06]. The spectral measure F of a stationary Gaussian process is obtained from Fourier transforming its correlation function and $\varrho(\lambda) = \int_{\mathbf{R}} e^{i\lambda x} dF(x)$. According to [CKM03]:

$$\varrho(s) = \frac{2\Gamma(2H+1)\sin(\pi H)}{2\pi} \int_{\mathbf{R}} e^{isx} \frac{|x|^{1-2H}}{1+x^2} dx, \tag{5.8}$$

so the spectral measure is absolutely continuous with respect to the Lebesgue measure with spectral density $s(x) = c\frac{|x|^{1-2H}}{1+x^2}$.

For $k = 1, 2, \ldots$, we define the $\mathcal{F}_k$-adapted processes:

$$I(k) = \int_{k-1}^{k} U(y_s)ds = \Phi_U(k) - \Phi_U(k-1)$$
$$J(k) = \int_{k-1}^{k} V(y_s)ds = \Phi_V(k) - \Phi_V(k-1).$$

Remark 5.15 We note the following useful identities. For $k = 1, 2, \dots,$

$$\hat{U}(k) = I(k) + \mathbf{E}[\hat{U}(k+1) \,|\, \mathcal{F}_k], \tag{5.9}$$

$$M_{k+1} - M_k = I(k) + \hat{U}(k+1) - \hat{U}(k), \tag{5.10}$$

$$\int_0^k U(y_r)dr = M_k - \hat{U}(k) + \hat{U}(1) - M_1.$$

Proposition 5.16 *Suppose that U and V satisfy the assumptions imposed above, then the triple below converges in finite dimensional distributions.*

$$\lim_{\varepsilon \to 0} \left(\sqrt{\varepsilon} M_{[\frac{t}{\varepsilon}]},\ \sqrt{\varepsilon} N_{[\frac{t}{\varepsilon}]},\ \varepsilon \sum_{k=1}^{[\frac{t}{\varepsilon}]} (M_{k+1} - M_k) N_k \right) = \left(W_t^1,\ W_t^2,\ \int_0^t W_s^1 \, dW_s^2 \right).$$

Here W^1, W^2 are standard Wiener processes with covariance $\int_0^\infty \mathbf{E}(U(y_r)V(y_0))dr$, and variances respectively $\int_0^\infty \mathbf{E}(U(y_r)U(y_0))dr$ and $\int_0^\infty \mathbf{E}(V(y_r)V(y_0))dr$. The integration is in Itô sense.

Proof. Define

$$M_t^\varepsilon = \sqrt{\varepsilon} M_{[\frac{t}{\varepsilon}]}, \qquad N_t^\varepsilon = \sqrt{\varepsilon} N_{[\frac{t}{\varepsilon}]},$$

Using the identity (5.10) we show that

$$\begin{aligned} M_t^\varepsilon &= \sqrt{\varepsilon} \sum_{k=1}^{[\frac{t}{\varepsilon}]} (M_{k+1} - M_k) + \sqrt{\varepsilon}\, M_1 \\ &= \sqrt{\varepsilon} \int_0^{[\frac{t}{\varepsilon}]} U(y_r) ds + \sqrt{\varepsilon} \hat{U}([\frac{t}{\varepsilon}]) - \sqrt{\varepsilon} \hat{U}(1) + \sqrt{\varepsilon} M_1. \end{aligned}$$

Since $\hat{U}$ is L^2 bounded, the joint convergence of M_t^ε and N_t^ε, in finite dimensional distributions follows from Lemma 5.2. Next observe that,

$$\varepsilon \sum_{k=1}^{[\frac{t}{\varepsilon}]} (M_{k+1} - M_k) N_k = \int_0^t M_s^\varepsilon dN_s^\varepsilon.$$

The joint convergence follows since $\mathbf{E}(M_t^\varepsilon)^2 \lesssim t + o(\varepsilon)$, see Lemma 5.19 and Lemma 5.20, we can use the continuity theorems on integrals with respect to martingales with jumps. □

Henceforth, in this section we set $L = L(\varepsilon) = [\frac{t}{\varepsilon}]$.

Lemma 5.17 *There exists a function $\mathbf{Er}_1(\varepsilon)$, which converges to zero in probability as $\varepsilon \to 0$, such that*

$$\begin{aligned} &\varepsilon \int_0^{\frac{t}{\varepsilon}} \int_0^s U(y_s) V(y_r) dr ds \\ &= \varepsilon \sum_{k=1}^{L} I(k) \sum_{l=1}^{k-1} J(l) + t \int_0^1 \int_0^s \mathbf{E}(U(y_s)V(y_r))\, dr ds + \mathbf{Er}_1(\varepsilon) \end{aligned} \tag{5.11}$$

Proof. Let us divide the integration region $0 \le r \le s \le L$ main region and the other negligible regions.

$$\int_0^L \int_0^s U(y_s)V(y_r)drds + \varepsilon \int_L^{\frac{t}{\varepsilon}} \int_0^s U(y_s)V(y_r)drds.$$

The second term, integration in the small region, is of order $o(\varepsilon)$, since $\| \int_L^{\frac{t}{\varepsilon}} U(y_s)ds\|_{L^2}$ is bounded by stationarity of y_r and $\|\sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}} V(y_r)dr\|_{L^2}$ is bounded by Lemma 4.7. We compute the integration in the main region:

$$\begin{aligned}
&\int_0^L \int_0^s U(y_s)V(y_r)drds \\
=&\sum_{k=1}^L \int_{k-1}^k U(y_s)\Big(\int_0^{k-1} V(y_r)dr + \int_{k-1}^s V(y_r)dr\Big)ds \\
=&\sum_{k=1}^L \int_{k-1}^k U(y_s)ds \int_0^{k-1} V(y_r)dr + \sum_{k=1}^L \int_{\{k-1\le r\le s\le k\}} U(y_s)V(y_r)drds \\
=&\sum_{k=1}^L I(k)\sum_{l=1}^{k-1} J(l) + \sum_{k=1}^L \int_{\{k-1\le r\le s\le k\}} U(y_s)V(y_r)drds.
\end{aligned}$$

The stochastic processes $Z_k = \int_{\{k-1\le r\le s\le k\}} U(y_s)V(y_r)drds$ are shift invariant and the shift operator is ergodic with respect to the probability distribution on the path space, generated by the fOU process, hence, by Birkhoff's ergodic theorem,

$$\frac{1}{L}\sum_{k=1}^L Z_k \overset{(\varepsilon\to 0)}{\longrightarrow} \mathbf{E}Z_1 = \int_0^1 \int_0^s \mathbf{E}(U(y_s)V(y_r))drds.$$

This complete the proof. □

Lemma 5.18 *The following converges in probability:*

$$\lim_{\varepsilon\to 0} \varepsilon \sum_{k=1}^L \left(I(k)\sum_{l=1}^{k-1} J(l) - (M_{k+1}-M_k)N_k \right) = \int_1^\infty \int_0^1 \mathbf{E}(U(y_s)V(y_r))drds.$$

Proof. **A.** Summing from 1 to k of the identity (5.10), we see that

$$\sum_{l=1}^{k-1} J(l) = N_k - \hat{V}(k) + \hat{V}(1) - N_1,$$

With this and relation (5.10) for the martingale difference $M_{k+1} - M_k$, we obtain:

$$\begin{aligned}
&\sum_{k=1}^{L} I(k)\sum_{l=1}^{k-1} J(l) - (M_{k+1}-M_k)N_k\\
&=\sum_{k=1}^{L} I(k)\left(N_k - \hat V(k) + \hat V(1) - N_1\right) - \left(I(k) + \hat U(k+1) - \hat U(k)\right)N_k\\
&=\sum_{k=1}^{L} -I(k)\,\hat V(k) + \sum_{k=1}^{L} I(k)(\hat V(1) - N_1) - \sum_{k=1}^{L}(\hat U(k+1) - \hat U(k))N_k.
\end{aligned}$$

Firstly, by the shift invariance of the summands, below, and Birkhoff's ergodic theorem, we obtain

$$-\varepsilon\sum_{k=1}^{L} I(k)\hat V(k) \longrightarrow (-t)\,\mathbf{E}[I(1)\hat V(1)] = (-t)\mathbf{E}\left(\int_0^1 U(y_r)dr\int_0^\infty V(y_s)ds\right). \tag{5.12}$$

Next, since $\hat V(1) - N_1 = \mathbf{E}[\hat V(1)\,|\,\mathcal{F}_0]$,

$$\begin{aligned}
\mathbf{E}\left|\varepsilon\sum_{k=1}^{L} I(k)(\hat V(1) - N_1)\right|^2 =&\mathbf{E}\left|\varepsilon\int_0^L U(y_r)\;dr\;\mathbf{E}[\hat V(1)\,|\,\mathcal{F}_0]\right|^2\\
\lesssim&\varepsilon^2\,\mathbf{E}[\hat V(1)]^2\int_0^L\int_0^L \mathbf{E}[U(y_r)U(y_s)]\,ds\,dr,
\end{aligned}$$

which by Lemma 4.7 is of order ε.

B. It remains to discuss the convergence of

$$\varepsilon\sum_{k=1}^{L}(\hat U(k+1) - \hat U(k))N_k.$$

We do not have shift invariant and therefore break it down into increments. We change the order of summation to obtain the following decomposition

$$\begin{aligned}
&\sum_{k=1}^{L}(\hat U_{k+1} - \hat U(k))N_k\\
&=\sum_{k=1}^{L}(\hat U(k+1) - \hat U(k))\left[\sum_{j=1}^{k-1}(N_{j+1} - N_j) + N_1\right]\\
&=\sum_{j=1}^{L-1}(N_{j+1} - N_j)\sum_{k=j+1}^{L}(\hat U(k+1) - \hat U(k)) + \sum_{k=1}^{L}(\hat U(k+1) - \hat U(k))N_1\\
&=\sum_{j=1}^{L-1}(N_{j+1} - N_j)\hat U(L+1) - \sum_{j=1}^{L-1}(N_{j+1} - N_j)\,\hat U(j+1) + \left(\hat U(L+1) - \hat U(1)\right)N_1.
\end{aligned}$$

We may now apply Birkhoff's ergodic theorem to the first term, taking $\varepsilon \to 0$,

$$\lim_{\varepsilon\to 0} -\varepsilon \sum_{j=1}^{L-1}(N_{j+1}-N_j)\hat{U}(L+1) = 0,$$

in probability. By the same ergodic theorem, the second term

$$\varepsilon(N_L - N_1)\hat{U}(L+1) \longrightarrow t\,\mathbf{E}\Big(\hat{U}(2)(N_2-N_1)\Big),$$

in probability. By Proposition 5.11, $\hat{U}(j)$ is bounded in L^2, hence we obtain for the third term,

$$\varepsilon\Big|\Big(\hat{U}(L+1)-\hat{U}(1)\Big)N_1\Big|_{L^2} \lesssim \varepsilon.$$

Overall we end up with

$$\varepsilon\sum_{k=1}^{L}(\hat{U}_{k+1}-\hat{U}(k))N_k \longrightarrow -\,t\,\mathbf{E}\Big(\hat{U}(2)(N_2-N_1)\Big), \tag{5.13}$$

and so

$$\begin{aligned}&\lim_{\varepsilon\to 0}\varepsilon\sum_{k=1}^{L}\left(\sum_{l=0}^{k} I(k)J(l) - (M_{k+1}-M_k)N_k\right)\\ &= t\Big[\mathbf{E}\Big(\hat{U}(2)(N_2-N_1)\Big) - I(1)\hat{V}(1)\Big].\end{aligned} \tag{5.14}$$

C. We look a better expression for this limit,

$$\begin{aligned}\mathbf{E}\hat{U}(2)(N_2-N_1) &= \mathbf{E}\left(\int_1^\infty U(y_r)dr\Big(\hat{V}(2)-\mathbf{E}(\hat{V}(2)|\mathcal{F}_1)\Big)\right)\\ &= \mathbf{E}\left(\int_0^\infty U(y_r)dr\Big(\hat{V}(2)-\mathbf{E}(\hat{V}(2)|\mathcal{F}_1)\Big)\right).\end{aligned}$$

Since

$$\hat{V}(2)-\mathbf{E}(\hat{V}(2)|\mathcal{F}_1) - \int_0^1 V(y_s)ds = \hat{V}(2)-\hat{V}(1),$$

hence we work with

$$\mathbf{E}\left(\int_0^\infty U(y_r)dr\Big(\hat{V}(2)-\hat{V}(1)\Big)\right).$$

we compute $\sum_{k=2}^{L}\mathbf{E}[I(k)J(1)]$. We first use Remark 5.15 to write individual terms by the martingale differences. Specifically we use

$$\begin{aligned}J(1) &= (\hat{V}(1)-\hat{V}(2)) + (N_2-N_1),\\ I(k) &= M_{k+1}-M_k-(\hat{U}(k+1)-\hat{U}(k))\end{aligned}$$

for obtaining

$$
\begin{aligned}
&\sum_{k=2}^{L} I(k)J(1) - \sum_{k=2}^{L} (M_{k+1} - M_k)(N_2 - N_1) \\
=&\sum_{k=2}^{L} I(k)(\hat{V}(1) - \hat{V}(2) + \sum_{k=2}^{L} I(k)(N_2 - N_1) - \sum_{k=2}^{L} (M_{k+1} - M_k)(N_2 - N_1) \\
=&\sum_{k=2}^{L} I(k)(\hat{V}(1) - \hat{V}(2)) - \sum_{k=2}^{L} \Big(-\hat{U}(k+1) + \hat{U}(k)\Big)(N_2 - N_1) \\
=&\sum_{k=2}^{L} I(L)[\hat{V}(L+1-k) - \hat{V}(L+2-k)] - \sum_{k=2}^{L}(\hat{U}(k+1) - \hat{U}(k))(N_2 - N_1) \\
=&I(L)\hat{V}(1) - I(L)\hat{V}(L) + [\hat{U}(2)(N_2 - N_1)] - [\hat{U}(L+1)(N_2 - N_1)].
\end{aligned}
$$

