arXiv:2104.06364v1 [math.PR] 13 Apr 2021

Functional Limit Theorems for Volterra Processes and Applications to Homogenization

Johann Gehringer, Xue-Mei Li, and Julian Sieber
Imperial College London

April 14, 2021

Abstract

We prove an enhanced limit theorem for additive functionals of a multidimensional Volterra process $(y_t)_{t\geqslant 0}$. As an application, we establish weak convergence of the solutions of rough differential equations (RDE) of the form

$$dx_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} f(x_t^\varepsilon, y_{\frac{t}{\varepsilon}})\, dt + g(x_t^\varepsilon)\, d\mathbf{B}_t,$$

and identify their limits as solutions of an RDE driven by a Gaussian field with a drift coming from the Lévy area correction of the limiting rough driver. The equation models a passive tracer in a random field.

In particular if h is random field such that $h(x,\cdot)$ a semi-martingale with spatial parameter x, we show that the solutions of the equations

$$dx_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} f(x_t^\varepsilon, y_{\frac{t}{\varepsilon}})\, dt + h(x_t^\varepsilon, dt),$$

converge weakly to that of a Kunita type Itô SDE $dx_t = G(x_t, dt)$ where $G(x,t)$ is a semi-martingale with spatial parameters. Furthermore the N-point motions converge.

Keywords: fractional noise, multi-scale, correlation, mixed functional central and non-central limit theorems, rough creation, rough homogenization, semi-martingales with spatial parameters, simultaneous convergence

MSC2020 Subject classification: 34F05, 60F05, 60F17

Contents

1 Introduction **2**

2 Preliminaries **4**
2.1 The Normalized Noise Process 4
2.2 Malliavin Calculus 7
2.3 Hölder Spaces and Their Tensor Product 8
2.4 Rough Path Theory 9

3 The Multidimensional Limit Theorem **10**
3.1 Statement of the results 10
3.2 Functional Central Limit Theorem 11
3.3 Lifted Functional Central Limit Theorem 15
3.4 Proof of Theorem 3.1 19

4 Rough Homogenization **20**
4.1 Casting the Homogenization Problem as RDE 20
4.2 Statement of the result 21
4.3 Weak Convergence in Rough Path Spaces 24
4.4 Proof of Theorem 4.6 26

1 Introduction

We first study the convergence of the stochastic processes $(X_t^\varepsilon, \mathbb{X}^\varepsilon)$ where X^ε is of the form

$$X_t^\varepsilon = \left(\sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_1(y_s)\, ds, \dots, \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_N(y_s)\, ds \right)$$

and the $\mathbb{X}^\varepsilon \overset{\text{def}}{=} \int_s^t (X_r^\varepsilon - X_s^\varepsilon) \otimes dX_r^\varepsilon$'s are the canonical lift and y is a correlated multi-dimensional Voterra processes, a correlated Gaussian noise with a spectral density, with Hölder regularity greater than $\frac{1}{2}$ and slow correlation decay. We obtain a Functional limit theorem together with an enhanced and a rough CLT. The set of admissible functions G_i for the CLT are L^2 functions with Hermite rank bounded from below by $\frac{1}{\beta}$ (or by $\frac{2}{\beta}$ for the rough CLT) that are sufficiently smooth.

As an application we find an approximate effective dynamics for the evolution of a particle moving in a fast oscillating random field $f(\cdot, y_t(\omega))$. More precisely we study the small ε limit of the solutions of the following equations on $\mathbf{R}^d$:

$$dx_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} f(x_t^\varepsilon, y_{\frac{t}{\varepsilon}})\, dt + g(x_t^\varepsilon)\, d\mathbf{B}_t, \tag{1.1}$$

where $\mathbf{B}$ is a rough path with Hölder regular sample paths in the Hölder space $\mathcal{C}^{\frac{1}{3}+}$. The noise $\mathbf{B}$ does not play any essential role and can be turned off and can be also replaced a random field $g(x, dt)$.

We emphasize that even when $\mathbf{B}_t = (B_t, \mathbb{B})$ is the Itô lift of a Brownian motion and so (1.1) can be interpreted as an Itô stochastic differential equation (and even when y is a strong mixing Markov process so we are in classical domain), there is advantage to consider the Itô stochastic differential equation as a RDE. With the approach presented here, we automatically have simultaneous convergence of the solutions with any finite number of initial conditions to the solutions of the same limiting SDE of Kunita type, which is not easy to prove with the usual martingale method.

To illustrate the mechanism behind the homogenization problem we first review the following popular perturbation model:

$$\dot{z}_t^\varepsilon = \varepsilon g(z_t^\varepsilon, y_t) + f(x_t^\varepsilon, y_t)$$

where $f, g : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}^d$ are sufficiently regular with suitable initial conditions. On the scale of $[0, \frac{1}{\sqrt{\varepsilon}}]$, z_t^ε can be identified with the solution of the slow/fast system:

$$\dot{x}_t^\varepsilon = g(x_t^\varepsilon, y_{\frac{t}{\varepsilon}}) + \frac{1}{\sqrt{\varepsilon}} f(x_t^\varepsilon, y_{\frac{t}{\varepsilon}}), \qquad x_0^\varepsilon = x_0, \qquad x_0 \in \mathbf{R}^d.$$

The parameter ε is a positive number tuning the relative speed with respect to the fast motion y and is assumed to be small. If the stochastic process (y_t) exhibits oscillatory features and f averages to zero with respect to a measure determined by it, the slow motion feels the Averaged and the Central Limit Theorem Velocity. It is then plausible to deduce a simpler equation, called the effective equation, whose solutions approximate (z_t^ε) on the scale $[0, \frac{1}{\sqrt{\varepsilon}}]$. This is the principle of slow/fast homogenisation. Unlike for (x_t^ε), the effective equation for $\lim_{\varepsilon \to 0} x_t^\varepsilon$ is autonomous. There are two well known theories on this: (1) the fast motion is deterministic, (2) the fast motion is strong mixing Markovian. If the fast motion is periodic this is classic. More recently, diffusive homogenisation is obtained for a class of non-uniformly hyperbolic fast flows including solutions of the classical the Lorenz equations [KM17]. The second theory, in which y_t is an ergodic Markov process y_t satisfying a strong mixing condition, is by now also classic [KP78, FW12]. The theory we intend to study is a fast dynamics with slow decaying correlation. Such dynamics appears naturally in natural and economical social sciences, as well as in statistics, [CR96, Rio17, Sam06, Rob95].

In this article, y_t is taken to be an admissible 'moving average' Gaussian process. Let W_t be a Brownian motion in $\mathbf{R}^d$, and $K(r) \in \mathbb{L}(\mathbf{R}^d, \mathbf{R}^n)$ an admissible kernel. A Volterra process is of the form

$$y_t = \int_{-\infty}^t K(t-u) dW_u. \tag{1.2}$$

Increments of Brownian motions and fractional Brownian motion, and the fractional Ornstein-Uhlenbeck processes are Volterra processes. In fact, a centred one dimensional stationary L^2 continuous Gaussian process (y_t) has the integral representation (1.2) if and only if its spectral measure has a density [Kal21, Prop. 14.20] by simply taking K the Fourier transform of the square root of the density. We will assume an integrable condition on the kernel, ddescribing the correlation decay: for some $\beta > 0$, $s, t \geqslant 0$,

$$\big|\mathbf{E}[y_s \otimes y_t]\big| = \left| \int_{\mathbf{R}} K(s-u)K^*(t-u)\, du \right| \leqslant \Theta\big(1 \wedge |t-s|^{-\beta}\big).$$

For a stationary Gaussian processes, the regularity of the spectral measure can be obtained from the decay rate of the correlation function. In comparison, 'strong mixing', which can be characterised by the spectral density, cannot be classified by the decay rate of its correlation function.

Before going ahead explaining how to propose the convergence theorem, we go back to the well known situation in which y_t is a Markov process or strong mixing and take $\sigma = 0$ for simplicity. Then the ansatz is a diffusion process whose generator can be formally written down and the road map for the convergence of x_t^ε is to show that $\{(x_t^\varepsilon)\}$ is relatively compact followed by an application of the martingale problem method. The paramount questions is then to determine for which L^2 function G a functional central limit theorem holds for $\sqrt{\varepsilon}\int_0^{t/\varepsilon} G(y_s)ds$. This is the Kipnis-Varadhan theory [KV86] allowing to include a large family of functions and therefore beautiful applications in both continuous and discrete stochastic models, see [KLO12] and the reference therein. We would like to highlight the rough path approach which turns out to be fruitful and explored in [KM17, GL19, DOP19].

In case y_s is the non-strong mixing Volterra Gaussian noise, how do we formulate the condition on the driver? Let us first take y_t to be the one dimensional stationary fractional Ornstein-Uhlenbeck process and consider

$$\alpha(G,\varepsilon)\int_0^{t/\varepsilon} G(y_s)ds$$

Then the scale $\alpha(G,\varepsilon)$ and the scaling limit depend on its Hermit rank, this is the lowest m in the Hermite polynomial expansion of $G = \sum_{i=m}^\infty c_i H_i$. Furthermore, the solutions of the following equations $\dot{x}_t^\varepsilon = \sum_{i=1}^N \alpha(G_i,\varepsilon) f_i(x_t^\varepsilon) G_i(y_{t/\varepsilon})$ converge [GL20a]. For $N = 1$, this is straightforward. The $N > 1$, this models different driving vector fields across different regions, is already more complex and involves CLT's for the iterated integrals in a topology stronger than the weak topology. An enhanced CLT is deduced from and built on the vast literature before, c.f. [Taq79, MT07, CNN18] under a fast chaos decay condition. The effect equation can be described in some cases a diffusion, in other cases an anomalous super diffusion, or even a stochastic differential equation with mixed Itô, Lebesgue, and Young integrals. For related work see [ILRMS13]

The result we described applies to the product from $f(x,y) = \sum_{i=1}^N f_i(x)G(y)$ and the one dimensional toy model noise. The problem remains open for the evolution in the random field generated by multi-dimensional noise and for the non-product form. For a preliminary examination, let denote the random field by $F(x,t,\varepsilon,\omega) = f(x, y_{\frac{t}{\varepsilon}(\omega)})$, assumed to be in L^2. Taylor expanding F in x in some direction v, we see

$$F(x,t,\varepsilon,\omega) \sim f(x_0, y_{\frac{t}{\varepsilon}}(\omega)) + \nabla_x f(x_0, y_{\frac{t}{\varepsilon}}(\omega)) \cdot v + \dots.$$

The CLT applies only if f is sufficiently fast oscillatory so the oscillation compensates the insufficient decay in the auto-correlation of (y_t). It is more efficient to take the Hermite expansion of f with respect to the Gaussian process y_t. Wrapping the Hermite polynomials around a Gaussian process amplifies its correlation decay. We use this to determine the decay rate of the random field. In case of $\mathbf{E}[y_0 \otimes y_0] = \mathrm{id}$, the Hermite rank of $f(x,\cdot)$ is the lowest $|\boldsymbol{\ell}|$ with non-vanishing $c_{\boldsymbol{\ell}}$ in the following Hermite polynomial expansion.

$$f(x,y_t) = \sum_{\boldsymbol{\ell}} c_{\boldsymbol{\ell}}(x) H_{\boldsymbol{\ell}}(y_t) \qquad \forall\, t \geqslant 0. \tag{1.3}$$

where $\boldsymbol{\ell} = (\ell_1, \dots, \ell_n)$ is a multi-index and $c_{\boldsymbol{\ell}} = (c_{\boldsymbol{\ell}}^1, \dots, c_{\boldsymbol{\ell}}^n)$. Otherwise we use a transformed process z_t for defining the Hermite rank. The Hermite rank precisely characterises whether the functional CLT holds. We will specify the regularity conditions later.

Main Results. Our main results are presented in **Theorem 3.1** and **Theorem 4.6**. The former is an enhanced and rough Functional limit theorem for the multi-dimensional Voterra processes: $X^\varepsilon \to X$ and under more stringent decay conditions on the correlation, we show $(X^\varepsilon, \mathbf{X}^\varepsilon) \to (X, \mathbf{X})$ where X is a Gaussian process. Theorem 4.6 is an application to the homogenisation problem for (1.1). We show that for compactly coefficients the effective limit is an RDE of the form

$$dx_t = \bar{f}(x_t)\, d\mathbf{X}_t + \Lambda(x_t)dt + g(x_t)\, d\mathbf{B}_t$$

where $\mathbf{X}$ is the Itô lift of a Gaussian random field which we denote by $X_t(x)$. In case B is a BM (allowing to have a drift) and $\mathbf{B} = (B, \mathbb{B})$ its Itô rough path lift, for $f \in \mathcal{C}_b^3$ and $g \in \mathcal{C}_b^{2+}$, we show the limiting effective equation is an SDE of Kunita type:

$$dx_t = \Gamma(x_t)dt + F(x_t, dt) + g(x_t)dB_t,$$

where $F(x,\cdot) \overset{\text{def}}{=} X(x)$ is a martingale with spatial parameters whose characteristics are of the form $a_{i,j}(x,z,t) = \sigma_{i,j}(x,z)t$. The advantage of the formulation with Kunita SDE is that for any finite set of initial positions we have the weak convergence of the N-point motion, by which we mean $(\varphi^\varepsilon(z^1), \dots, \varphi^\varepsilon(z^N))$, where φ^ε are the solutions flows of the SDEs and $z^i \in \mathbf{R}^d$ are initial points. This is due to the convergence of the drivers in the compactly supported case and solution theory on SDEs of Kunita type. These theorems are proved under further conditions on the Hermite ranks of G_i and f.

The proof of Theorem 3.1 is the content of Section 3 which relies on computational techniques from Malliavin Calculus, and that for Theorem 4.6 is presented in Section 4. For the proof of convergence of (x_t^ε), we take the continuity theorem route for solution flows of a rough differential equations (RDE). For this we cast the equation (1.1) as a Banach space-valued rough differential equations. To put this in the perspective of diffusion processes, this is analogous to lifting a smooth stochastic differential equation to the diffeomorphism group. The rough differential equation method has previously been employed in the very nice work of Kelly and Melbourne [KM17], in a different context, there is however minor inaccuracies in that article (see Example B), the arguments are rather formal leaving verifications wanted. By fixing these, we also resolved a question raised in [KM17] in Proposition 4.9, with which we also bypass the invocation of the martingale problem method used there, for identifying the limit equation. We then turn this problem back to a finite state problem, and interpret the resulting RDE as classical stochastic differential equations of Kunita type and identify the characteristics of the driving semi-martingales with spatial parameters. Note that we have kept the infinite dimensional noise in the Kunita type SDE, for we would loose the simultaneous convergence if it is converted to an SDE driven by a finite number of noise.

2 Preliminaries

This section features a brief overview of background material, and pre-treatment of the noise.

2.1 The Normalized Noise Process

Definition 2.1 An $\mathbf{R}^n$-valued stochastic process $(y_t)_{t\geqslant 0}$ is called a *(stationary) Volterra process* if there is an integer $d \in \mathbf{N}$, a square-integrable kernel $K : \mathbf{R} \to \mathbb{L}(\mathbf{R}^d, \mathbf{R}^n)$ with $K(t) = 0$ for $t < 0$, and a d-dimensional Wiener process $(W_t)_{t\in\mathbf{R}}$ such that

$$(y_t)_{t\geqslant 0} \overset{d}{=} \left(\int_{\mathbf{R}} K(t-s)\, dW_s \right)_{t\geqslant 0}. \tag{2.1}$$

Let $\beta, \Theta > 0$. We write $\mathcal{V}_n(\beta, \Theta)$ for the space of n-dimensional Volterra processes which satisfy

$$\left| \int_{-\infty}^{-t} \langle K_j(-u), K_j(-u) \rangle\, du \right| \leqslant \Theta\big(1 \wedge t^{-2\beta}\big) \qquad \forall\, t \geqslant 0,\ j = 1, \dots, n, \tag{2.2}$$

where K_j is the j^{th} row of K. It is also convenient to declare $\mathcal{V}_n \overset{\text{def}}{=} \bigcup_{\beta,\Theta>0} \mathcal{V}_n(\beta,\Theta)$.

It is clear that any $y \in \mathcal{V}_n$ is a centered, stationary Gaussian process, which is actually ergodic under the canonical time-shift [CFS82]. We set $\Sigma \stackrel{\text{def}}{=} \mathbf{E}[y_0 \otimes y_0]$ throughout the article.