In the second step, we used the stationary property by which we also have

$$\mathbf{E}[I(L)(\hat{V}(1) - \hat{V}(L))] = \mathbf{E}[I(1)\hat{V}(L)] - \mathbf{E}[I(1)\hat{V}(1)].$$

Since

$$\mathbf{E}\Big(I(L)\hat{V}(1) - \hat{U}(L+1)(N_2 - N_1)]\Big)^2 \to 0$$

by Corrolary 5.12. This concludes that

$$\sum_{k=2}^{L} \mathbf{E}(I(k)J(1)) = -\mathbf{E}[I(1)\hat{V}(1)] + \mathbf{E}[\hat{U}(2)(N_2 - N_1)] + o(\varepsilon).$$

On the other hand,

$$\sum_{k=2}^{L} \mathbf{E}(I(k)J(1)) = \int_1^L \int_0^1 \mathbf{E}(U(y_s)V(y_r))drds.$$

Since $L = \lfloor \frac{t}{\varepsilon} \rfloor \to \infty$ as $\varepsilon \to 0$, this complete the proof for Lemma 5.18. □

Now we return to Proposition 5.13. Taking $\varepsilon \to 0$ in Equation (5.11) we obtain

$$
\begin{aligned}
&\lim_{\varepsilon \to 0} \varepsilon \int_0^{\frac{t}{\varepsilon}} \int_0^s U(y_s)V(y_r)drds \\
&= \lim_{\varepsilon \to 0} \varepsilon \sum_{k=0}^{\lfloor \frac{t}{\varepsilon} \rfloor} (M_{k+1} - M_k)N_k + t\int_0^1 \int_0^s \mathbf{E}(U(y_s)V(y_r))drds + t\int_1^\infty \int_0^1 \mathbf{E}(U(y_s)V(y_r))drds \\
&= \int_0^t W_s^1 dW_s^2 + t\int_0^\infty \mathbf{E}(U(y_0)V(y_u))du,
\end{aligned}
$$

which completes the proof for Proposition 5.13.

5.2.2 Enhanced functional limit theorem (in f.d.d)

To put everything together we first need to state a lemma.

Lemma 5.19 *Suppose the stochastic processes* $(X^n, Y^n, Z^n, R_n^1, R_n^2)$ *satisfy the conditions (1)-(3) below. Then if the trio* $(X^n, Y^n, Z^n) \to (X, Y, Z)$ *in law as* $n \to \infty$*, in the càdlàg topology, so does the quadruple*

$$\Big(X^n, Y^n + R_n^1, \int Y_s^n \, dX_s^n, \, Z^n + R_n^2\Big) \to \Big(X, \, Y, \, \int Y_s dX_s, \, Z\Big).$$

The integrals are in Itô sense.

(1) Each $(X_s^n, s \in [0, t])$ *is a càdlàg* d*-dimensional semi-martingales, on a filtered probability space* $(\Omega_n, \mathcal{F}_n, P_n)$*, satisfying the following predictable uniform tightness (P-UT) condition: for every* $t > 0$*,*

$$\lim_{C \uparrow \infty} \sup_{H^n} P_n\bigg(\int_0^t H_s^n \, dX_s^n > C\bigg) = 0,$$

where the supremum is taken over all elementary processes H^n *uniformly bounded by* 1 *and adapted to* $\mathcal{F}_n$*.*

(2) Y^n *is a family of* $d' \times d$*- dimensional càdlàg stochastic processes,* Z^n *another family of multi-dimensional càdlàg stochastic processes, both adapted to* $\mathcal{F}^n$*.*

(3) R_n^1, R_n^2 *are adapted stochastic processes converging to zero in probability.*

The Lemma is essentially [JS03, Theorem 6.22] , it is only left to include the terms R_n^1 and R_n^2. On a separable probability space, convergence in finite dimensional distributions is equivalent to convergence in the Prohorov metric. The Prohorov metric is defined by $d(\mu_1, \mu_2) = \inf_{\delta > 0}\{\mu_1(A) \le \mu_2(A_\delta) + \delta, \mu_2(A) \le \mu_1(A_\delta) + \delta\}$ where A is any Borel measurable set and A_δ its δ-enlargement set. Thus X^n converges in distribution and $R_n \to 0$ implies that $X_n + R_n \to 0$ in distribution. We can then apply Theorem 6.22 from [JS03]. Without the R_n terms, the convergence of the quadruple is trivial for H_n uniformly bounded elementary processes. The rest follows from a loalization procedure, a density argument applied to integrands, the predictable uniform tightness condition on the integrands, X_n, allows reversing the order of taking limits. For $Z_t^n = R_n^1 = R_n^2 = 0$ one can also apply [KP91, Theorem 2.7].

Lemma 5.20 *If a sequence of martingales* N_s^n *is uniformly bounded in* L^2*, then they satisfy the P-UT condition.*

Proof. Given an elementary process $H^n = a_0 \mathbf{1}_{\{0\}} + \sum_{i=1}^k a_i \mathbf{1}_{(s_i, s_{i+1}]}(t)$ where $a_i \in \mathcal{F}_{s_i}$ is assumed to be bounded by 1 and $s_i \le [tn]$, we see that

$$\begin{aligned}
&P_n\bigg(\int_0^t H_s^n dN_s^n > C\bigg) \le \frac{1}{C^2}\mathbf{E}\bigg|\int_0^t H_s^n dN_s^n\bigg|^2 \\
&\le \frac{1}{C^2}\mathbf{E}\bigg|\sum_{i=1}^k a_i\big(N_{s_{i+1}\wedge(nt)} - N_{s_i \wedge(nt)}\big)\bigg|^2 \\
&\le \frac{1}{C^2}\sum_{i=1}^k \mathbf{E}\big(a_i^2\big(N_{s_{i+1}\wedge(nt)} - N_{s_i\wedge(nt)}\big)\big)^2 \le \frac{1}{C^2}\sum_{i=1}^k \mathbf{E}\big(N_{s_{i+1}\wedge(nt)} - N_{s_i\wedge(nt)}\big)^2 \\
&\le \frac{1}{C^2}\mathbf{E}\big(N_{s_{k+1}\wedge(nt)}\big)^2 - \frac{1}{C^2}\mathbf{E}\big(N_{s_1\wedge(nt)}\big)^2 \overset{(C\to\infty)}{\to} 0.
\end{aligned}$$

□

Now, recall,

$$X^\varepsilon := \left(\alpha_1(\varepsilon)\int_0^t G_1(y_s^\varepsilon)ds, \dots, \alpha_N(\varepsilon)\int_0^t G_N(y_s^\varepsilon)ds\right)$$
$$A^{i,j} = \begin{cases} \int_0^\infty \mathbf{E}(G_i(y_s)G_j(y_0))ds, & \text{if } i,j \le n, \\ 0, & \text{otherwise.}\end{cases}$$
$$X = \lim_{\varepsilon\to 0} X^\varepsilon.$$

Proposition 5.21 *Assume that $G_k \in L^{p_k}(\mu)$ satisfy Assumption 2.5. Set*

$$\mathbb{X}^{i,j,\varepsilon} = \alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s G_i(y_s^\varepsilon)G_j(y_r^\varepsilon)drds,$$
$$\mathbb{X}^{i,j} = \int_0^t X^i dX^j,$$

where the second integral is to be understood in the Itô-sense if two Wiener processes appear and in the Young sense otherwise. Then, as $\varepsilon \to 0$,

$$\mathbb{X}^\varepsilon \to \mathbb{X} + A(t-s),$$

in the sense of finite dimensional distributions . Furthermore,

$$\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon) \to \mathbf{X} = (X, \mathbb{X} + A(t-s)),$$

in finite dimensional distributions.

Proof. To apply Lemma 5.19 we first define the multi-dimensional martingales to deal with the part converging to a Wiener process. For $i \le n$ set

$$M_L^i = \sum_{k=1}^L \hat{G}_i(k) - \mathbf{E}[\hat{G}_i(k)|\mathcal{F}_{k-1}].$$

Now, by Lemma 5.17 and Lemma 5.18, for $i,j \le n$,

$$\mathbb{X}^{i,j,\varepsilon} = \varepsilon\sum_{k=1}^{[\frac{t}{\varepsilon}]}(M_{k+1}^j - M_k^j)M_k^i + t\int_0^\infty \mathbf{E}(G_i(y_s)G_j(y_0))ds + E_\varepsilon,$$

where $E_\varepsilon \to 0$ in probability. Hence, it is enough to establish convergence of $\varepsilon\sum_{k=1}^L(M_{k+1}^j - M_k^j)M_k^i$ in finite dimensional distributions. To do so we define the piecewise constant càdlàg L^2-martingales $(M_t^{j,\varepsilon}, t \ge 0)$ via

$$M_t^{j,\varepsilon} = \sqrt{\varepsilon}M^j_{[\frac{t}{\varepsilon}]}$$

to obtain

$$\varepsilon\sum_{k=1}^{[\frac{t}{\varepsilon}]}(M^j_{k+1}-M^j_k)M^i_k=\int_0^{[\frac{t}{\varepsilon}]}M^{i,\varepsilon}_s dM^{j,\varepsilon}_s.$$

According to Remark 5.15,

$$M^{j,\varepsilon}_{\lfloor\frac{t}{\varepsilon}\rfloor}=\sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}}G^j(y_s)ds+o(\sqrt{\varepsilon}).$$

Thus, by Lemma 5.2, $(M^{j,\varepsilon}_t)_{j=1,\dots,n}$ converge jointly in some Hölder space, which implies of course convergence in the Skorokhod topology. Therefore it is only left to establish the uniform L^2 bounds, which follows from Corollary 5.12 or as below:

$$\|M^{j,\varepsilon}_t\|_{L^2}=\sqrt{\varepsilon}\|M_{[\frac{t}{\varepsilon}]}\|_{L^2}=\sqrt{\varepsilon}\left\|\int_0^{\frac{t}{\varepsilon}}G_j(y_s)ds+o(\sqrt{\varepsilon})\right\|_{L^2}\lesssim 1.$$

Now by Lemma 5.20 our martingales satisfy the predictable uniform tightness condition. We apply Lemma 5.5, Lemma 5.19 with $X^\varepsilon=Y^\varepsilon=M^\varepsilon=(M^{1,\varepsilon},\dots M^{n,\varepsilon})$, $Z^\varepsilon=(X^k,\mathbb{X}^{i,j})$ for $k>n$ and i,j such that $i\vee j>n$ and $R^1_n=\mathbf{Er}_1(\varepsilon)$ as in Proposition 5.13 to conclude the claim. □

5.3 Tightness of iterated integrals

In this section we establish moment bounds on the iterated integrals to show that the above proven convergence takes place in a suitable space of rough paths.