By a simple application of the Cauchy-Schwarz inequality, the estimate (2.2) yields

$$\left| \int_{\mathbf{R}} \langle K_i(s-u), K_j(t-u) \rangle \, du \right| \leqslant \Theta\big(1 \wedge |t-s|^{-\beta}\big) \qquad \forall\, s,t \geqslant 0,\ i,j = 1,\dots,n.$$

This in turn implies the following decay of the temporal correlations of $(y_t)_{t \geqslant 0}$

$$\big|\mathbf{E}[y_s \otimes y_t]\big| \leqslant \Theta\big(1 \wedge |t-s|^{-\beta}\big).$$

Let us give a few examples comprised by Definition 2.1:

Example A

1. Given a fractional Brownian motion B with Hurst parameter $H \in (0,1) \setminus \{\frac{1}{2}\}$, one can show that for functions f satisfying

$$\begin{cases} \int_{\mathbf{R}} \Big(\int_{\mathbf{R}} f(u)(u-s)_+^{H-\frac{3}{2}} du \Big)^2 ds & H > \frac{1}{2}, \\ \exists \varphi_f \in L^2(\mathbf{R}) : f(s) = \int_{\mathbf{R}} \varphi_f(u)(u-s)_+^{-H-\frac{1}{2}} & H < \frac{1}{2}, \end{cases}$$

   the following identity holds

$$\int_{\mathbf{R}} f(s) dB_s^H = c_H \begin{cases} \displaystyle\int_{\mathbf{R}} \int_{\mathbf{R}} f(u)(u-s)_+^{H-\frac{3}{2}} \, du \, dW_s, & H > \frac{1}{2}, \\ \displaystyle\int_{\mathbf{R}} \int_0^\infty \frac{f(s) - f(s+u)}{u^{\frac{3}{2}-H}} \, du \, dW_s, & H < \frac{1}{2}, \end{cases}$$

   see [PT00]. Here, $c_H > 0$ is some explicitly known constant. Hence, the fractional Ornstein-Uhlenbeck process $dX_t = -X_t + dB_t$ [CKM03] is a Volterra process with kernel

$$K(t) = c_H \begin{cases} \displaystyle\int_0^\infty e^{-u}(u-t)_+^{H-\frac{3}{2}} \, du, & H > \frac{1}{2}, \\ \displaystyle e^{-t} \int_0^\infty \frac{\mathbf{1}_{[0,\infty]}(t) - e^{-u}\mathbf{1}_{[-\infty,t]}(u)}{u^{\frac{3}{2}-H}} \, du, & H < \frac{1}{2}. \end{cases}$$

   It is not hard to check that (2.2) holds with $\beta = 1 - H$.
2. Another example is

$$K(t) = c_H \begin{cases} 0 & \text{if } t < 0, \\ t^{H-\frac{1}{2}} & \text{if } 0 \leqslant t \leqslant 1, \\ t^{H-\frac{1}{2}} - (t-1)^{H-\frac{1}{2}} & \text{if } t \geqslant 1, \end{cases} \qquad H \in (0,1),$$

   which leads (for a suitable choice of the normalization constant $c_H \in \mathbf{R}$) to a fractional Brownian increment $\int_{\mathbf{R}} K(t-s)\, dW_s = B_t - B_{t-1}$ [MVN68]. It is again easy to see that (2.2) holds with $\beta = 1 - H$.

Fact 2.2 *Let $\Sigma \in \mathbb{L}(\mathbf{R}^n, \mathbf{R}^n)$ be a positive semi-definite matrix with rank m. Let $O \in \mathbb{L}(\mathbf{R}^m, \mathbf{R}^n)$ be a linear transformation such that $D \stackrel{\text{def}}{=} O^\top \Sigma O$ is diagonal and features precisely the non-zero eigenvalues of Σ. Let G be a real-valued function on $\mathbf{R}^n$. Then*

$$G \in L^2\big(\mathbf{R}^n, N(0,\Sigma)\big) \quad \Longleftrightarrow \quad G\big(OD^{\frac{1}{2}} \cdot\big) \in L^2\big(\mathbf{R}^m, N(0, \mathrm{id})\big).$$

Let $(y_t)_{t\geqslant 0}$ be a centered, stationary Gaussian process and recall that $\Sigma = \mathbf{E}[y_0 \otimes y_0]$. Let D and O denote the matrices furnished by Fact 2.2. For any $G \in L^2\big(\mathbf{R}^n, N(0,\Sigma)\big)$, we have the following $L^2\big(\mathbf{R}^n, N(0,\mathrm{id})\big)$-convergent expansion:

$$G\big(OD^{\frac{1}{2}}y\big) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} c_{\boldsymbol{\ell}} H_{\boldsymbol{\ell}}(y), \qquad \text{where} \quad c_{\boldsymbol{\ell}} \stackrel{\text{def}}{=} \int_{\mathbf{R}^n} G\big(OD^{\frac{1}{2}}y\big) H_{\boldsymbol{\ell}}(y)\, N(0,\mathrm{id})(dy).$$

Here, $H_{\boldsymbol{\ell}} : \mathbf{R}^n \to \mathbf{R}$ denotes the *Hermite polynomial* of degree $\boldsymbol{\ell}$:

$$H_{\boldsymbol{\ell}}(x) \stackrel{\text{def}}{=} \prod_{i=1}^n H_{\boldsymbol{\ell}_i}(x_i), \qquad x = (x_1,\dots,x_n) \in \mathbf{R}^n,$$

where

$$H_m(x) \stackrel{\text{def}}{=} (-1)^m e^{\frac{x^2}{2}} \frac{d^m}{dx^m} e^{-\frac{x^2}{2}}, \qquad x \in \mathbf{R},\ m \in \mathbf{N}_0.$$

Note that $H_0(x) = 1$, $H_1(x) = x$ and $\langle H_m, H_n\rangle_{L^2(\mathbf{R},N(0,1))} = \delta_{m,n} m!$. We want to remark that in [Nua06] Hermite polynomials are defined with a different normalization.

Let us introduce the *normalized process* $(z_t)_{t\geqslant 0}$ by declaring

$$z_t \stackrel{\text{def}}{=} D^{-\frac{1}{2}} O^\top y_t. \tag{2.3}$$

Then $(z_t)_{t\geqslant 0}$ is clearly a centered, stationary Gaussian process with $\mathbf{E}[z_0 \otimes z_0] = \mathrm{id}$ and

$$G(y_t) = G\big(OD^{\frac{1}{2}} z_t\big) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} c_{\boldsymbol{\ell}} H_{\boldsymbol{\ell}}(z_t), \qquad t \geqslant 0. \tag{2.4}$$

Note that, by definition,

$$z_t = \int_{\mathbf{R}} \hat{K}(t-u) dW_u, \qquad \hat{K}(t) \stackrel{\text{def}}{=} D^{-\frac{1}{2}} O^\top K(t). \tag{2.5}$$

and $z \in \mathcal{V}_n(\beta, \hat{\Theta})$ where $\hat{\Theta} = \big|D^{-\frac{1}{2}} O^\top\big|^2 \Theta$. Finally, we note that there is no loss of generality in assuming that $\mathrm{rank}(\Sigma) = n$ for otherwise y lives almost surely in a proper subspace of $\mathbf{R}^n$. Without further notice, we shall resort to this case in the sequel.

Definition 2.3 Let $G \in L^2\big(\mathbf{R}^n, N(0,\Sigma)\big)$ and consider the expansion (2.4).

1. The *Hermite rank* of G (with respect to y) is defined by
$$\mathsf{H}(G) \stackrel{\text{def}}{=} \inf\big\{|\boldsymbol{\ell}| :\ c_{\boldsymbol{\ell}} \neq 0\big\}.$$
2. We say that G satisfies the *fast chaos decay condition* with parameter $p > 1$ if
$$\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} |c_{\boldsymbol{\ell}}| (p-1)^{\frac{|\boldsymbol{\ell}|}{2}} \sqrt{\boldsymbol{\ell}!} < \infty.$$

Remark 2.4 Denoting the Ornstein-Uhlenbeck operator by

$$T_\theta G = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} c_{\boldsymbol{\ell}} e^{-\theta|\boldsymbol{\ell}|} H_{\boldsymbol{\ell}},$$

we have $\mathcal{D}(T_\theta) = \{G \in L^2(\mathbf{R}^n, N(0,\mathrm{id})) : \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} |c_{\boldsymbol{\ell}}|^2 e^{-2\theta|\boldsymbol{\ell}|} \boldsymbol{\ell}! < \infty\}$. Thus, by an application of Cauchy-Schwarz we obtain that a function G satisfies the fast chaos parameter condition with parameter p if for some $\delta > 0$, $G \in \mathcal{D}(T_{-\frac{1}{2}\ln(p-1)-\delta})$.

Remark 2.5 Recalling the generating function

$$e^{\langle x,a\rangle - \frac{|a|^2}{2}} = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} \frac{a^{\boldsymbol{\ell}}}{\boldsymbol{\ell}!} H_{\boldsymbol{\ell}}(x), \qquad a, x \in \mathbf{R}^n,$$

2.2 Malliavin Calculus

In this section we recall some concepts from Malliavin calculus. For details we refer to [Nuao6].

Let $\mathscr{H} = L^2(\mathbf{R}, \mathbf{R}^d)$, w.r.t. the Lebesgue measure. If $W_t = (W_t^1, \dots, W_t^d)$ is a two sided Wiener process, we construct an *isonormal Gaussian process* $\{W(h) : h \in \mathscr{H}\}$ by Itô-Wiener integrals

$$W(h) \stackrel{\text{def}}{=} \int_{\mathbf{R}} \langle h(t), dW_t \rangle,$$

where h is identified with an L^2 function from $\mathbf{R}$ to $\mathbf{R}^d$. If $f \in \mathscr{H}^{\otimes m} \cong L^2(\mathbf{R}^m, \mathbf{R}^{d^m})$, it corresponds to a family of functions $f_{i_1,\dots,i_m} : \mathbf{R}^m \to \mathbf{R}$ where $i_j \in \{1, \dots, d\}$. We denote by f^{Sym} its symmetrisation. We define the multiple Wiener integral with respect to W as follows:

$$I : \bigoplus_{m \geq 0} \mathscr{H}^{\otimes m} \to L^2(\Omega; \mathbf{R}), \qquad \text{for } f \in \mathscr{H}^{\otimes m},$$

$$I(f) \stackrel{\text{def}}{=} m! \sum_{i_1,\dots,i_m=1}^{d} \int_{-\infty}^{\infty} \int_{-\infty}^{t_{m-1}} \cdots \int_{-\infty}^{t_2} f^{\mathrm{Sym}}_{i_1,\dots,i_m}(t_1, \dots, t_m) dW_{t_1}^{i_1} \cdots dW_{t_m}^{i_m}.$$

Given $f = (f_1, \dots, f_n) \in \mathscr{H}^n$ and a multi-index $\boldsymbol{\ell} = (\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_n) \in \mathbf{N}_0^n$, we define $f^{\otimes \boldsymbol{\ell}} \in \mathscr{H}^{\otimes |\boldsymbol{\ell}|}$ by

$$f^{\otimes \boldsymbol{\ell}} \stackrel{\text{def}}{=} f_1^{\otimes \boldsymbol{\ell}_1} \otimes \cdots \otimes f_n^{\otimes \boldsymbol{\ell}_n}. \tag{2.6}$$

To keep the verification of the fourth moment theorem simple we introduce contraction operators between functions that are not necessarily symmetric. If S and $\tilde{S}$ are two index sets, $f = \otimes_{i \in S} f_i$ and $g = \otimes_{i \in \tilde{S}} g_i$ are primitive tensors with $f_i, g_i \in \mathscr{H}$, for any pair $p = \{a, b\}$ where $a \in S$ and $b \in \tilde{S}$ we define the contraction of $f \otimes g$:

$$\mathrm{Tr}_p(f \otimes g) = \langle f_a, g_b \rangle \left(\bigotimes_{i \in S \setminus \{a\}} f_i \right) \otimes \left(\bigotimes_{i \in \tilde{S} \setminus \{b\}} g_i \right). \tag{2.7}$$

We define the multi-contraction with multi-pairs of k-elements. Let $P = \{(a_i, b_i) : a_i \in S, b_i \in \tilde{S}, i = 1, \dots, k\}$, where $\{a_i\}$ are distinct and so are $\{b_i\}$, by:

$$\mathrm{Tr}_P(f \otimes g) = \Pi_i \langle f_{a_i}, g_{b_i} \rangle \left(\bigotimes_{i \in S \setminus \{a_i\}} f_i \right) \otimes \left(\bigotimes_{i \in \tilde{S} \setminus \{b_i\}} g_i \right). \tag{2.8}$$

Then the following product formula holds for $f \in \mathscr{H}^{\otimes m}$ and $g \in \mathscr{H}^{\otimes n}$.

$$I(f)I(g) = \delta_{n,m} \sum_{P \in \mathcal{P}} I(\mathrm{Tr}_P f \otimes g), \qquad \mathcal{P} = \bigcup_{k=0}^{n} \mathcal{P}_k, \tag{2.9}$$

where P runs through all multi-pairs from $S = \{1, \dots, m\}$ and $\tilde{S} = \{1, \dots, n\}$ so $\mathcal{P} = \bigcup_{k=0}^{m \wedge n} \mathcal{P}_k$ where $\mathcal{P}_k$ denotes the collection of all k distinct pairs of indices from $S \times \tilde{S}$. The 0^{th} contraction is $f \otimes g$. In particular,

$$\mathbf{E}[I(f)I(g)] = \delta_{m,n} m! \langle \mathrm{Sym}(f), \mathrm{Sym}(g) \rangle_{\mathscr{H}^{\otimes m}}. \tag{2.10}$$

We have the following straight-forward generalization of [Nuao6, Proposition 1.1.4]:

Lemma 2.6 *Let $\boldsymbol{\ell} \in \mathbf{N}_0^n$ and $f \in \mathscr{H}^n$ with $\|f_i\|_{\mathscr{H}} = 1$ for each $i = 1, \dots, n$. Then we have*

$$H_{\boldsymbol{\ell}}\big(W(f)\big) = I(f).$$

Given $\ell \in \mathbf{N}_0^n$ we call a graph of complete pairings, with no self-loop, with nodes $\{1, \dots, n\}$ ℓ-*admissible* if the k^{th} node has exactly ℓ_k edges. We denote the collection of all such graphs by Γ_ℓ. We will use the fact

$$|\Gamma_\ell| \leqslant \sqrt{\ell!}(n-1)^{\frac{|\ell|}{2}}. \tag{2.11}$$

For a graph $\mathbb{G} \in \Gamma_\ell$, we write $\gamma_{i,j}(\mathbb{G})$ for the number of edges between the nodes i and j.

Proposition 2.7 (Diagram Formula [BH02, Taq77]) *Let $\ell \in \mathbf{N}_0^n$ and $X = (X_1, \dots, X_n)$ be multivariate Gaussian. Then*

$$\mathbf{E}[H_\ell(X)] = \sum_{\mathbb{G} \in \Gamma_\ell} \prod_{i=1}^{n} \prod_{j=1}^{i} \mathbf{E}[X_i X_j]^{\gamma_{i,j}(\mathbb{G})}.$$

In particular, if $k \in \mathbf{N}_0^n$ and $Y = (Y_1, \dots, Y_n)$ is a multivariate Gaussian independent of X, then we have

$$\big|\mathbf{E}[H_\ell(X) H_k(Y)]\big| \lesssim \delta_{|k|,|\ell|} \sqrt{k!\ell!}(2n-1)^{|\ell|} \max_{1 \leqslant i < j \leqslant n} \big|\mathbf{E}[X_i Y_j]\big|^{|\ell|}.$$

2.3 Hölder Spaces and Their Tensor Product

The algebraic tensor product $\mathcal{X} \otimes_a \mathcal{Y}$ of two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined as the subspace of the dual of the bilinear mappings $\mathcal{X} \times \mathcal{Y} \to \mathbf{R}$ spanned by the elements $x \otimes y$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, which act by

$$\langle B, x \otimes y \rangle \stackrel{\text{def}}{=} B(x, y).$$

We can also declare a dual action on $\mathcal{X} \otimes_a \mathcal{Y}$ by $\langle x \otimes y, x^* \otimes y^* \rangle \stackrel{\text{def}}{=} \langle x, x^* \rangle \langle y, y^* \rangle$ for x^*, y^* in the dual spaces of $\mathcal{X}$ and $\mathcal{Y}$, respectively.

If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, we call a norm $\|\cdot\|_{\mathcal{X}\otimes\mathcal{Y}}$ on $\mathcal{X} \otimes_a \mathcal{Y}$ a *reasonable crossnorm* if

$$\begin{aligned} \|x \otimes y\|_{\mathcal{X}\otimes\mathcal{Y}} &= \|x\|_{\mathcal{X}} \|y\|_{\mathcal{Y}} && \forall\, x \in \mathcal{X},\ y \in \mathcal{Y}, \\ \|x^* \otimes y^*\|_{(\mathcal{X}\otimes\mathcal{Y})^*} &\stackrel{\text{def}}{=} \sup_{z \in \mathcal{X}\otimes_a\mathcal{Y}} \frac{|\langle z, x^* \otimes y^* \rangle|}{\|z\|_{\mathcal{X}\otimes\mathcal{Y}}} = \|x^*\|_{\mathcal{X}^*} \|y^*\|_{\mathcal{Y}^*} && \forall\, x^* \in \mathcal{X}^*,\ y^* \in \mathcal{Y}^*. \end{aligned}$$

It is easy to see that both canonical examples, the injective and the projective tensor norm, define reasonable cross norms. Finally, the tensor product of $\mathcal{X}$ and $\mathcal{Y}$ is defined as the completion of $\mathcal{X} \otimes_a \mathcal{Y}$ with respect to the norm $\|\cdot\|_{\mathcal{X}\otimes\mathcal{Y}}$. Without further notice, we always assume that $\mathcal{X} \otimes \mathcal{Y}$ is is equipped with a reasonable crossnorm. Further details on the tensor product of Banach spaces can be found in the classical monographs [LC85, Rya02].