Now, let G_i, G_j be two functions in L^2 with Hermite ranks m_{G_i} and m_{G_j} respectively. Set $\alpha_i=\alpha(\varepsilon,H^*(m_{G_i}))$ and $\alpha_j(\varepsilon)=\alpha(\varepsilon,H^*(m_{G_j}))$. To obtain tightness for

$$\mathbb{X}^{i,j,\varepsilon}(t)=\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s G_i(y^\varepsilon_s)G_j(y^\varepsilon_r)drds$$

we assume a certain decay in the Hermite expansion of our functions. If G_i and G_j are in finite chaos, $\mathbf{E}\big(\mathbb{X}^{i,j,\varepsilon}\big)^p$ is the sum of a finite number of terms which are controlled by integrals of the form,

$$\alpha_i(\varepsilon)^p\alpha_j(\varepsilon)^p\int_0^t\cdots\int_0^t\mathbf{E}\left(\prod_{k=1}^{2p}H_{m_k}(y^\varepsilon_{s_k})\right)ds_1\dots ds_{2p},$$

where H_{m_k} are Hermite polynomials with $m_k\geq\min(m_{G_i},m_{G_j})$and $s_1\leq s_2\leq\cdots\leq s_{2p}$. Let us try to compute $\mathbf{E}\Big(\prod_{k=1}^{2p}H_{m_k}(y^\varepsilon_{s_k})\Big)$. If X_i are random variables, there is the product formula

$$\mathbf{E}\left[\prod_{k=1}^{2p}X_k\right]=\sum_{\pi\in P(1,\dots,2p)}\prod_{B\in\pi}\mathbf{E}^c[X_k:k\in B],$$

where π denotes a partition of $\{1,2,\dots,2p\}$ and $\mathbf{E}^c[X_k:k\in B]$ denotes the joint cumulant of the variable X_k with k in a block B. These joint cumulants can also be given by linear combination of the form $\prod_{B\in\pi}\mathbf{E}[\prod_{k\in B}X_k]$. Gaussian processes have vanishing third or higher order cumulants, and since expectations for products of Hermite polynomials of X_k are related to these cumulants, there are simpler ways for computing them. Using convolution with heat kernels fewer terms from the graph remain. The expectations of the product are then given by product of expectations of all pairings, and summed over all such graphs. For a particular graph, we denote by $n(l,k)$ the number of edges connecting l to k, so it takes values in $\{0,1,\dots\min(m_l,m_k)\}$, and consider the pairings in an ordered way so that each pairing is counted only once. We have $\sum_{k=1}^{2p}n(l,k)=m_l$. Since edges are only allowed to connect with different nodes we have $n(k,k)=0$ for every k. We then observe that $\mathbf{E}\Big(\prod_{k=1}^{2p}H_{m_k}(y^\varepsilon_{s_k})\Big)$ is the sum over the finite number of graphs of pairings, for any given graph this is

$$\prod_{k=1}^{2p}\prod_{l=k+1}^{2p}\mathbf{E}(y^\varepsilon_{s_k}y^\varepsilon_{s_l})=\prod_{k=1}^{2p}\prod_{\{l:l>k,\,l\in\Gamma_k\}}\varrho(s^\varepsilon_l-s^\varepsilon_k)^{n(l,k)},$$

where Γ_k denotes the subgraph of nodes connected to k from nodes in the forward direction. A *complete pairing* of a graph with $2p$ nodes, with respectively m_k edges, is a graph in which each edge from a node is connected to an edge with a different nod.

Lemma 5.22

1. *Let Γ denote a graph of pairings of edges (no self-connection is allowed). Define:*

$$I(\varepsilon,p)=\overbrace{\int_0^t\cdots\int_0^t}^{2p}\left|\prod_{\{(s_k,s_l)\}\in\Gamma}\mathbf{E}(y^\varepsilon_{s_k}y^\varepsilon_{s_l})\right|ds_1\dots ds_{2p}.$$

Let m_k denote the number of edges issuing from the node s_k. Then

$$I(\varepsilon,p)\lesssim\prod_{k=1}^{2p}\sqrt{t\int_{-t}^t|\varrho^\varepsilon(s)|^{m_k}\,ds}\lesssim\prod_{k=1}^{2p}\frac{t^{H^*(m_k)\vee\frac12}}{\alpha(\varepsilon,H^*(m_k))}. \tag{5.15}$$

2. *Let $G_i,G_j:\mathbf{R}\to\mathbf{R}$ be functions in finite chaos with Hermite ranks m_{G_i} and m_{G_j} respectively. Then,*

$$\begin{aligned}\|\mathbb{X}^{i,j,\varepsilon}\|_{L^p}:=\alpha_i(\varepsilon))\alpha_j(\varepsilon)\left\|\int_0^t\int_0^s G_i(y^\varepsilon_s)G_j(y^\varepsilon_r)drds\right\|_{L^p}\\ \lesssim t^{H^*(m_{G_i})\vee\frac12+H^*(m_{G_j})\vee\frac12}.\end{aligned}$$

Proof. Different graphs yield different asymptotics, the 'worst' graph is the one with exactly m edges at each node, and all edges of a given node are linked to the same node. For a general graph, let us deal with the first variable s_1. We first count forward and observe

$$\prod_{\{(s_k,s_l)\}\in\Gamma}\mathbf{E}(y^\varepsilon_{s_k}y^\varepsilon_{s_l})=\prod_{k=1}^{2p}\prod_{l=k+1}^{2p}\mathbf{E}(y^\varepsilon_{s_k}y^\varepsilon_{s_l})=\prod_{k=1}^{2p}\prod_{\{l:l>k,\,l\in\Gamma_k\}}(\varrho^\varepsilon(s_l-s_k))^{n(l,k)},$$

where Γ_k denotes the subgraph of nodes connected to k, from the nodes with index greater than k. Using Hölder's inequality we obtain

$$\begin{aligned}
&\int_0^t \prod_{\{k:k>1,\,k\in\Gamma_1\}} |\varrho^\varepsilon(s_1-s_k)|^{n(1,k)} ds_1 \\
&\le \prod_{\{k:k>1,\,k\in\Gamma_1\}} \left(\int_0^t |\varrho^\varepsilon(s_1-s_k)|^{m_1}\, ds_1\right)^{\frac{n(1,k)}{m_1}} \\
&\le \int_{-t}^t |\varrho^\varepsilon(s_1)|^{m_1} ds_1.
\end{aligned}$$

We have used $\sum_{\{k>1:k\in\Gamma_1\}} n(1,k) = m_1$, the number of edges at node 1. We then peel off the integrals layer by layer, and proceed with the same technique to the next integration variable. For example suppose the remaining integrator containing s_2 has the combined exponent $\tau_2 = \sum_{k=2}^{2p} n(2,k)$, ($\tau_1 = m_1$). By the same procedure as for s_1 we score a factor

$$\int_{-t}^t |\varrho^\varepsilon(s_2)|^{\tau_2} ds_2.$$

By induction and putting estimates for each integral together,

$$\overbrace{\int_0^t \cdots \int_0^t}^{2p} \prod_{k=1}^{2p} \prod_{\{l:l>k,\,l\in\Gamma_k\}} (\varrho^\varepsilon(s_l-s_k))^{n(l,k)}\, ds_1 \dots ds_{2p} \lesssim \prod_{k=1}^{2p} \int_{-t}^t |\varrho^\varepsilon(s)|^{\tau_k} ds.$$

Following [BH02], we reverse the procedure in the estimation for the integral kernel, take ξ_k to be the number of edges connected to nodes in the backward direction, then $\xi_k = \sum_{l=1}^k n(l,k)$ and the same reasoning leads to the followign estimate:

$$\overbrace{\int_0^t \cdots \int_0^t}^{2p} \prod_{k=1}^{2p} \prod_{\{l:l>k,\,l\in\Gamma_k\}} (\varrho^\varepsilon(s_l-s_k))^{n(l,k)}\, ds_1 \dots ds_{2p} \lesssim \prod_{k=1}^{2p} \int_{-t}^t |\varrho^\varepsilon(s)|^{\xi_k} ds.$$

Since $\tau_k + \xi_k = m_k$ and

$$\int_{-t}^t |\varrho^\varepsilon(s)|^{\tau_k} ds \int_{-t}^t |\varrho^\varepsilon(s)|^{\xi_k} ds \le 2t \int_{-t}^t |\varrho^\varepsilon(s)|^{m_k} ds,$$

and therefore

$$\left(\overbrace{\int_0^t \cdots \int_0^t}^{2p} \prod_{k=1}^{2p} \prod_{\{l:l>k,\,l\in\Gamma_k\}} (\varrho^\varepsilon(s_l-s_k))^{n(l,k)}\, ds_1 \dots ds_{2p}\right)^2 \lesssim \prod_{k=1}^{2p} \left(t \int_{-t}^t |\varrho^\varepsilon(s)|^{m_k} ds\right).$$

By Lemma 3.4 for each k,

$$\alpha(\varepsilon, H^*(m_k))^2 t \int_{-t}^{t} |\varrho^\varepsilon(s)|^{m_k} ds \lesssim \left(t^{H^*(m_k)\vee\frac{1}{2}}\right)^2,$$

the first part of the lemma follows.

Next let us consider $G_i = H_k$ and $G_j = H_l$. Then we are in a position to apply the first part of the lemma:

$$\begin{aligned}
&\mathbf{E}\left(\int_0^t\int_0^s H_k(y_s^\varepsilon)H_l(y_r^\varepsilon)dr ds\right)^p \\
&= \int_0^t\int_0^{s_p} \cdots \int_0^t\int_0^{s_1} \mathbf{E}\left(\prod_{i=1}^p H_k(y_{s_i}^\varepsilon)H_l(y_{r_i}^\varepsilon)\right) \prod_{i=1}^p dr_i\, ds_i \\
&\le \sum_\Gamma \int_0^t\int_0^{s_p} \cdots \int_0^t\int_0^{s_1} \prod_{k=1}^{2p} \prod_{\{l:l>k,\, l\in\Gamma_k\}} \varrho^\varepsilon(s_l - s_k)^{n(l,k)} \prod_{i=1}^p dr_i\, ds_i \\
&\lesssim C_{k,l} \frac{t^{p(H^*(k)\vee\frac{1}{2})+p(H^*(l)\vee\frac{1}{2})}}{\alpha(\varepsilon, H^*(k))^p \alpha(\varepsilon, H^*(l))^p}\ ,
\end{aligned}$$

where the summation is over all graphs of complete parings and $C_{k,l}$ denotes the number of graphs needed for computing the expectations.

For $G_i = \sum_{k=m_{G_i}}^N a_{i,k} H_k$ and $G_j = \sum_{k=m_{G_j}}^N a_{i,k} H_k$, we expand the products in the multiple integrales. Each summand then has exactly p factors from G_i, for those $k \ge m_{G_i}$, and p from G_j for those $k \ge m_{G_j}$. Splitting them accordingly we have,

$$\|\mathbb{X}^{i,j,\varepsilon}\|_{L^p} \lesssim \alpha_i(\varepsilon)\left(\prod_{k_1=1}^p t\int_{-t}^t |\varrho^\varepsilon(s)|^{m_{k_1}} ds\right)^{\frac{1}{p}} \alpha_j(\varepsilon)\left(\prod_{k_2=1}^p t\int_{-t}^t |\varrho^\varepsilon(s)|^{m_{k_2}}\, ds\right)^{\frac{1}{p}}. \quad (5.16)$$

where $m_{k_1} \ge m_{G_i}$ and $m_{k_2} \ge m_{G_i}$. The treatment for the two products are the same. Let us consider the first factor on the right hand side. Since

$$\int_{-t}^t |\varrho^\varepsilon(s)|^m ds$$

decreases with m, then those terms with the Hermite ranks of G_i and G_j as exponents give the fastest possible blow up, or the slowest convergence rate. Since both G_i and G_j belong to the finite chaos,$(\|\mathbb{X}^{i,j,\varepsilon}\|_{L^p})^p$ is the sum of finite terms of the form

$$\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s \cdots \int_0^t\int_0^s \prod_l \mathbf{E}[H_k(y^\varepsilon s_i)H_l(y_{r_l}^\varepsilon)] ds_l\, dr_l,$$

each of these has the same type type bound (by the previous computation), with the constant $C_{k,l}$ in front of the relevant expansions uniformly bounded. We may conclude that $\|\mathbb{X}^{i,j,\varepsilon}\|_{L^p} \lesssim t^{(H^*(k)\vee\frac{1}{2})+(H^*(l)\vee\frac{1}{2})}$, finishing the proof. □

For functions not belonging to the infinite chaos we must count the number of graphs in the computation, and need some assumptions. Let $M(\{m_1,\dots,m_{2p}\})$ denote the cardinality of complete pairings of a graph ith $2p$ nodes, with respectively m_k edges. In [Taq77] it was shown that

$$M(m_1,m_2,\dots,m_{2p}) \le \prod_{k=1}^{2p}(2p-1)^{\frac{m_k}{2}}\sqrt{m_k}.$$

This leads to Assumption 2.5 (1), which restricts the G_i to the class of functions whose coefficients in the Hermite expansion decays sufficiently fast.

Proposition 5.23 *Suppose that $G_k \in L^{p_k}$ and satisfies Assumption 2.5. Then one has for each $i,j=1,\dots N$,*

$$\left\|\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\!\!\int_0^s G_i(y_s^\varepsilon)G_j(y_r^\varepsilon)drds\right\|_{L^p} \lesssim t^{H^*(m_{G_i})\vee\frac12+H^*(m_{G_j})\vee\frac12}.$$

In particular the family

$$\left\{\alpha_k(\varepsilon)\int_0^t G_k(y_s^\varepsilon)ds, \alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s G_i(y_s^\varepsilon)G_j(y_r^\varepsilon)drds, k,i,j=1,\dots N\right\}$$

is tight in $\mathscr{C}^\gamma$ for any $\gamma\in\left(\frac13,\left(H^*(m_{G_i})\vee\frac12+H^*(m_{G_j})\vee\frac12\right)-\frac1p\right)$.

Proof. Using the above estimates we compute, and the fact that $\varrho(s)>0$,

$$\begin{aligned}&\mathbf{E}\left(\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^t\int_0^s G_i(y_s)G_j(y_r)drds\right)^p\\ &\le \alpha_i(\varepsilon)^p\alpha_j(\varepsilon)^p\left|\mathbf{E}\left(\int_0^t\!\!\int_0^s\sum_{k,k'}c_{i,k}c_{j,k'}H_k(y_s)H_{k'}(y_r)drds\right)^p\right|.\end{aligned}$$

We estimate the integrals on the right hand side:

$$\begin{aligned}&\left|\mathbf{E}\left(\int_0^t\!\!\int_0^s\sum_{k,k'}c_{i,k}c_{j,k'}H_k(y_s^\varepsilon)H_{k'}(y_r^\varepsilon)drds\right)^p\right|\\
&\le\left|\sum_{k_1,\dots k_p=m_{G_i}}^\infty\sum_{k_1',\dots k_p'=m_{G_j}}^\infty\prod_{l=1}^p c_{i,k_l}c_{j,k_l'}\int_0^t\!\!\int_0^s\cdots\int_0^t\!\!\int_0^s\prod_{l=1}^p\mathbf{E}\Big(H_{k_l}(y_{s_l}^\varepsilon)(H_{k_l'}(y_{r_l}^\varepsilon)\Big)dr_lds_l\right|\\
&\le\sum_{k_1,\dots k_p=m_{G_i}}^\infty\sum_{k_1',\dots k_p'=m_{G_j}}^\infty\prod_{l=1}^p|c_{i,k_l}c_{j,k_l'}|\int_0^t\!\!\int_0^s\cdots\int_0^t\!\!\int_0^s\sum_\Gamma\prod_{v=1}^{2p}\prod_{\{u:u>v,\,u\in\Gamma_v\}}\varrho^\varepsilon(s_u-s_v)^{n(u,v)}ds_uds_v\\
&\le\sum_{k_1,\dots k_p=m_{G_i}}^\infty\sum_{k_1',\dots k_p'=m_{G_j}}^\infty\prod_{l=1}^p|c_{i,k_l}c_{j,k_l'}|\overbrace{\int_0^t\cdots\int_0^t}^{2p}\sum_\Gamma\prod_{v=1}^{2p}\prod_{\{u:u>v,\,u\in\Gamma_v\}}\varrho^\varepsilon(s_u-s_v)^{n(u,v)}ds_uds_v.\end{aligned}$$

We then apply estimates from the first part of Lemma 5.22,

$$\mathbf{E}\Big(\alpha_i(\varepsilon)\alpha_j(\varepsilon)\int_0^s G_i(y_s^\varepsilon)G_j(y_r^\varepsilon)drds\Big)^p$$
$$\lesssim t^{p\left(H^*(m_{G_i})\vee\frac{1}{2}+H^*(m_{G_j})\vee\frac{1}{2}\right)}\sum_{k_1,\dots k_p=m_{G_i}}^\infty\sum_{k_1',\dots k_p'=m_{G_j}}^\infty\prod_{l=1}^p|c_{i,k_l}c_{j,k_l'}|M(k_1,\dots,k_p,k_1',\dots,k_p')$$
$$\lesssim t^{p\left(H^*(m_{G_i})\vee\frac{1}{2}+H^*(m_{G_j})\vee\frac{1}{2}\right)}\sum_{k_1,\dots k_p=m_{G_i}}^\infty\sum_{k_1',\dots k_p'=m_{G_j}}^\infty\prod_{l=1}^p|c_{i,k_l}c_{j,k_l'}|\sqrt{k_l!k_l'!}(2p-1)^{\frac{k_l+k_l'}{2}},$$

By assumption, the double power series is finite. This proves the required moment bounds for the second order process. For tightness we argue by Lemmas 4.7 and 3.11, concluding the proof. □

5.4 Weak convergence in $\mathscr{C}^\gamma$, concluding Theorem B.

We are ready to show weak convergence of the rough path lifts. Denote by X^W the limiting Wiener processes in Lemma 5.5. Then, for $i,j\le n$, we can form the Itô integrals $\int X_s^i dX_s^j$ and denote it by $\mathbb{X}^{i,j}$. If either the ith or the jth component limit is not given by a Wiener process, we will see that $\mathbb{X}_{0,t}^{i,j,\varepsilon}=\int_0^t X^{i,\varepsilon}dX^{j,\varepsilon}$ converges weakly to a process with higher regularity, which, as a rough path, does not exert any influence on the interpretation of the rough integral, and therefore has no effect on the effective equation.

Theorem 5.24 *Assume that $G_k\in L^{p_k}(\mu)$ satisfy Assumption 2.5, then*

$$\mathbf{X}^\varepsilon=(\mathbf{X}^\varepsilon,\mathbb{X}^\varepsilon)\to\mathbf{X}=(X,\mathbb{X}+A(t-s)),$$

weakly in $\mathscr{C}^\gamma$ for $\gamma\in(\frac{1}{3},\frac{1}{2}-\frac{1}{\min_{k\le n}p_k})$.

Proof. By Proposition 5.21 $\mathbf{X}^\varepsilon$ converges in finite dimensional distributions and by Lemma 3.10, Lemma 3.11, Proposition 5.23 and Lemma 4.7 show that the convergence takes place is the respective Hölder spaces. □

With Theorem 5.24 together with Proposition 5.5 we conclude Theorem B.

6 Multi-scale homogenization theorem

We are now in the position to complete the proof for the multi-scale homogenization theorem for the long range dependent case, this is Theorem A, which we formulate here how it is proved.

Theorem 6.1 *Let $H\in\left(\frac{1}{2},1\right)$, $f_k\in\mathcal{C}_b^3(\mathbf{R}^d;\mathbf{R}^d)$, and G_k satisfies Assumption 2.5. Then, the following statements hold.*

1. *The solutions x_t^ε of (2.4) converge weakly in $\mathcal{C}^\gamma$ on any finite interval and for any $\gamma\in(\frac{1}{3},\frac{1}{2}-\frac{1}{\min_{k\le n}p_k})$.*

2. *The limit solves the rough differential equation*

$$dx_t = f(x_t)d\mathbf{X}_t \quad x_0 = x_0 \tag{6.1}$$

Here $f = (f_1, \dots, f_N)$ *and* $\mathbf{X} = (X, \mathbb{X}_{s,t} + (t-s)A)$ *is a rough path over* $\mathbf{R}^N$*. (They will be described in Theorem B below.)*

3. *Equation (6.1) is equivalent to the stochastic equation below:*

$$dx_t = \sum_{k=1}^{n} f_k(x_t) \circ dX_t^k + \sum_{l=n+1}^{N} f_l(x_t) dX_t^l, \quad x_0 = x_0,$$

where $\circ$ *denotes Stratonovich integral, otherwise a Young integral.*

Proof. We want to formulate our slow/fast random differential equation as a family of rough differential equations, such that the drivers converge in the rough path topology. Using the continuity of the solution map, we obtain weak convergence of the solutions to a rough differential equation. In the appendix, we interpret the rough differential equation as a mixed Stratonovich and Young integral equation and the coefficients of the equation will be computed.

Let us set $F : \mathbf{R}^d \to \mathbb{L}(\mathbf{R}^N, \mathbf{R}^d)$ as below:

$$F(x)(u_1, \dots, u_m) = \sum_{k=1}^{N} u_k f_k(x),$$

and for the standard o.n.b. $\{e_i\}$ of $\mathbf{R}^d$ we set $F_i(x) = F(x)(e_i)$. If we set

$$G^\varepsilon = \Big(\alpha_1(\varepsilon) G_1, \dots, \alpha_N(\varepsilon) G_N\Big),$$

we may then write the equation as follows: $\dot{x}_t^\varepsilon = F(x_t^\varepsilon) G^\varepsilon(y_t^\varepsilon)$. Now define the rough path $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon)$, where

$$\begin{aligned} X_t^\varepsilon &= \Big(\alpha_1(\varepsilon) \int_0^t G_1(y_s^\varepsilon) ds, \dots, \alpha_N(\varepsilon) \int_0^t G_N(y_s^\varepsilon) ds\Big) \\ &= (X_t^{1,\varepsilon}, \dots, X_t^{N,\varepsilon}) \\ \mathbb{X}_{s,t}^{i,j,\varepsilon} &= \int_s^t (X_r^{i,\varepsilon} - X_s^{i,\varepsilon}) dX_r^{j,\varepsilon}. \end{aligned}$$

We may therefore rewrite our equation as a rough differential equation with respect to $\mathbf{X}^\varepsilon$:

$$dx_t^\varepsilon = F(x_t^\varepsilon) d\mathbf{X}^\varepsilon(t).$$

with covariance as specified in Theorem B.

By Theorem 5.24, $\mathbf{X}^\varepsilon$ converges to $\mathbf{X} = (X, \mathbb{X} + (t-s)A)$ in $\mathscr{C}^\gamma$ where $\gamma \in (\frac{1}{3}, \frac{1}{2} - \frac{1}{\min_{k \le n} p_k})$.

Since $\gamma > \frac{1}{3}$ by Assumption 2.5, We may apply the continuity theorem for rough differential equations, Theorem 3.8, to conclude that the solutions converge to the solutions of the rough differential differential equation

$$\dot{x}_t = F(x_t) d\mathbf{X}_t.$$

Since F belongs to $\mathcal{C}_b^3$, this is well posed in the rough path equation sense. We completed the proof for the convergence. □

7 Appendix

7.1 Interpreting the effective dynamics by classical equations

We now explain what the limiting equation means in the classical sense. Although it merits a verification for the integrand consists of correlated Winer with drift block and correlated Hermite block, the answer is obvious and believed for those working in rough path theory. For the slow/fast and homogenization community this is the mysterious part, and in any case the multi-dimension path notation needs some classification. Our set up is the following.

Assumption 7.1 *Let $X_t = (X_t^W, X_t^Z)$ where X_t^W is a multi-dimensional possibly correlated Wiener process and X_t^Z a multi-dimensional Hermite process. The two components X_t^W and X_t^Z are not correlated, we denote by A the block matrix*

$$A := \begin{pmatrix} \mathrm{Cov}(X^W) & 0 \\ 0 & 0 \end{pmatrix}.$$

We write $A^{i,j}$ for the component of A. We are concerned with the classical interpretation for the rough differential equation

$$\dot{x}_t = F(x_t) d\mathbf{X}_t,$$

where $F : \mathbf{R}^d \to \mathbb{L}(\mathbf{R}^N, \mathbf{R}^d)$ is a BC^3 map and $\mathbf{X} = (X, \mathbb{X} + (t-s)A)$ where $\mathbb{X} = (\mathbb{X}^{i,j})$ denotes the 'Canonical lift' of X,

$$\mathbb{X}_t^{i,j} = \int_0^t X_s^i \, dX_s^j$$

interpreted as Itô integrals if $i, j \le n$, otherwise as Young integrals.

According to the general theorems on rough differential equation there exists a unique solution in the controlled rough path space $D_X^{2\alpha}([0,1]; \mathbf{R}^d)$ where $\alpha > \frac{1}{3}$. The solution exists global in time and the full controlled process is given by $(x_s, F(x_s))$. See [Lyo94, FH14].

We begin with setting the notation and at the same time explaining the raison d'être for the definition of rough integrals. Given a rough path $(X, \mathbb{X})$ and a controlled rough path in $\mathcal{D}_X^{2\alpha}$ is a pair of processes (Y, Y') with the properties

$$Y_{s,t} = Y_s' X_{s,t} + R_{s,t},$$

where $Y' \in \mathcal{C}^\alpha(\mathbb{L}(\mathbf{R}^d, \mathbb{L}(\mathbf{R}^N, \mathbf{R}^d))$ and the two parameter function R, satisfies $\|R\|_{2\alpha} < \infty$. Let $\alpha > \frac{1}{3}$. The rough path integral is given by the enhanced Riemann sums

$$\int_s^t Y \, d\mathbf{X} = \lim \sum_{[u,v] \in \mathcal{P}} Y_u X_{u,v} + Y_u' \mathbb{X}_{u,v}.$$

The remainder $R_{s,t}$'s contribution is of order $|t-s|^{3\alpha}$ term, which sums to zero for $\alpha > \frac{1}{3}$ and can be ignored. The limit is along any sequences of partitions with mesh converging to zero.