We now turn to the tensor product of interest in the sequel of this work. Let $\alpha > 0$ and $\mathcal{X}$ be normed space. The classical Hölder space $\mathcal{C}^\alpha(\mathbf{R}^d, \mathcal{X})$ is defined as the collection

$$\begin{aligned} \mathcal{C}^\alpha(\mathbf{R}^d, \mathcal{X}) &\stackrel{\text{def}}{=} \big\{ f \in \mathcal{C}^{\lfloor \alpha \rfloor}(\mathbf{R}^d, \mathcal{X}) : |f|_{\mathcal{C}^\alpha} < \infty \big\}, \\ |f|_{\mathcal{C}^\alpha} &\stackrel{\text{def}}{=} \sup_{\substack{\ell \in \mathbf{N}_0^d \\ |\ell| \leqslant \lfloor \alpha \rfloor}} \sup_{x \in \mathbf{R}^d} |D^\ell f(x)|_{\mathcal{X}} + \sup_{\substack{\ell \in \mathbf{N}_0^d \\ |\ell| = \lfloor \alpha \rfloor}} \sup_{x \neq y} \frac{|D^\ell f(x) - D^\ell f(y)|_{\mathcal{X}}}{|x - y|^{\alpha - \lfloor \alpha \rfloor}}. \end{aligned}$$

We also write $f \in \mathcal{C}^{\alpha+}(\mathbf{R}^d, \mathcal{X})$ if there is an $\alpha' > \alpha$ such that $f \in \mathcal{C}^{\alpha'}(\mathbf{R}^d, \mathcal{X})$.

We will use the following result. It was shown in [KM17, Cor. 4.6] for $\mathcal{C}^\alpha$ instead of for $\mathcal{C}_b^\alpha$ with a different $\mathcal{C}^\alpha$ norm.

Lemma 2.8 *The canonical embedding*

$$\begin{aligned} \imath &: \mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d) \otimes \mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d) \hookrightarrow \mathcal{C}_b^\alpha\big(\mathbf{R}^d, \mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d \otimes \mathbf{R}^d)\big), \\ \imath(f \otimes g)(x, y) &\stackrel{\text{def}}{=} f(x) \otimes g(y) \qquad \forall\, f, g \in \mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d),\ x, y \in \mathbf{R}^d, \end{aligned}$$

extending to the whole space by linearity, defines a reasonable cross norm on the former, with respect to this the embedding is continuous.

2.4 Rough Path Theory

The theory of rough paths has by now certainly found its way into the mathematical mainstream. This is, of course, also due to the very nice monographs [FV10, FH14] to which we refer for further details. We shall work in the framework of controlled rough paths [Gubo4, FdLPo6] popularized in the book of Friz and Hairer.

Let $X, Y : [0,T] \to \mathcal{X}$ be Hölder continuous with exponents γ_1 and γ_2, respectively. If $\gamma_1 + \gamma_2 > 1$, Young's integration theory enables us to define $\int_0^T Y dX$ as the limit of Riemann sums $\sum_{[u,v]\in\mathcal{P}} Y_u(X_v - X_u)$ along any sequence of partitions $\mathcal{P}$ of $[0,T]$ with mesh $|\mathcal{P}|$ tending to 0 [You36]. Furthermore, the mapping $(X,Y) \mapsto \int_0^\cdot Y dX$ is continuous. Thus, for $X \in \mathcal{C}^{\frac{1}{2}+}$, one expects a unique solution theory to the Young differential equation $dY_s = f(Y_s)dX_s$ (given enough regularity on f). Indeed, if $f \in \mathcal{C}_b^2$, this equation is well posed and the solution is continuous in both the driver X and the initial data. In the case $\gamma_1 \leqslant \frac{1}{2}$, this fails and one cannot define the integral $\int X dX$ by the above Riemann sum anymore. This is partially remedied by rough path theory, which allows to define an integral with respect to less regular integrators by enhancing the Riemann sum, see (2.14) below.

Let $T > 0$, $\gamma \in (0,1)$, and $\mathcal{X}$ be a Banach space. We write $|X|_{\mathcal{C}^\gamma}$ for the γ-Hölder norm of a function $X : [0,T] \to \mathcal{X}$. The space of all γ-Hölder functions $[0,T] \to \mathcal{X}$ is denoted by $\mathcal{C}^\gamma\big([0,T],\mathcal{X}\big)$. We set $\Delta_T \stackrel{\text{def}}{=} \{(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ and

$$\mathcal{C}^\gamma(\Delta_T, \mathcal{X}\otimes\mathcal{X}) \stackrel{\text{def}}{=} \big\{\mathbb{X} : \Delta_T \to \mathcal{X}\otimes\mathcal{X} : |\mathbb{X}|_\gamma < \infty\big\}, \qquad |\mathbb{X}|_\gamma \stackrel{\text{def}}{=} \sup_{(s,t)\in\Delta_T} \frac{\|\mathbb{X}_{s,t}\|_{\mathcal{X}\otimes\mathcal{X}}}{|t-s|^\gamma},$$

where $\|\cdot\|_{\mathcal{X}\otimes\mathcal{X}}$ is the norm on $\mathcal{X}\otimes\mathcal{X}$. Note that any γ-Hölder function $X \in \mathcal{C}^\gamma\big([0,T],\mathcal{X}\big)$ has a natural lift to $\mathcal{C}^\gamma(\Delta_T,\mathcal{X})$ by declaring $X_{s,t} \stackrel{\text{def}}{=} X_t - X_s$. An γ-rough path with values in $\mathcal{X}$ is a pair $\mathbf{X} = (X,\mathbb{X}) \in \mathcal{C}^\gamma\big([0,T],\mathcal{X}\big) \oplus \mathcal{C}^{2\gamma}\big([0,T],\mathcal{X}\otimes\mathcal{X}\big)$ obeying the algebraic constraint (*Chen's relation*)

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u}\otimes X_{u,t} \qquad \forall\, 0 \leqslant s \leqslant u \leqslant t \leqslant T. \tag{2.12}$$

Owing to this identity, the values of the two-parameter process $\mathbb{X}$ can actually be recovered from the knowledge of $\mathbb{X}_t \stackrel{\text{def}}{=} \mathbb{X}_{0,t}$, $t \in [0,T]$:

$$\mathbb{X}_{s,t} = \mathbb{X}_t - \mathbb{X}_s - X_s \otimes X_t.$$

It is customary to denote the set of γ-rough paths with values in $\mathcal{X}$ by $\mathscr{C}^\gamma\big([0,T],\mathcal{X}\big)$. Albeit this space is certainly not linear, it becomes a complete metric space in the topology inherited from the Banach space $\mathcal{C}^\gamma\big([0,T],\mathcal{X}\big) \oplus \mathcal{C}^{2\gamma}\big([0,T],\mathcal{X}^2\big)$ which is canonically equipped with the norm

$$|\mathbf{X}|_{\mathcal{C}^\gamma\oplus\mathcal{C}^{2\gamma}} \stackrel{\text{def}}{=} |X|_{\mathcal{C}^\gamma} + |\mathbb{X}|_{\mathcal{C}^{2\gamma}}.$$

We emphasize that—unless $\mathcal{X} = \{0\}$—the space $\mathscr{C}^\gamma\big([0,T],\mathcal{X}\big)$ is not separable. In order to avoid norm versus seminorm considerations, we shall tacitly assume that $X_0 = 0$ which anyways the case in our ultimate application of the theory.

Let, $f \in \mathcal{C}^2\big(\mathbf{R}^d, \mathbb{L}(\mathcal{X},\mathbf{R}^d)\big)$, $Y \in \mathcal{C}^\gamma([0,T],\mathbf{R}^d)$, and $\mathbf{X} \in \mathscr{C}^\gamma([0,T],\mathcal{X})$. For $\gamma \in (\frac{1}{3},\frac{1}{2}]$ the integral

$$\int_0^\cdot f(Y_s)\, d\mathbf{X}_s \in \mathcal{C}^\gamma([0,T],\mathbf{R}^d) \tag{2.13}$$

can be defined, provided that Y is controlled by X. This is to say, there is a $Y' \in \mathcal{C}^\gamma\big([0,T],\mathbb{L}(\mathcal{X},\mathbf{R}^d)\big)$ such that

$$\big|Y_{s,t} - Y'_s X_{s,t}\big| = \mathcal{O}\big(|t-s|^{2\gamma}\big), \qquad 0 \leqslant s,t \leqslant T.$$

In this case, it is also customary to write $(Y,Y') \in \mathscr{D}_X^{2\gamma}([0,T],\mathbf{R}^d)$. Under these conditions, the integral (2.13) is well defined as the limit of *compensated* Riemann sums along an arbitrary sequence of partitions with mesh tending to 0:

$$\int_s^t f(Y_s)\, d\mathbf{X}_s \stackrel{\text{def}}{=} \lim_{|\mathcal{P}|\to 0} \sum_{[u,v]\in\mathcal{P}([s,t])} \Xi_{u,v}, \qquad \Xi_{u,v} \stackrel{\text{def}}{=} \Big(f(Y_u)X_{u,v} + \big(Y'_u \odot Df(Y_u)\big)\mathbb{X}_{u,v}\Big). \tag{2.14}$$

Here, we introduced the 'product'

$$\big(Y'_u \odot Df(Y_u)\big)(x \otimes y) \stackrel{\text{def}}{=} \sum_{i=1}^{d} \big(Y'_u(x)\big)_i \partial_i f(Y_u)(y), \qquad x, y \in \mathcal{X}.$$

Employing the algebraic relation (2.12), it is then an easy exercise to check that

$$\delta\Xi_{s,u,t} \stackrel{\text{def}}{=} \Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t} = \mathcal{O}\big(|t-s|^{3\gamma}\big), \qquad 0 \leqslant s \leqslant u \leqslant t \leqslant T,$$

whence the integral (2.13) is indeed well defined by the sewing lemma [Gubo4, FdLPo6].