We therefore seek two processes L_s, R and an expression

$$F(x_t) - F(x_s) = L_s(X_t - X_s) + R_{s,t},$$

where $\sup_{s\neq t, s,t\in[0,1]} |\frac{R_{s,t}}{|t-s|^{2\alpha}}| \le C$. Taylor expanding $F(x_t)$, one can deduce that $L_s = DF(x_s)F(x_s)$. By Taylor's theorem,

$$\begin{aligned} F(x_t) &= F(x_s) + DF(x_s)(x_t - x_s) \\ &\quad + \frac{1}{2}\int_0^t (1-u)\mathrm{H}ess(F)(x_s + u(x_t - x_s))(x_t - x_s, x_t - x_s)du. \end{aligned}$$

Here $DF(x)(v) = \sum_{i=1}^d \frac{\partial F}{\partial x_i} v_i$, $\mathrm{H}essF$ is the Hessian of F, and

$$\mathrm{H}essF(x)(e,v) = \sum_{i,j=1}^d \frac{\partial^2 F}{\partial x_i \partial x_j} e_i v_j.$$

The last terms in the Taylor expansion is of order $\mathcal{C}^{2\alpha}$, one α each from $x_t - x_s$, and so goes into the R term. We express $x_t - x_s$ in terms of $X_t - X_s$:

$$\begin{aligned} DF(x_s)(x_t - x_s) &= DF(x_s)(\int_s^t F(x_r)dX_r) \\ &\sim DF(x_s)(F(x_s)(X_t - X_s)) + R^1_{s,t}. \end{aligned}$$

The R^1 term is of order $|t-s|^{2\alpha}$.

$$x_t - x_s \sim \sum F(x_u)(X_v - X_u) + DF(x_u)(F(x_u)\mathbb{X}_{v,u}).$$

If $\mathbf{X} = (W, \mathbb{W})$ to be the standard Brownian motion with its Itô lift, i.e. $\mathbb{W}^{i,j}_{s,t} = \int_s^t (W^i_r - W^i_s)dW^j_r$ then

$$\sum DF(x_u)F(x_u)\mathbb{W}_{u,v} \to 0$$

in probability. This means the equation is the Itô integral. If X is a correlated Wiener process, choose U such that $U^T U = A$. Suppose that the Wiener process block is: $(X, \mathbb{X} + At) = (UW, \int_s^t U(W_r - W_s)dUW_r) + \frac{1}{2}At)$. This leads to an SDE with Stratonovich integral

$$dx_t = F(x_t)U \circ dW_t.$$

This comes from the following fact. Let $\mathbf{X} = (X, \mathbb{X})$, $Z = (Z, \mathbb{Z})$ be two rough paths in $\mathcal{C}^\alpha$, and (Y, Y') is controlled by X, i.e. $(Y, Y') \in \mathcal{D}^{2\alpha}_X$. If g is a 2α-Hölder continuous functions such that

$$Z_t = X_t. \qquad \mathbb{Z}_{s,t} = \mathbb{X}_{s,t} + g(t) - g(s),$$

then according to [FH14], $(Y, Y') \in \mathcal{D}^{2\alpha}_Z$, and

$$\begin{aligned} \int Y\, d\mathbf{Z} &= \lim \sum_{[s,t]\in\mathcal{P}} Y_s Z_{s,t} + Y'_s \mathbb{Z}_{s,t} \\ &= \lim \sum_{[s,t]\in\mathcal{P}} (Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t} + (g(t) - g(s))Y'_s) \\ &= \int Y_s\, d\mathbf{X}_s + \int Y'_s dg. \end{aligned}$$

the last integral is a Young integral which is well defined since $Y' \in C^\alpha$ and $g \in C^{2\alpha}$.

For the Hermite component Z the secondary process makes no visible contribution in the limit of the enhanced Riemann sum and the rough integral $\int_0^t f_i(x_s) d\mathbf{X}_s^i$ agrees with the Young integral. Take X^i or X^j, whose sum of regularity is then greater than 1, the secondary process $\int_s^t X^i_{s,r} dX^j_r$, thus the enhanced Riemann sum limit is a Young integral. We may now conclude.

Remark 7.2 The solution of the rough differential equation $\dot{x}_t = F(x_t) d\mathbf{X}_t$, where $\mathbf{X}_t = (UW_t + \frac{1}{2}At, Z_t)$ with canonical lift, agrees almost surely with the solution

$$dx_t = F_1(f_t) U \circ dW_t + F_2(x_t) dZ_t,$$

where F_1 denotes F restricted to its first n components and F_2 denotes F restricted to the remaining $N-n$ components. By the standard theorem, also, the mixed integral equations is well posed, global, and continuous in the initial data.

7.2 Auto-correlation and moments of fOU

Lemma 3.3 Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$. For any $t \neq s$,

$$|\varrho(s,t)| \lesssim 1 \wedge |t-s|^{2H-2}. \tag{7.1}$$

Proof. We give an indicative proof and fix $H > \frac{1}{2}$. It is sufficient to prove this for $|t-s|$ large, then

$$\begin{aligned} \frac{\varrho(s,t)}{\sigma^2 H(2H-1)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{(2t-v)\wedge(2s+u)} e^{-(s+t-u)} |v|^{2H-2} \, du \, dv \\ &= \int_{-\infty}^{\infty} e^{-|v-(t-s)|} |v|^{2H-2} dv = \int_{-\infty}^{\infty} |v+t-s|^{2H-2} \, e^{-|v|} \, dv. \end{aligned}$$

The integration region breaks up into three:

$$(A) \qquad \frac{1}{2}|t-s| \le |v| \le 2|t-s|, \qquad (B) \quad |v| \le \frac{1}{2}|t-s| \quad \text{or} \quad |v| \ge 2|t-s|.$$

In region (A),

$$\int_{\frac{1}{2}|t-s| \le |v| \le 2|t-s|} |v+t-s|^{2H-2} \, e^{-\frac{1}{2}|t-s|} \, dv \le |t-s|^{2H-1} e^{-\frac{1}{2}|t-s|}.$$

For $|t-s|$ large this gives better bound than $|t-s|^{2H-2}$. In region (B), since $2H-2<0$,

$$\int_B |v+t-s|^{2H-2} \, e^{-|v|} \, dv \le |t-s|^{2H-2} \int_{-\infty}^{\infty} e^{-|v|} \, dv,$$

giving the correct rate. For $H < \frac{1}{2}$, we have on one hand the large time asymptotics from [CKM03]: $\varrho(s) = 2\sigma^2 H(2H-1) s^{2H-2} + O(s^{2H-4})$, on the other hand $\mathbf{E}(y_s y_t) \le \|y_s\|_{L^2} \|y_y\|_{L^2} \le 1$, showing that ϱ is locally bounded and concluding the proof. □

Proof for Lemma 3.4. This comes down to the following statement: we only need to show that for $\varepsilon \in (0, \frac{1}{2}]$, the following holds uniformly :

$$\left(\int_0^t \int_0^t |\varrho^\varepsilon(u,r)|^m \,dr\,du\right)^{\frac{1}{2}} \lesssim \begin{cases} \sqrt{t\varepsilon \int_0^\infty \varrho^m(s)ds}, & \text{if} \quad H^*(m) < \frac{1}{2}, \\ \sqrt{t\varepsilon |\ln\left(\frac{t}{\varepsilon}\right)|}, & \text{if} \quad H^*(m) = \frac{1}{2}, \\ \left(\frac{t}{\varepsilon}\right)^{H^*(m)-1}, & \text{if} \quad H^*(m) > \frac{1}{2}. \end{cases} \tag{7.2}$$

Proof. We first observe that

$$\int_0^\infty \varrho^m(s)ds < \infty \quad \Longleftrightarrow \quad H^*(m) < \frac{1}{2} \quad \Longleftrightarrow \quad H < 1 - \frac{1}{2m}. \tag{7.3}$$

By a change of variables and using estimate (3.6) on the decay of the auto correlation function (3.6),

$$\begin{aligned}\int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} |\varrho(|r-u|)|^m drdu &= 2\frac{t}{\varepsilon}\int_0^{\frac{t}{\varepsilon}} |\varrho(s)|^m ds \\ &\lesssim \begin{cases} \frac{t}{\varepsilon}\int_0^\infty \varrho^m(s)ds, & \text{if } H^*(m) < \frac{1}{2}, \\ \left(\frac{t}{\varepsilon}\right)^{2H^*(m)}, & \text{otherwise.} \end{cases} ,\end{aligned}$$

For the case $H^*(m) = \frac{1}{2}$ we use

$$\int_0^{\frac{t}{\varepsilon}} |\varrho(s)|^m ds \le \int_0^{\frac{T}{\varepsilon}} |\varrho(s)|^m ds \lesssim \int_0^{\frac{T}{\varepsilon}} (1 \wedge \frac{1}{s})ds \lesssim |\ln\left(\frac{T}{\varepsilon}\right)| \lesssim |\ln\left(\frac{1}{\varepsilon}\right)|.$$

To complete the proof we observe that by a simple change of variables,

$$\int_0^t \int_0^t |\varrho^\varepsilon(u,r)|^m \,dr\,du = \varepsilon^2 \int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} |\varrho(u,r)|^m \,dr\,du,$$

concluding the proof. □

Lemm 3.5. For any $\gamma \in (0, H)$, $p > 1$, and $s, t \in [0, \infty)$, the following estimates hold:

$$\|y_s - y_t\|_{L^p} \lesssim 1 \wedge |s-t|^H, \qquad \mathbf{E}\sup_{s\neq t}\left(\frac{|y_s - y_t|}{|t-s|^\gamma}\right)^p \lesssim C(\gamma,p)^p M.$$

Proof. We use the Ornstein-Uhlenbeck equation

$$y_s - y_r = -\int_r^s y_u du + B_s - B_r,$$

to obtain $\mathbf{E}|y_s - y_r|^2 \lesssim (s-r)^2 \mathbf{E}|y_1|^2 + q|s-r|^{2H}$. Using stationality of y_t, one has also $\mathbf{E}|y_s - y_r|^2 \le 2\mathbf{E}|y_1|^2 = 2$. Since for Gaussian random variables, the estimate for the L^{2p} norm is the same as for its L^2 norm, we have

$$\|y_s - y_r\|_{L^p} \lesssim \begin{cases} 1, & \text{if } |s-r| \ge 1; \\ |s-r|^H, & \forall |s-r| \le 1. \end{cases}$$

By symmetry, and change of variables,

$$\begin{aligned}
&\int_0^T \int_0^T \frac{\mathbf{E}|y_s - y_r|^p}{|s-r|^{\gamma p+2}} \\
&\lesssim \int_0^1 \int_{-v}^{v} v^{Hp-\gamma p-2}\, du\, dv + \left(\int_1^T \int_{-v}^{v} + \int_T^{2T} \int_{2T+v}^{2T-v} \right) v^{-\gamma p-2} du dv \\
&\lesssim 1 + T^{-\gamma p} \lesssim 1.
\end{aligned}$$

The first term is finite as soon as $\gamma < H$, the second term is finite as soon as γ is positive. The remaining claim follows from the well known Garcia-Rodemich-Romsey inequality below. □

Lemma 7.3 (Garcia-Rodemich-Romsey-Kolmogorov inequality) *Let $T > 0$.*

(1) Let $\theta : [0,T] \to \mathbf{R}^d$. For any positive numbers γ, p, there exists a constant $C(\gamma,p)$ such that

$$\sup_{s\neq t, s,t\in[0,T]} \frac{|\theta(t)-\theta(s)|}{|t-s|^\gamma} \le C(\gamma,p) \left(\int_0^T \int_0^T \frac{|\theta_s - \theta_r|^p}{|s-r|^{\gamma p+2}} ds dr \right)^{\frac{1}{p}}.$$

(2) Let θ be a stochastic process. Suppose that for $s,t \in [0,T]$, $p > 1$ and $\delta > 0$,

$$\mathbf{E}|\theta(t)-\theta(s)|^p \le c_p |t-s|^{1+\delta},$$

where c_p is a constant. Then for $\gamma < \frac{\delta}{p}$,

$$\| |\theta|_{C^\gamma([0,T])} \|_p \le C(\gamma,p)(c_p)^{\frac{1}{p}} \left(\int_0^T \int_0^T |u-v|^{\delta-\gamma p-1} du dv \right)^{\frac{1}{p}},$$

the right hand side is finite when $\gamma \in (0, \frac{\delta}{p})$.

7.3 Proof of the conditional integrability of fOU

The aim of this section is to prove the estimate (5.4), we therefore restrict ourselves to the case $H > \frac{1}{2}$. Firstly, we compute the conditional expectations of $\mathbf{E}(G(y_t)|\mathcal{F}_k)$ where $G \in L^2(\mu)$.

We begin with decomposing the fractional Ornstein-Uhlenbeck process into a part which is $\mathcal{F}_s$ measurable and another one independent of $\mathcal{F}_s$, where $\mathcal{F}_s$ is the filtration generated by B^H. In [Hai05b] it was shown that such a decomposition is available for fractional Brownian motion with $H > \frac{1}{2}$ by using the Mandelbrot-Van Ness representation it was shown that, for $k < t$,

$$\begin{aligned}
B_t - B_k &= \frac{1}{c_1(H)} \left(\int_{-\infty}^{k} (t-r)^{H-\frac{3}{2}} - (k-r)^{H-\frac{3}{2}} dW_r + \int_k^t (t-r)^{H-\frac{3}{2}} dW_r \right) \\
&= \overline{B}_t^k + \tilde{B}_t^k,
\end{aligned}$$

where $\overline{B}_t^k$ is $\mathcal{F}_k$ measurable and $\tilde{B}_t^k$ is independent of $\mathcal{F}_k$. Furthermore the filtration generated by the fractional Brownian motion is the same as the one generated by the two-sided Wiener process

W_t. Using the above we have,

$$
\begin{aligned}
y_t &= \int_{-\infty}^{t} e^{-(t-r)} dB_r = \int_{-\infty}^{k} e^{-(t-r)} dB_r + \int_k^t e^{-(t-r)} d(B_r - B_k) \\
&= \left(\int_{-\infty}^{k} e^{-(t-r)} dB_r + \int_k^t e^{-(t-r)} d\overline{B}_t^k \right) + \int_k^t e^{-(t-r)} d\tilde{B}_t^k \\
&= \overline{y}_t^k + \tilde{y}_t^k,
\end{aligned}
$$

where the first term $\overline{y}_t^k$ is $\mathcal{F}_k$ measurable and $\tilde{y}_t^k$ is independent of $\mathcal{F}_k$. In case $s \leq k$ we set $\overline{y}_t^k = y_s$.

It is only left so show that for $q \geq m$,

$$
\int_{k-1}^{\infty} \int_{k-1}^{\infty} \mathbf{E}\left(\overline{y}_s^k \overline{y}_r^k \right)^q dr ds
$$

is bounded in q and k. We make use of the following classical result,

Lemma 7.4 ([HC78]) *If X_s is a Gaussian process with covariance $R(t,s) = \mathbf{E}(X_s X_t)$. Suppose that $\partial_t \partial_s R(t,s)$ is integrable over every bounded region. Then we have, for any $f, g : \mathbf{R} \to \mathbf{R}$ satisfying the assumption*

$$
\int_a^b \int_a^b |f_t \, g_s \, \partial_t \partial_s R(t,s)| ds dt < \infty,
$$

the following Itô-isometry,

$$
\mathbf{E}\left(\int_a^b f_t \, dX_t \int_a^b g_s \, dX_s \right) = \int_a^b \int_a^b f_t \, g_s \, \partial_t \partial_s R(t,s) \, ds \, dt.
$$

Lemma 7.5 *Setting $R(t,s) = \mathbf{E}\left(\overline{B}_t^k \overline{B}_s^k \right)$ and $S(t,s) = \partial_t \partial_s R(t,s)$. Then S is regular in the region $\{t \neq s\}$, and*

$$
S(t,s) \lesssim (t \wedge s - k)^{2H-2} \qquad \forall s, t > k.
$$

Proof. Recall that $H > \frac{1}{2}$ and observe that for $k < t$,

$$
\begin{aligned}
\overline{B}_t^k &= \frac{1}{c_1(H)} \int_{-\infty}^{k} \left((t-r)^{H-\frac{1}{2}} - (k-r)^{H-\frac{1}{2}} \right) dW_r \\
&= \frac{H - \frac{1}{2}}{c_1(H)} \int_{-\infty}^{k} \int_k^t (u-r)^{H-\frac{3}{2}} \, du \, dW_r.
\end{aligned}
$$

By the Itô isometry for Wiener processes, for $k < \min(t,s)$,

$$
\begin{aligned}
&\mathbf{E}\left(\overline{B}_t^k \overline{B}_s^k \right) \\
&= \left(\frac{H - \frac{1}{2}}{c_1(H)} \right)^2 \int_k^t \int_k^s \mathbf{E}\left(\int_{-\infty}^{k} (u-r)^{H-\frac{3}{2}} dW_r \int_{-\infty}^{k} (w-v)^{H-\frac{3}{2}} dW_v \right) dw \, du \\
&= \left(\frac{H - \frac{1}{2}}{c_1(H)} \right)^2 \int_k^t \int_k^s \int_{-\infty}^{k} (u-y)^{H-\frac{3}{2}} (w-y)^{H-\frac{3}{2}} \, dy \, dw \, du,
\end{aligned}
$$

By a change of variables, for $t > s > k$

$$\begin{aligned}
S(t,s) &= \partial_t \partial_r R(t,r)\\
&= \left(\frac{H-\frac{1}{2}}{c_1(H)}\right)^2 \int_{-\infty}^{k} (t-y)^{H-\frac{3}{2}}(s-y)^{H-\frac{3}{2}}\, dy\\
&= \left(\frac{H-\frac{1}{2}}{c_1(H)}\right)^2 (t-s)^{2H-2} \int_{\frac{s-k}{t-s}}^{\infty} (1+w)^{H-\frac{3}{2}} w^{H-\frac{3}{2}}\, dw,
\end{aligned}$$

which is regular on $\{t > s\}$. Now

$$\int_{\frac{s-k}{t-s}}^{\infty} (1+w)^{H-\frac{3}{2}} w^{H-\frac{3}{2}}\, dw \le \int_{\frac{s-k}{t-s}}^{\infty} w^{2H-3}\, dw = \frac{1}{2-2H}\left(\frac{s-k}{t-s}\right)^{2H-2}$$

is finite for $H > \frac{1}{2}$ giving the bound

$$S(t,s) \lesssim (t \wedge s - k)^{2H-2},$$

concluding the proof for (t,s) off diagonal.

On the diagonal,

$$\begin{aligned}
\partial_t \partial_t \mathbf{E}\Big(c_1(H)\overline{B}_t^k\Big)^2 =& 2\int_{-\infty}^{k} (t-y)^{H-\frac{3}{2}}(t-y)^{H-\frac{3}{2}} dy\\
&+ 2\int_k^t \int_{-\infty}^{k} \left(H-\frac{3}{2}\right)(t-v)^{H-\frac{5}{2}}(s-v)^{H-\frac{3}{2}}\, dv\, ds\\
=& 2\int_{t-k}^{\infty} w^{2H-3} dw\\
&+ 2\frac{H-\frac{3}{2}}{H-\frac{1}{2}} \int_{-\infty}^{k} (t-v)^{H-\frac{5}{2}}\Big[(t-v)^{H-\frac{1}{2}} - (k-v)^{H-\frac{1}{2}}\Big]\, dv.
\end{aligned}$$

The second term has no singularity at k. Consequently,

$$\begin{aligned}
&\partial_t \partial_t \mathbf{E}\Big(c_1(H)\overline{B}_t^k\Big)^2\\
&= 2\frac{(t-k)^{2H-2}}{2H-2} + \frac{2H-3}{H-\frac{1}{2}} \int_{-\infty}^{k} (t-v)^{H-\frac{5}{2}}\Big[(t-v)^{H-\frac{1}{2}} - (k-v)^{H-\frac{1}{2}}\Big]\, dv\\
&\le 2\frac{(t-k)^{2H-2}}{2H-2} + \frac{2H-3}{H-\frac{1}{2}} \int_{-\infty}^{k} (t-v)^{2H-3}\, dv = C(t-k)^{2H-2}.
\end{aligned}$$

We have used the fact that $\int_{-\infty}^{k} (t-v)^{H-\frac{5}{2}}(k-v)^{H-\frac{1}{2}}\, dv$ is a finite number. This means the double derivative is only singular at $t = k$, but not along the whole diagonal. Moreover, the singularity is integrable. □

Remark 7.6 For the fractional Brownian motion B_t and for $t > s$,

$$\partial_t \partial_s \mathbf{E}(B_t B_s) = 2H(2H-1)(t-s)^{2H-2},$$

which is singular along the diagonal, whereas as shown in the proof of Lemma 7.5, $\partial_t\partial_t\mathbf{E}\Big(\overline{B}^k_t\Big)^2$ has only singularity at $t=k$, but not along the whole diagonal. Moreover, the singularity is integrable, we may therefore use the Itô isometry, Lemma 7.4, freely.

Now we go ahead and establish estimates for the decay of $\mathbf{E}\big(\overline{y}^k_t\overline{y}^k_s\big)$.

Lemma 7.7 *For any $k<s$,*

$$\int_k^s e^{-(s-v)}(v-k)^{2H-2}\,dv \lesssim 1\wedge(s-k)^{2H-2}.$$

The constant is independent of k.

Proof. This can be seen by splitting the integral into three regions $\int_k^{k+1}+\int_{k+1}^{\frac{s}{2}}+\int_{\frac{s}{2}}^s$, and integration by part with the second two terms. This leads to: $\int_k^{k+1}e^{-(s-v)}(v-k)^{2H-2}\,dv\le\frac{1}{2H-1}e^{-(s-k-1)}$. Furthermore for $s>3k$,

$$\begin{aligned}
&\int_{k+1}^s e^{-(s-v)}(v-k)^{2H-2}\,dv\\
&=(s-k)^{2H-2}-e^{-(s-k-1)}(2H-3)-(2H-2)\left(\int_{k+1}^{\frac{s}{2}}+\int_{\frac{s}{2}}^s\right)e^{-(s-v)}(v-k)^{2H-3}dv\\
&\lesssim(s-k)^{2H-2}-e^{-(s-k)}-e^{-\frac{s}{2}}(v-k)^{2H-2}|_{k+1}^{\frac{s}{2}}-(v-k)^{2H-2}|_{\frac{s}{2}}^{s}\\
&\lesssim(s-k)^{2H-2}+e^{-\frac{s}{2}}+(\frac{s}{2}-k)^{2H-2}\\
&\lesssim(s-k)^{2H-2}.
\end{aligned}$$

This gives the required estimate. □

Proposition 7.8 *Let $H>\frac{1}{2}$ and suppose that $H^*(m)<0$. Then,*

$$\sup_k\sup_{q\ge m}\int_{k-1}^\infty\int_{k-1}^\infty\mathbf{E}\Big(\overline{y}^k_s\overline{y}^k_t\Big)^q\,dt\,ds<\infty.$$

Proof. We first compute,

$$\begin{aligned}
\mathbf{E}\Big(\overline{y}^k_t\overline{y}^k_s\Big)=&\mathbf{E}\bigg(\Big(e^{-(t-k)}y_k+\int_k^t e^{-(t-r)}d\overline{B}^k_r\Big)\Big(e^{-(s-k)}y_k+\int_k^s e^{-(s-v)}d\overline{B}^k_v\Big)\bigg)\\
&=e^{-(t-s)}+I\!I+I\!I\!I+I\!V,
\end{aligned}$$

where the first term is due to $\mathbf{E}(y_k)^2=1$ and $e^{-(t-k)}e^{-(s-k)}\le e^{-(t-s)}$, and

$$\begin{aligned}
I\!I&=\mathbf{E}\Big(e^{-(t-k)}y_k\int_k^s e^{-(s-v)}d\overline{B}^k_v\Big),\\
I\!I\!I&=\mathbf{E}\Big(e^{-(s-k)}y_k\int_k^t e^{-(t-r)}d\overline{B}^k_r\Big),\\
I\!V&=\mathbf{E}\Big(\int_k^t e^{-(t-r)}d\overline{B}^k_r\int_k^s e^{-(s-v)}d\overline{B}^k_v\Big),
\end{aligned}$$

for $s \gg k$. For $I\!I$ we use Cauchy-Schwartz inequality and the Lemma 7.5,

$$\begin{aligned}(I\!I)^2 &= \mathbf{E}\Big(e^{-(t-k)} y_k \int_k^s e^{-(s-v)} d\overline{B}^k_v\Big)^2 = e^{-2(t-k)} \mathbf{E}\Big(\int_k^s e^{-(s-v)} d\overline{B}^k_v\Big)^2 \\ &\lesssim e^{-2(t-k)} \int_k^s \int_k^s e^{-(s-v)} e^{-(s-u)} \partial_u \partial_v\, \mathbf{E}\Big(\overline{B}^k_u \overline{B}^k_v\Big)\, du\, dv \\ &\lesssim e^{-2(t-k)} \int_k^s \int_k^s e^{-(s-v)} e^{-(s-u)} (u \wedge v - k)^{2H-2}\, du\, dv.\end{aligned}$$

We continue with the computation, making use of Lemma 7.7 in the final step:

$$\begin{aligned}&\int_k^s \int_k^s e^{-(s-v)} e^{-(s-u)} (u \wedge v - k)^{2H-2}\, du\, dv \\ &\lesssim 2 \int_k^s e^{-(s-v)} (v-k)^{2H-2} \Big(\int_v^s e^{-(s-u)}\, du\Big)\, dv \\ &\lesssim \int_k^s e^{-(s-v)} (s-v)^{2H-2}\, dv \\ &\lesssim (s-k)^{2H-2}.\end{aligned}$$

Putting them together we have,

$$I\!I \leq e^{-(t-k)} (s-k)^{H-1}.$$

Analogous arguments lead to $I\!I\!I \lesssim (t-k)^{2H-2}$. Assuming, without loss of generality that $t \geq s$,

$$\begin{aligned}IV &= \mathbf{E}\Big(\int_k^t e^{-(t-r)} d\overline{B}^k_r \int_k^s e^{-(s-v)} d\overline{B}^k_v\Big) \\ &= \int_k^t \int_k^s e^{-(t-r)} e^{-(s-r)} S(r,v) dv dr \\ &\lesssim e^{-(t-s)} \int_k^t \int_k^s e^{-(s-r)} e^{-(s-r)} (r \wedge v - k)^{2H-2} dv dr \\ &= 2e^{-(t-s)} \int_k^t \int_r^s e^{-(s-r)} e^{-(s-r)} (r \wedge v - k)^{2H-2} dv dr \\ &\lesssim 2e^{-(t-s)} \int_k^t (r-k)^{2H-2} e^{-(s-r)} \int_r^s e^{-(s-r)}\, du\, dr \\ &\lesssim 2 \int_k^t (r-k)^{2H-2} e^{-(t-r)}\, dr \\ &\lesssim 2(t-k)^{2H-2}.\end{aligned}$$