Proposition 2.9 *Let $Y_0 \in \mathcal{Y}$ and $\gamma \in (\frac{1}{3}, \frac{1}{2}]$. If $\mathbf{X} \in \mathscr{C}^\gamma([0,T], \mathcal{X})$ and $f \in \mathcal{C}^{\frac{1}{\gamma}+}\big(\mathcal{Y}, \mathbb{L}(\mathcal{X}, \mathcal{Y})\big)$, then there is a controlled rough path $\big(Y, f(Y)\big) \in \mathcal{D}_X^{2\gamma}([0,T], \mathcal{Y})$ such that*

$$Y_t = Y_0 + \int_0^t f(Y_s)\, d\mathbf{X}_s, \qquad t \in [0,T].$$

Moreover, the solution map $\Phi(Y_0, \mathbf{X}) \stackrel{\text{def}}{=} Y$ is continuous from $\mathcal{Y} \times \mathscr{C}^\gamma([0,T], \mathcal{X})$ to $\mathcal{C}^{\gamma'}([0,T], \mathcal{Y})$ for each $\gamma' < \gamma$.

Remark 2.10 We note that the theory of controlled rough paths usually assumes the continuous embedding

$$\mathbb{L}\big(\mathcal{X}, \mathbb{L}(\mathcal{X}, \mathcal{Y})\big) \hookrightarrow \mathbb{L}(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y}), \tag{2.15}$$

see [FH14, Section 1.5]. It is well known that this is ensured by working with the projective tensor product on $\mathcal{X} \otimes \mathcal{X}$. This, however, turns out to be rather inconvenient when $\mathcal{X}$ is a Hölder space, which is the case of interest in this article. It is in fact beneficial to work with the more explicit reasonable crossnorm introduced in Section 2.3. Thankfully, as already observed by Kelly and Melbourne [KM17, Proof of Theorem 3.3], the proof of Proposition 2.9 in [FH14] does not really require the embedding (4.14), but it is enough to equip the tensor product with a reasonable crossnorm.

3 The Multidimensional Limit Theorem

3.1 Statement of the results

First we need to introduce a bit of notation: $\mathbb{W}$ denotes the Stratonovich lift of an n-dimensional standard Wiener process W, that is, $\mathbb{W}_{s,t}^{i,j} = \int_s^t \big(W_r^i - W_s^i\big) \circ dW_r^j$. For a smooth function X, we let $\mathbb{X}_{s,t} \stackrel{\text{def}}{=} \int_s^t (X_r - X_s) \otimes dX_r$ denote the canonical lift and set $(\Upsilon \otimes \Upsilon)\mathbb{X}_{s,t} \stackrel{\text{def}}{=} \int_s^t \Upsilon X_{s,r} d\Upsilon X_r$ where Υ is a matrix.

Theorem 3.1 *Let $y \in \mathcal{V}_n(\beta, \Theta)$ for some $\beta, \Theta > 0$ and $\hat{\Theta} = \big|D^{-\frac{1}{2}} O^\top\big|^2 \Theta$. Then for any real-valued functions $G_1, \dots, G_N \in L^2\big(\mathbf{R}^n, N(0,\Sigma)\big)$ with*

$$\min_{k=1,\dots,N} \mathsf{H}(G_k) > \frac{1}{\beta}$$

and any terminal time $T > 0$ all of the following hold:

- ▷ **(Finite dimensional distributions)** *If each G_k is of fast chaos decay with parameter $\hat{\Theta}(2n-1)+1$, then*

$$X_t^\varepsilon = \left(\sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_1(y_s)\, ds, \dots, \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_N(y_s)\, ds\right)_{t \in [0,T]} \xrightarrow{\text{f.d.d.}} \Upsilon W, \tag{3.1}$$

where $W = (W_t)_{t \in [0,T]}$ is a standard Wiener process and Υ is the unique non-negative square root of the matrix

$$\Upsilon_{i,j}^2 \stackrel{\text{def}}{=} \int_0^\infty \mathbf{E}[G_i(y_r) G_j(y_0) + G_i(y_0) G_j(y_r)]\, dr, \qquad i,j = 1, \dots, N.$$

▷ **(Functional Central Limit Theorem)** *If each G_k is of fast chaos decay with parameter $\hat{\Theta}(2n-1)(p_k-1)+1$ for $p_k > 2$, then the convergence (3.1) is actually weakly in $\mathcal{C}^\gamma\big([0,T],\mathbf{R}^N\big)$ for any $\gamma < \frac{1}{2} - (\min_k p_k)^{-1}$.*

▷ **(Rough Central Limit Theorem)** *If $\min_{k=1,\dots,N} \mathsf{H}(G_k) > 2\beta^{-1}$ and each G_k is of fast chaos decay with parameter $\hat{\Theta}(4n-1)(p_k-1)+1$, then, for any $\gamma < \frac{1}{2} - (\min_k p_k)^{-1}$, where $p_k > 2$*

$$\mathbf{X}^\varepsilon = \big(X^\varepsilon, \mathbb{X}^\varepsilon\big) \Rightarrow \mathbf{X}$$

in the rough path space $\mathscr{C}^\gamma\big([0,T],\mathbf{R}^N\big)$, where

$$\begin{aligned} \mathbf{X}_{s,t} &= \big(\Upsilon W_{s,t}, \quad (\Upsilon\otimes\Upsilon)\mathbb{W}_{s,t} + \Xi(t-s)\big), \\ \Xi_{i,j} &= \frac{1}{2}\int_0^\infty \mathbf{E}[G_i(y_0)G_j(y_r) - G_i(y_r)G_j(y_0)]\,dr. \end{aligned} \tag{3.2}$$

Theorem 3.1 is a multidimensional version of the result in [GL20b], see also [Geh20] for related limit theorems for Hermite processes.

Remark 3.2 The limiting rough path (3.2) can be rewritten in Itô form as follows. Let $\mathbb{W}^{\text{Itô}}$ be the Itô lift of W. Then

$$\mathbf{X}_{s,t} = \big(\Upsilon W_{s,t}, \quad (\Upsilon\otimes\Upsilon)\mathbb{W}^{\text{Itô}}_{s,t} + \Lambda(t-s)\big),$$

where

$$\Lambda_{i,j} = \int_0^\infty \mathbf{E}[G_i(y_0)G_j(y_r)]\,dr.$$

3.2 Functional Central Limit Theorem

We fix $y \in \mathcal{V}_n(\beta,\Theta)$ and write K for its kernel, and z the normalized process (2.3). Recall, $z \in \mathcal{V}_n(\beta,\hat{\Theta})$ where $\hat{\Theta} = \big|D^{-\frac{1}{2}}O^\top\big|^2\Theta$.

Lemma 3.3 *For each $\boldsymbol{\ell} \in \mathbf{N}_0^n$, $H_{\boldsymbol{\ell}}(z_t)$ can be written in terms of the multiple Itô-Wiener integral below:*

$$H_{\boldsymbol{\ell}}(z_t) = \prod_{i=1}^n I\big((\tau_t\hat{K}^i)^{\otimes\boldsymbol{\ell}_i}\big) = I\big(\tau_t\hat{K}^{\otimes\boldsymbol{\ell}}\big). \tag{3.3}$$

Consequently, $\{H_{\boldsymbol{\ell}}(z_t), \boldsymbol{\ell}\in\mathbf{N}_0^n\}$ is an orthonormal set for any t.