We have again applied Lemma 7.7 . Putting everything we and using

$$\mathbf{E}\Big(\overline{y}^k_t \overline{y}^k_s\Big) \leq \|\overline{y}^k_t\|_{L^2} \|\overline{y}^k_s\|_{L^2} \leq 1,$$

we obtain

$$\mathbf{E}\Big(\overline{y}_t^k\overline{y}_s^k\Big) \lesssim 1 \wedge (t\wedge s-k)^{2H-2}.$$

Now recall that $H^*(m)<0$ is equivalent to $(H-1)m+1<0$. Consequently, for $q\geq m$ and $(2H-2)m+2<0$,

$$\begin{aligned}\int_{k-1}^\infty\int_{k-1}^\infty \mathbf{E}\Big(\overline{y}_t^k\overline{y}_s^k\Big)^q &\leq C\int_{k-1}^\infty\int_{k-1}^\infty 1\wedge(t\wedge s-k)^{(2H-2)q}dsdt\\ &<\infty,\end{aligned}$$

where C is a constant independent of q. We have reached the conclusion of the proposition. □

7.4 Kernel convergence of scaling path integrals

We prove Lemma 4.2. In [Taq79], instead of y_t^ε, a moving average of the form

$$X_t=\int_{\mathbf{R}} p(t-\xi)dW_\xi.$$

for a suitable function p was considered and limit theorems (convergence in finite dimensional distributions) were proven for

$$\varepsilon^{H^*(m)}\int_0^{\frac{t}{\varepsilon}} G(X_s)ds.$$

In this setup, in order to prove weak convergence one uses the self-similarity of a Wiener process, $\sqrt{\lambda}W_{\frac{t}{\lambda}}\sim W_t$, leading to weak convergence as this equivalence of course is only in law. Nevertheless, in our case we can write directly, without using self-similarity properties, $y_t^\varepsilon=\int_{\mathbf{R}}\hat{g}(\frac{t-\xi}{\varepsilon})dW_\xi$, and thus avoid using self-similarity and, hence, obtain convergence in L^2. In [Taq79, Theorem 4.7] using [Taq79, Lemma 4.5, Lemma 4.6] the following result was obtained

$$\mathbf{E}\left(\int_{\mathbf{R}^m}\int_0^t\prod_{i=1}^m p\left(\frac{s-\xi}{\varepsilon}\right)\varepsilon^{H-\frac{3}{2}}dsdW_\xi-\frac{Z^{H^*(m),m}}{K(H,m)}\right)\to 0,$$

using the Wiener integral representation of the Hermite processes this is equivalent, by multiple Wiener-Ito isometry, to

$$\int_{\mathbf{R}^m}\left(\int_0^t\prod_{i=1}^m p\left(\frac{s-\xi_i}{\varepsilon}\right)\varepsilon^{H-\frac{3}{2}}ds-\int_0^t\prod_{i=1}^m(s-\xi_i)_+^{H-\frac{3}{2}}ds\right)^2 d\xi_1\dots d\xi_m.$$

To apply [Taq79, Theorem 4.7] we rewrite our kernels in the above moving average form. For the rescaled fractional Ornstein-Uhlenbeck process we obtain by above computation,

$$\begin{aligned} y_t^\varepsilon &= \frac{1}{c_1(H)}\varepsilon^{-H}\int_{\mathbf{R}}\int_{-\infty}^{t} e^{-\frac{t-u}{\varepsilon}}(u-s)_+^{H-\frac{3}{2}}du dW_s \\ &= \frac{1}{c_1(H)}\varepsilon^{-H}\int_{\mathbf{R}} e^{-\frac{t-s}{\varepsilon}}\int_{-\infty}^{t-s} e^{\frac{v}{\varepsilon}}v_+^{H-\frac{3}{2}}dv dW_s \\ &= \frac{1}{c_1(H)}\varepsilon^{-\frac{1}{2}}\int_{-\infty}^{t} e^{-\frac{t-s}{\varepsilon}}\int_0^{\frac{t-s}{\varepsilon}} e^{v}v_+^{H-\frac{3}{2}}dv dW_s \\ &= \varepsilon^{-\frac{1}{2}}\int_{-\infty}^{t} g\left(\frac{t-s}{\varepsilon}\right)dW_s, \end{aligned}$$

where

$$g(s) = \frac{1}{c_1(H)}e^{-s}\int_0^s e^u u_+^{H-\frac{3}{2}}du. \tag{7.4}$$

To apply [Taq79, Theorem 4.7] we will verify the conditions for this g.

Now, the term we consider in (4.2) has the following form,

$$\varepsilon^{H^*(m)-1}\int_{\mathbf{R}^m}\int_0^t \prod_{i=1}^m g\left(\frac{s-\xi_i}{\varepsilon}\right)\varepsilon^{-\frac{m}{2}}ds dW_{\xi_1}\dots dW_{\xi_m}.$$

Using $H^*(m) = (H-1)m+1$ we obtain that this is equal to

$$\int_{\mathbf{R}^m}\int_0^t \prod_{i=1}^m g\left(\frac{s-\xi_i}{\varepsilon}\right)\varepsilon^{H-\frac{3}{2}}ds dW_{\xi_1}\dots dW_{\xi_m},$$

the required Lemma 4.2 will follow from the reformulated convergence below.

Lemma 7.9 *Let λ denote the Lebesgue measure, then*

$$\left\|\int_0^t \prod_{i=1}^m g\left(\frac{s-\xi_i}{\varepsilon}\right)\varepsilon^{H-\frac{3}{2}}ds - \int_0^t \prod_{i=1}^m (s-\xi_i)_+^{H-\frac{3}{2}}ds\right\|_{L^2(\mathbf{R}^m,\lambda)} \to 0. \tag{7.5}$$

Proof. We are now in the above framework and it is only left the check the conditions imposed on the functions p in [Taq79, Theorem 4.7]. Examining Taqqu's proof, we note that in fact the L^2 convergence of (7.5) is obtained under the following conditions.

1 $\int_{\mathbf{R}} p(s)^2 ds < \infty$.
2 $|p(s)| \le Cs^{H-\frac{3}{2}}L(u)$ for almost all $s > 0$.
3 $p(s) \sim s^{H-\frac{3}{2}}L(s)$ as $s \to \infty$.
4 There exists a constant γ such that $0 < \gamma < (1-H)\wedge(H-(1-\frac{1}{2m}))$ such that $\int_{-\infty}^0 |p(s)g(xy+s)|ds = o(x^{2H-2}L^2(x))y^{2H-2-2\gamma}$ as $x \to \infty$ uniformly in $y \in (0,t]$.

where L denotes a slowly varying function (for every $\lambda > 0$ $\lim_{x\to\infty}\frac{L(\lambda x)}{L(x)}) = 1$). Now we go ahead and show that g defined by (7.4) satisfies these conditions, to increase readability we suppress the constant $\frac{1}{c_1(H)}$ in the computations. For $s<1$,

$$e^{-s}\int_0^s e^u u^{H-\frac{3}{2}}du \le \int_0^s u^{H-\frac{3}{2}}du \lesssim s^{H-\frac{1}{2}}.$$

We calculate for $s>1$ via integration by parts

$$\begin{aligned} e^{-s}\int_0^s e^u u^{H-\frac{3}{2}}du &\le e^{-s}\int_0^1 e^u u^{H-\frac{3}{2}}du + e^{-s}\int_1^s e^u u^{H-\frac{3}{2}}du \\ &\lesssim e^{-s} + s^{H-\frac{3}{2}} - 1 + e^{-s}\int_1^s e^u u^{H-\frac{5}{2}}du \\ &\lesssim s^{H-\frac{3}{2}}. \end{aligned}$$

This of course implies that g is L^2 integrable. Finally observe that

$$\int_{-\infty}^0 |g(s)g(xy+s)|ds = 0$$

as $g(s)=0$ for $s<0$. With these we apply [Taq79, Theorem 4.7] to conclude the L^2 convergence of the kernels. □

7.5 Joint convergence by asymptotic independence

If a sequence of random variable x_n converges to x and another sequence y_n converges to y, does (x_n, y_n) converge jointly and what is the correlation of the limit. When both x_n and y_n are Gaussian sequences, they converge jointly to (x,y) where x and y are taken to be independent, provided the correlation between x_n and y_n converges to zero. In [NR14] these are generalised to moment determinant random variables as limits, in which case it is sufficient to show $\mathrm{Cov}(x_n^2, y_n^2) \to 0$. By Torsten Carleman's theorem, a real valued random variable is moment determinant if its p-moment grows no faster than $(\frac{n}{C})^n(\log n)^n$ (e.g. if it has exponentially decaying density). An L^2 random variable in the second chaos has exponential tails, beyond the second chaos, the tails grow decays slower than exponentially. Extensions to high order chaos were obtained in [NNP16]. Neither of these results are sufficient for the need in this paper, we therefore present a generalisation, which can be easily deduced from the reasoning in [NNP16]. The result in [BT13] has moment determinant limit as restrictions, has restrictions to Wiener chaos 1 and 2, i.e. corresponding to BM, fBM and Rosenblatt processes in our case.

These results benefits from three insights. The first being the characterisation, of Üstünel-Zaksi [UZ89] for independence of two iterated integrals with vanishing of contractions of their integral kernels. The second is the characterisation of the vanishing of the contractions of their integral kernels by covariances of their squares the third is the characterisation of independence of random variables by the covariance of their squares [NP05, NNP16].

Denote by $I_p(f)$ the p^{th} iterated Itô-Wiener integral

$$I_p(f) = p!\int_{-\infty}^{\infty}\int_{-\infty}^{s_{p-1}}\cdots\int_{-\infty}^{s_2} f(s_1,\dots,s_p)dW_{s_1}dW_{s_2}\dots dW_{s_p}.$$

We take the L^2 function f to be symmetric functions of appropriate number of variables throughout. Given $f \in L^2(\mathbf{R}^p)$ and $g \in L^2(\mathbf{R}^q)$, where $p, q \geq 1$, their is

$$f \otimes_1 g = \int_{\mathbf{R}} f(x_1, \ldots, x_{p-1}, s) g(y_1, \ldots, y_{q-1}, s) ds.$$

Similarly

$$f \otimes_r g = \int_{\mathbf{R}^r} f(x_1, \ldots, x_{p-r}, s_1, \ldots, s_r) g(y_1, \ldots, y_{q-r}, s_1, \ldots, s_r) ds_1 \ldots ds_r.$$

If $f \otimes_1 g = 0$, so do all higher order contractions.

By [UZ89, Thm.6] two integrals $I_p(f)$ and $I_q(g)$ are independent, if and only if the 1-contraction between f, g vanishes almost surely. The necessity comes from the product formula,

$$I_p(f) I_q(g) = \sum_{m=1}^{p \wedge q} \frac{p! q!}{m!(p-m)!(q-m)!} I_{p+q-2m}(f \otimes_m g).$$

and the independence: $\mathbf{E}(I_p(f) I_q(g))^2 = p! q! \| f \otimes g|_{L^2}$. All terms in the binomial expansion for the product drop to zero except for the $m = 0$ term. The following asymptotic independent result is proven in [NNP16, Thm. 3.1].

Lemma 7.10 *[NNP16] Given $F_\varepsilon = I_p(f^\varepsilon)$ and $G_\varepsilon = I_q(g^\varepsilon)$, then*

$$\mathrm{Cov}(F^{\varepsilon 2}, G_\varepsilon{}^2) \to 0$$

is equivalent to

$$\| f^\varepsilon \otimes_r g^\varepsilon \| \to 0,$$

for $1 \leq r \leq p \wedge q$.

It is also observed in two integrals $I_p(f)$ and $I_q(g)$ are independent, their Malliavin derivative begin orthogonal. This explain why Malliavin calculus comes into prominent play, which has been developed to its perfection in [NNP16, Lemma 3.2]. The space of test functions is taken to be $C_q^\infty := C^\infty \cap BC^{q-1}(\mathbf{R}^m)$, which is sufficient to approximiate indicator functions of any measurable sets. We also set for $\varphi \in C_q^\infty$,

$$\|\varphi\|_q = \|\varphi\|_\infty + \sum_{|k|=1}^{q} \left\| \frac{\partial^k}{\partial^k x} \right\|_\infty,$$

where the sum runs over multi-indices $k = (k_1, \ldots, k_m)$. Let $L = -\delta D$.

Lemma 7.11 *[NNP16] Let $\theta \in C_q^\infty(\mathbf{R}^m)$, $G = I_q(g)$, $F = (F_1, \ldots, F_m)$, where $F_i = I_{p_i}(f_i)$ and $\mathbf{E} F_i^2 = 1$, where $p_i \geq q$. Then*

$$\mathbf{E} |\langle (I-L)^{-1} \theta(F) DF_j, DG \rangle_{\mathscr{H}}| \leq c \, \|\theta\|_{q-1} \mathrm{Cov}(F_j^2, G^2),$$

where c is a constant depending on $\|F\|_{L^2}$ and $\|G\|_{L^2}$.

Throughout this section $f_i : \mathbf{R}^{p_i} \to \mathbf{R}$, $g : \mathbf{R}^q \to \mathbf{R}$ are symmetric functions.

The final piece of the puzzle is the observation that the defect in being independent is quantitatively controlled by the covariance of the squares of the relative components. The following is from [NNP16], our only modification is to take G to be vector valued. Let $g_i : \mathbf{R}^{q_i} \to \mathbf{R}$ be symmetric functions.

Lemma 7.12 *Let $F = (I_{p_1}(f_1), \dots I_{p_m}(f_m))$ and $G = (I_{q_1}(g_1), \dots, I_{q_n}(g_n))$ such that $p_k \geq q_l$ for every pair of k, l. Then for every $\varphi \in \mathcal{C}_q^\infty(\mathbf{R}^m)$, $\psi \in \mathcal{C}_1^\infty(\mathbf{R}^n)$, the following holds for some constant c, depending on $\|F\|_{L^2}$, $\|G\|_{L^2}$:*

$$\mathbf{E}(\varphi(F)\psi(G)) - \mathbf{E}(\varphi(F))\mathbf{E}(\psi(G)) \leq c\|D\psi\|_\infty \|\varphi\|_q \sum_{i=1}^m \sum_{j=1}^n \mathrm{Cov}(F_i^2, G_j^2)$$

Proof. Define $L^{-1}(\sum_{k=0}^\infty I_k(h_m)) = \sum_{k=1}^\infty \frac{1}{k} I_k(h_m) \in \mathbb{D}^{2,2}$. The key equality is $-DL^{-1} = (I-L)^{-1}D$. As in [NNP16],

$$\varphi(F) - \mathbf{E}(\varphi(F)) = LL^{-1}\varphi(F) = \sum_{j=1}^m \delta((I-L)^{-1}\partial_j \varphi(F) DF_j)$$

Multiply both sides by $\psi(G)$ and use integration by parts we see

$$\begin{aligned}
&\mathbf{E}(\varphi(F)\psi(G)) - \mathbf{E}(\varphi(F))\mathbf{E}(\psi(G)) \\
&= \sum_{j=1}^m \sum_{i=1}^n \mathbf{E}\big(\langle (I-L)^{-1}\partial_j \varphi(F) DF_j, DG_i\rangle_{\mathscr{H}} \partial_i \psi(G)\big) \\
&\leq \|D\psi\|_\infty \sum_{j=1}^m \sum_{i=1}^n \big|\mathbf{E}\big(\langle (I-L)^{-1}\partial_j \varphi(F) DF_j, DG_i\rangle_{\mathscr{H}}\big)\big|.
\end{aligned}$$

To conclude, apply to each summand Lemma 7.11 with $\theta = \partial_j \varphi$ and $G = G_i$. □

Lemma 7.13 *Let $F_\varepsilon = (I_{p_1}(f_1^\varepsilon), \dots I_{p_m}(f_m^\varepsilon))$ and $G_\varepsilon = (I_{q_1}(g_1^\varepsilon), \dots, I_{q_n}(g_n^\varepsilon))$ with $q_1 \leq q_2, \dots q_m \leq p_1 \leq p_2, \dots p_m$. Then for every $i \leq m, j \leq n$,*

$$\|f_j^\varepsilon \otimes_r g_i^\varepsilon\| \to 0, \quad 1 \leq r \leq p_j \wedge q_i$$

implies that for any $\varphi \in \mathcal{C}^\infty(R^m)_{p_m}$ $\psi \in \mathcal{C}^\infty_{q_n}$,

$$\mathbf{E}(\psi(F_\varepsilon)\psi(G_\varepsilon)) - \mathbf{E}(\psi(F_\varepsilon))\mathbf{E}(\psi(G_\varepsilon)) \to 0.$$

Proof. Just combine Lemma 7.12 and Lemma 7.10. □

The following generalises results from [NNP16].

Proposition 7.14 *Given $F_\varepsilon = (I_{p_1}(f_1^\varepsilon), \dots, I_{p_m}(f_m^\varepsilon))$ and $G_\varepsilon = (I_{q_1}(g_1^\varepsilon), \dots, I_{q_n}(g_n^\varepsilon))$ such that $q_1 \leq q_2, \dots q_m, \leq p_1 \leq p_2, \dots \leq p_m$ and such that*

$$\|f_j^\varepsilon \otimes_r g_i^\varepsilon\| \to 0.$$

If $F_\varepsilon \to U$ and $G_\varepsilon \to V$ weakly, then $(F_\varepsilon, G_\varepsilon) \to (U, V)$ jointly where U, V are taken to be independent.

Proof. Since $(F_\varepsilon, G_\varepsilon)$ is bounded in L^2 it is tight. Now choose a weakly converging subsequence (F_n, G_n) with limit denoted by (X, Y). Let $\varphi \in \mathcal{C}^\infty_{p_m}(R^m)$ $\psi \in \mathcal{C}^\infty_{q_n}(\mathbf{R}^n)$. Then By Lemma 7.13 and the bounds on φ, ψ, we pass to the limit under the expectation sign and obtain

$$\mathbf{E}(\varphi(X)\psi(Y)) = \mathbf{E}(\varphi(X))\mathbf{E}(\psi(Y)).$$

Thus every limit measure is the product measure determined by U, V and hence the $(F_\varepsilon, G_\varepsilon)$ converges as claimed. □

Acknowledgement: We would like to thank Martin Hairer for helpful discussions.

References

[AS84] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York; John Wiley & Sons, Inc., New York, 1984. Reprint of the 1972 edition, Selected Government Publications.

[ATH12] H. Al-Talibi and A. Hilbert. Differentiable approximation by solutions of Newton equations driven by fractional Brownian motion. Preprint, 2012.

[BC09] Boris Buchmann and Ngai Hang Chan. Integrated functionals of normal and fractional processes. *Ann. Appl. Probab.*, 19(1):49–70, 2009.

[BC17] I. Bailleul and R. Catellier. Rough flows and homogenization in stochastic turbulence. *J. Differential Equations*, 263(8):4894–4928, 2017.

[BH02] Samir Ben Hariz. Limit theorems for the non-linear functional of stationary Gaussian processes. *J. Multivariate Anal.*, 80(2):191–216, 2002.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BM83] Peter Breuer and Péter Major. Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.*, 13(3):425–441, 1983.

[BT05] Brahim Boufoussi and Ciprian A. Tudor. Kramers-Smoluchowski approximation for stochastic evolution equations with FBM. *Rev. Roumaine Math. Pures Appl.*, 50(2):125–136, 2005.

[BT13] Shuyang Bai and Murad S. Taqqu. Multivariate limit theorems in the context of long-range dependence. *J. Time Series Anal.*, 34(6):717–743, 2013.

[CFK+19] Ilya Chevyrev, Peter K. Friz, Alexey Korepanov, Ian Melbourne, and Huilin Zhang. Multiscale systems, homogenization, and rough paths. In *Probability and Analysis in Interacting Physical Systems*, 2019.

[CFS82] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskii.

[CKM03] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 8:no. 3, 14, 2003.

[CNN18] Simon Campese, Ivan Nourdin, and David Nualart. Continuous breuer-major theorem: tightness and non-stationarity. In Arxiv: arXiv:1807.09740, 2018.

[DM79] R. L. Dobrushin and P. Major. Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete*, 50(1):27–52, 1979.

[Dob79] R. L. Dobrushin. Gaussian and their subordinated self-similar random generalized fields. *Ann. Probab.*, 7(1):1–28, 1979.

[dSEE18] José Luís da Silva, Mohamed Erraoui, and El Hassan Essaky. Mixed stochastic differential equations: existence and uniqueness result. *J. Theoret. Probab.*, 31(2):1119–1141, 2018.

[EM02] Paul Embrechts and Makoto Maejima. *Selfsimilar processes*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2002.

[FGL15] Peter Friz, Paul Gassiat, and Terry Lyons. Physical Brownian motion in a magnetic field as a rough path. *Trans. Amer. Math. Soc.*, 367(11):7939–7955, 2015.

[FH14] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[FK00] Albert Fannjiang and Tomasz Komorowski. Fractional Brownian motions in a limit of turbulent transport. *Ann. Appl. Probab.*, 10(4):1100–1120, 2000.

[FV10] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.

[GN08] João Guerra and David Nualart. Stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *Stoch. Anal. Appl.*, 26(5):1053–1075, 2008.

[Gre51] Melville S. Green. Brownian motion in a gas of noninteracting molecules. *J. Chem. Phys.*, 19:1036–1046, 1951.

[Hai05a] Martin Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.*, 33(2):703–758, 2005.

[Hai05b] Martin Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.*, 33(2):703–758, 2005.

[HBS] Harold Edwin Hurst, R. P. Black, and Y.M. Sinaika. *Long Term Storage in Reservoirs, An Experimental Study,*. Constable.

[HC78] Steel T. Huang and Stamatis Cambanis. Stochastic and multiple Wiener integrals for Gaussian processes. *Ann. Probab.*, 6(4):585–614, 1978.

[HNX14] Yaozhong Hu, David Nualart, and Fangjun Xu. Central limit theorem for an additive functional of the fractional Brownian motion. *Ann. Probab.*, 42(1):168–203, 2014.

[Itô51] K. Itô. Multiple Wiener integral. *J. of the Mathematical Society of Japan*, 3(1), 1951.

[JL77] G. Jona-Lasinio. Probabilistic approach to critical behavior. In *New developments in quantum field theory and statistical mechanics (Proc. Cargèse Summer Inst., Cargèse, 1976)*, pages 419–446. NATO Adv. Study Inst. Ser., Ser. B: Physics, 26, 1977.

[JS03] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[KKR12] Tomasz Komorowski, Alexei Komorowski, and Lenya Ryzhik. Evolution of particle separation in slowly decorrelating velocity fields. *Commun. Math. Sci.*, 10(3):767–786, 2012.

[KLO12] Tomasz Komorowski, Claudio Landim, and Stefano Olla. *Fluctuations in Markov processes*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012. Time symmetry and martingale approximation.

[KM17] David Kelly and Ian Melbourne. Deterministic homogenization for fast-slow systems with chaotic noise. *J. Funct. Anal.*, 272(10):4063–4102, 2017.

[KNR12] Tomasz Komorowski, Alexei Novikov, and Lenya Ryzhik. Evolution of particle separation in slowly decorrelating velocity fields. *Commun. Math. Sci.*, 10(3):767–786, 2012.

[KP91] Thomas G. Kurtz and Philip Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.

[Kub57] Ryogo Kubo. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Japan*, 12:570–586, 1957.

[KV86] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.

[LH19] Xue-Mei Li and Martin Hairer. Averaging dynamics driven by fractional brownian motion. arXiv:1902.11251, 2019.

[Lyo94] Terry Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4):451–464, 1994.

[MT07] Makoto Maejima and Ciprian A. Tudor. Wiener integrals with respect to the Hermite process and a non-central limit theorem. *Stoch. Anal. Appl.*, 25(5):1043–1056, 2007.

[MVN68] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[NNP16] Ivan Nourdin, David Nualart, and Giovanni Peccati. Strong asymptotic independence on Wiener chaos. *Proc. Amer. Math. Soc.*, 144(2):875–886, 2016.

[NNZ16] Ivan Nourdin, David Nualart, and Rola Zintout. Multivariate central limit theorems for averages of fractional Volterra processes and applications to parameter estimation. *Stat. Inference Stoch. Process.*, 19(2):219–234, 2016.

[NP05] David Nualart and Giovanni Peccati. Central limit theorems for sequences of multiple stochastic integrals. *The Annals of Probability*, 33(1):177–193, 2005.

[NP12] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012. From Stein's method to universality.

[NR14] Ivan Nourdin and J. Rosinski. Asymptotic independence of multiple wiener-itô integrals and the resulting limit laws. *The Annals of Probability*, 42(2):497–526, 2014.

[Nua06] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[PT00] Vladas Pipiras and Murad S. Taqqu. Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields*, 118(2):251–291, 2000.

[PT17] Vladas Pipiras and Murad S. Taqqu. *Long-range dependence and self-similarity*. Cambridge Series in Statistical and Probabilistic Mathematics, [45]. Cambridge University Press, Cambridge, 2017.

[Ros61] M. Rosenblatt. Independence and dependence. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II*, pages 431–443. Univ. California Press, Berkeley, Calif., 1961.

[Sam06] Gennady Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1(3):163–257, 2006.

[Sin76] Ja. G. Sinaı. Self-similar probability distributions. *Teor. Verojatnost. i Primenen.*, 21(1):63–80, 1976.

[Taq77] Murad S. Taqqu. Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(3):203–238, 1977.

[Taq79] Murad S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete*, 50(1):53–83, 1979.

[Tay21] G. I. Taylor. Diffusion by Continuous Movements. *Proc. London Math. Soc. (2)*, 20(3):196–212, 1921.

[UZ89] Ali Süleyman Üstünel and Moshe Zakai. On independence and conditioning on Wiener space. *Ann. Probab.*, 17(4):1441–1453, 1989.

[You36] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1):251–282, 1936.

[Zha08] Songfu Zhang. Smoluchowski-Kramers approximations for stochastic equations with Lévy noise. PhD theisis, Purdue Unviersity, 2008.