Proof. Given a multi-index $\boldsymbol{\ell} = (\boldsymbol{\ell}_1,\dots,\boldsymbol{\ell}_n) \in \mathbf{N}_0^n$, using Lemma 2.6,

$$H_{\boldsymbol{\ell}}(z_t) = \Pi_{i=1}^n H_{\boldsymbol{\ell}_i}(W(\tau_t\hat{K}^i)) = \Pi_{i=1}^n I\big((\tau_t\hat{K}^i)^{\otimes\boldsymbol{\ell}_i}\big).$$

We used the product formulas for multiple integrals [Nua06, Prop. 1.1.4]. Since $\mathbf{E}[z_0\otimes z_0] = \mathrm{id}$ we see

$$\langle\tau_t\hat{K}^i, \tau_t\hat{K}^j\rangle = \int_{\mathbf{R}}\langle\hat{K}^i(\tau-r), \hat{K}^j(\tau-r)\rangle dr = \int_{\mathbf{R}}\langle\hat{K}^i(-r), \hat{K}^j(-r)\rangle dr = \delta_{i,j},$$

as claimed. □

Let $G \in L^2\big(\mathbf{R}^n, N(0,\Sigma)\big)$. The expansion (2.4) becomes

$$G(y_t) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n, |\boldsymbol{\ell}|\le\mathsf{H}(G)} c_{\boldsymbol{\ell}}H_{\boldsymbol{\ell}}(z_t) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n, |\boldsymbol{\ell}|\le\mathsf{H}(G)} c_{\boldsymbol{\ell}}I\big((\tau_t\hat{K}^i)^{\otimes\boldsymbol{\ell}}\big). \tag{3.4}$$

The first step towards the proof of Theorem 3.1 is to establish a central limit theorem for the finite-dimensional distributions of the vector-valued process

$$\left(\sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}} G_1(y_s)ds, \dots, \sqrt{\varepsilon}\int_0^{\frac{t}{\varepsilon}} G_N(y_s)ds\right)_{t\in[0,1]}. \tag{3.5}$$

The argument proceeds along a well-established pathway, see e.g. [BT13, Geh20, NNZ16]:

1. We first assume $G_1, \dots, G_N \in L^2(\mathbf{R}^n, N(0,\Sigma))$ live in a finite number of chaoses (that is, their Hermite expansion (3.4) is finite) and prove the statement of Theorem 3.1 by invoking the fourth moment theorem of Nualart, Peccati, and Tudor [NP05, PT05], see also [NOL08, NP12].
2. A simple truncation argument then shows that the general case can be reduced to 1.

Proposition 3.4 *Let $m \geqslant 2$ and $\boldsymbol{\ell} \in \mathbf{N}^m$ be a multi-index with ordered components, that is, $\boldsymbol{\ell}_1 \leqslant \cdots \boldsymbol{\ell}_m$. Let $(f^\varepsilon)_{\varepsilon>0} \subset \oplus_{i=1}^m \mathscr{H}^{\otimes \ell_i}$. Denote by f_i the projection of f to $\mathscr{H}^{\otimes \ell_i}$. Assume that for $i,j = 1,\dots,m$, the limits*

$$\Lambda_{i,j} \stackrel{\text{def}}{=} \lim_{\varepsilon\to 0} \mathbf{E}\big[I(f_i^\varepsilon) I(f_j^\varepsilon)\big]$$

exists, and for any $1 \leqslant r < k \leqslant m$,

$$\lim_{\varepsilon\to 0} \operatorname{Tr}_P f_k^\varepsilon \otimes f_k^\varepsilon = 0, \qquad \forall P \in \bigcup_{i=1}^{k-1} \mathcal{P}_i,$$

i.e. all contractions of f_k^ε with itself vanish, except for the 0^{th} and the k^{th}. Then we have

$$\big(I_{\boldsymbol{\ell}_1}(f_1^\varepsilon), \dots, I_{\boldsymbol{\ell}_m}(f_m^\varepsilon)\big) \Rightarrow \mathcal{N}(0,\Lambda).$$

We verify that the kernel of the normalized process falls in the regime of the Fourth Moment Theorem. This is done in the following two key lemmas. Note that, even though $\{\tau_t \hat{K}^i\}$ are orthonormal, for two different times $\tau_u \hat{K}^i$ and $\tau_u \hat{K}^j$ may not be orthogonal for $i \neq j$. If they did then the right hand side vanish for $\boldsymbol{k} \neq \boldsymbol{\ell}$ and the formula simplifies.

Lemma 3.5 *Let $\boldsymbol{k}, \boldsymbol{\ell} \in \mathbf{N}_0^n$ with $|\boldsymbol{\ell}| \wedge |\boldsymbol{k}| > \frac{1}{\beta}$ and $s,t \in [0,T]$, then*

$$\lim_{\varepsilon\to 0} \varepsilon \int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{s}{\varepsilon}} \mathbf{E}[H_{\boldsymbol{\ell}}(z_u) H_{\boldsymbol{k}}(z_v)]\, du\, dv = \delta_{|\boldsymbol{k}|,|\boldsymbol{\ell}|}(s\wedge t) \int_0^\infty \mathbf{E}[H_{\boldsymbol{\ell}}(z_u) H_{\boldsymbol{k}}(z_0)] + \mathbf{E}[H_{\boldsymbol{\ell}}(z_0) H_{\boldsymbol{k}}(z_u)]\, du.$$

Proof. Use $H_{\boldsymbol{\ell}}(z_u) = I\big(\tau_u \hat{K}^{\otimes \boldsymbol{k}}\big)$ and by (2.10), we may assume $|\boldsymbol{\ell}| = |\boldsymbol{k}|$ and also $s \leqslant t$.

$$\mathbf{E}\Big[I\big(\tau_u \hat{K}^{\otimes \boldsymbol{k}}\big) I\big(\tau_v \hat{K}^{\otimes \boldsymbol{\ell}}\big)\Big] = \mathbf{E}\left[\sum_{P\in\mathcal{P}} I\big(\operatorname{Tr}_P\big(\tau_u \hat{K}^{\otimes \boldsymbol{\ell}}\big) \otimes \big(\tau_v \hat{K}^{\otimes \boldsymbol{k}}\big)\big)\right]$$

Let $\bar{\hat{K}}(t) \stackrel{\text{def}}{=} \tau_0 \hat{K} = \hat{K}(-t)$. We will use the fact that for $v \leq u$, the right hand side are sums of the products of the form $\int_{\mathbf{R}} \langle \hat{K}^i(u-r), \bar{\hat{K}}^j(r-v)\rangle dr$ is a function of $u-v$. We first take $s=t$, and by change of variable,

$$\begin{aligned}\varepsilon \int_0^{\frac{s}{\varepsilon}} \int_0^{\frac{s}{\varepsilon}} \mathbf{E}\Big[I\big(\tau_u \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_v \hat{K}^{\otimes \boldsymbol{k}}\big)\Big]\, du\, dv &= \int_0^{\frac{s}{\varepsilon}} (s-\varepsilon u) \mathbf{E}\Big[I\big(\tau_u \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_0 \hat{K}^{\otimes \boldsymbol{k}}\big) + I\big(\tau_0 \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_u \hat{K}^{\otimes \boldsymbol{k}}\big)\Big]\, du \\ &\xrightarrow{\varepsilon\to 0} s \int_0^\infty \mathbf{E}\Big[I\big(\tau_u \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_0 \hat{K}^{\otimes \boldsymbol{k}}\big) + I\big(\tau_0 \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_u \hat{K}^{\otimes \boldsymbol{k}}\big)\Big]\, du.\end{aligned}$$

Since

$$\left|\varepsilon \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_0^{\frac{s}{\varepsilon}} \mathbf{E}\Big[I\big(\tau_u \hat{K}^{\otimes \boldsymbol{\ell}}\big) I\big(\tau_v \hat{K}^{\otimes \boldsymbol{k}}\big)\Big]\, du\, dv\right| \lesssim \varepsilon \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_0^{\frac{s}{\varepsilon}} \big(1 \wedge |v-u|^{-\beta|\boldsymbol{k}|}\big)\, du\, dv \to 0,$$

the claim follows at once. ∎

Next, we show that the contractions vanish as we send $\varepsilon \to 0$:

Lemma 3.6 *Let $\boldsymbol{\ell} \in \mathbf{N}_0^n$ and $t \in [0,T]$. If $|\boldsymbol{\ell}| > \beta^{-1}$, then for $r = 1, \dots, |\boldsymbol{\ell}| - 1$, and any multi-pair $P = \{(a_1, b_1), \dots (a_r, b_r)\}$, with both indices of a pair from $\{1, \dots, |\boldsymbol{\ell}|\}$, one has*

$$\mathrm{Tr}_P \int_0^{\frac{t}{\varepsilon}} \tau_u K^{\boldsymbol{\ell}}\, du \otimes \int_0^{\frac{t}{\varepsilon}} \tau_u K^{\boldsymbol{\ell}}\, du \to 0.$$

Proof. Let $\boldsymbol{k}^1, \boldsymbol{k}^2$ be the multi-indices obtained from $\boldsymbol{\ell}$ by deleting $\{a_i\}$ and $\{b_i\}$ respectively.

$$\begin{aligned}
&\left\| \mathrm{Tr}_P \int_0^{\frac{t}{\varepsilon}} \tau_u K^{\boldsymbol{\ell}}\, du \otimes \int_0^{\frac{t}{\varepsilon}} \tau_u K^{\boldsymbol{\ell}}\, du \right\|^2 = \int_{[0,\frac{t}{\varepsilon}]^4} \langle \mathrm{Tr}_P(\tau_{u_1} K^{\boldsymbol{\ell}} \otimes \tau_{u_2} K^{\boldsymbol{\ell}}), \mathrm{Tr}_P(\tau_{u_3} K^{\boldsymbol{\ell}} \otimes \tau_{u_4} K^{\boldsymbol{\ell}}) \rangle\, d^4u \\
&\lesssim \int_{[0,\frac{t}{\varepsilon}]^4} \big(1 \wedge |u_1 - u_2|^{-\beta r}\big)\big(1 \wedge |u_3 - u_4|^{-\beta r}\big) |\langle \tau_{u_1} K^{\boldsymbol{k}^1} \otimes \tau_{u_2} K^{\boldsymbol{k}^2}, \tau_{u_3} K^{\boldsymbol{k}^1} \otimes \tau_{u_4} K^{\boldsymbol{k}^2} \rangle|\, d^4u \\
&\lesssim \int_{[0,\frac{t}{\varepsilon}]^4} \big(1 \wedge |u_1 - u_2|^{-\beta r}\big)\big(1 \wedge |u_3 - u_4|^{-\beta r}\big)\big(1 \wedge |u_1 - u_3|^{-\beta(|\boldsymbol{\ell}|-r)}\big)\big(1 \wedge |u_2 - u_4|^{-\beta(|\boldsymbol{\ell}|-r)}\big)\, d^4u.
\end{aligned}$$

by the correlation decay of the kernel K. To see that this expression is of order $o(\varepsilon^{-2})$ as $\varepsilon \to 0$, we use the inequality $x^\alpha y^\beta \leqslant x^{\alpha+\beta} + y^{\alpha+\beta}$ for $x, y > 0$ and $\alpha, \beta \in \mathbf{R}$ to find

$$\begin{aligned}
&\int_{[0,\frac{t}{\varepsilon}]^4} \big(1 \wedge |u_1 - u_2|^{-\beta|\boldsymbol{\ell}|}\big)\big(1 \wedge |u_3 - u_4|^{-\beta r}\big)\big(1 \wedge |u_2 - u_4|^{-\beta(|\boldsymbol{\ell}|-r)}\big)\, d^4u \\
&\int_0^{\frac{t}{\varepsilon}} du_4 \int_{-u_4}^{\frac{t}{\varepsilon}-u_4} \big(1 \wedge |u_3|^{-\beta(|\boldsymbol{\ell}|-r)}\big) du_3 \int_{-u_4}^{\frac{t}{\varepsilon}-u_4} \big(1 \wedge |u_2|^{-\beta r}\big) du_2 \int_{-u_2+u_4}^{\frac{t}{\varepsilon}-u_2+u_4} \big(1 \wedge |u_1|^{-\beta|\boldsymbol{\ell}|}\big) du_1 \\
&\lesssim \left(\frac{t}{\varepsilon}\right)^{2-\beta|\boldsymbol{\ell}|} \int_{-\frac{t}{\varepsilon}}^{\frac{2t}{\varepsilon}} \big(1 \wedge |u_3|^{-\beta(|\boldsymbol{\ell}|-r)}\big) du_3 \int_{-\frac{t}{\varepsilon}}^{\frac{2t}{\varepsilon}} \big(1 \wedge |u_2|^{-\beta r}\big) du_2 = \mathcal{O}\Big(\varepsilon^{2\beta|\boldsymbol{\ell}|-4}\Big).
\end{aligned}$$

Thus, the claim follows since $\beta|\boldsymbol{\ell}| > 1$ by assumption. □

Proposition 3.7 *Let $0 \leqslant t_1 < \cdots < t_m \leqslant T$ and $\boldsymbol{\ell}^1, \dots, \boldsymbol{\ell}^k \in \mathbf{N}_0^n$. Then*

$$\begin{pmatrix} \sqrt{\varepsilon} \int_0^{\frac{t_1}{\varepsilon}} H_{\boldsymbol{\ell}^1}(z_u)\, du \\ \vdots \\ \sqrt{\varepsilon} \int_0^{\frac{t_m}{\varepsilon}} H_{\boldsymbol{\ell}^1}(z_u)\, du \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \sqrt{\varepsilon} \int_0^{\frac{t_1}{\varepsilon}} H_{\boldsymbol{\ell}^k}(z_u)\, du \\ \vdots \\ \sqrt{\varepsilon} \int_0^{\frac{t_m}{\varepsilon}} H_{\boldsymbol{\ell}^k}(z_u)\, du \end{pmatrix} \Rightarrow \mathcal{N}(0, \Lambda),$$

where

$$\begin{aligned}
\Lambda &\stackrel{\text{def}}{=} \begin{pmatrix} \Lambda^{1,1} & \cdots & \Lambda^{1,k} \\ \vdots & \ddots & \vdots \\ \Lambda^{k,1} & \cdots & \Lambda^{k,k} \end{pmatrix}, \\
\Lambda^{i,j}_{\alpha,\beta} &\stackrel{\text{def}}{=} (t_\alpha \wedge t_\beta) \delta_{|\boldsymbol{\ell}^i|,|\boldsymbol{\ell}^j|} \int_0^\infty \mathbf{E}[H_{\boldsymbol{\ell}^i}(z_u) H_{\boldsymbol{\ell}^j}(z_0) + H_{\boldsymbol{\ell}^i}(z_0) H_{\boldsymbol{\ell}^j}(z_u)]\, du, \qquad \alpha, \beta = 1, \dots, m.
\end{aligned}$$

Proof. Combining Lemmas 3.3, 3.5 and 3.6 with Proposition 3.4, we proves the claim. □

Remark 3.8 This together with the lemmas below immediately lead to part 1 and 2 of Theorem 3.1 when G_k are polynomial functions: $X^\varepsilon \to \Upsilon W$ in finite dimensional distributions. The second claim follows for polynomials from the Hölder bounds proved below.

The following L^2-estimate, which we will actually lift to an L^p-bound momentarily, plays a key rôle in proving the weak convergence in Hölder topology:

Lemma 3.9 (L^2-Hölder bound) *Let $0 \leqslant s \leqslant t \leqslant T$. Let $G \in L^2(\mathbf{R}^n, N(0,\Sigma))$ be a real-valued function satisfying the fast chaos decay assumption with parameter $(2n-1)\hat{\Theta}$. If $\mathsf{H}(G) > \beta^{-1}$, then*

$$\left\| \sqrt{\varepsilon} \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} G(y_r)\, dr \right\|_{L^2} \lesssim |t-s|^{\frac{1}{2}}.$$

Proof. Owing to stationarity of $(y_t)_{t\geqslant 0}$, there is no loss of generality in assuming $s = 0$. We expand the square as double integral and make use of the expansion (3.4):

$$\mathbf{E}\left[\left(\int_0^{\frac{t}{\varepsilon}} G(y_u)\, du \right)^2 \right] = \sum_{\substack{\boldsymbol{k},\boldsymbol{\ell} \in \mathbf{N}_0^n \\ |\boldsymbol{k}|,|\boldsymbol{\ell}| \geqslant \mathsf{H}(G)}} c_{\boldsymbol{k}} c_{\boldsymbol{\ell}} \int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} \mathbf{E}[H_{\boldsymbol{k}}(z_u) H_{\boldsymbol{\ell}}(z_v)]\, du\, dv. \tag{3.6}$$

The series is absolutely summable, thus one can exchange the order of summation and integration. Recall that

$$\left| \mathbf{E}[z_u \otimes z_v] \right| \leqslant \hat{\Theta}\big(1 \wedge |u-v|^{-\beta}\big). \tag{3.7}$$

Combining this with Proposition 2.7, we find

$$\mathbf{E}[H_{\boldsymbol{k}}(z_u) H_{\boldsymbol{\ell}}(z_v)] \leqslant \delta_{|\boldsymbol{k}|,|\boldsymbol{\ell}|} \sqrt{\boldsymbol{k}!\boldsymbol{\ell}!} \hat{\Theta}^{|\boldsymbol{k}|} (2n-1)^{|\boldsymbol{k}|} \Big(1 \wedge |u-v|^{-|\boldsymbol{k}|\beta} \Big).$$

Inserting this estimate back into (3.6), we get

$$\mathbf{E}\left[\left(\int_0^{\frac{t}{\varepsilon}} G(y_u)\, du \right)^2 \right] \leqslant \sum_{\substack{\boldsymbol{k},\boldsymbol{\ell} \in \mathbf{N}_0^n \\ |\boldsymbol{k}|,|\boldsymbol{\ell}| \geqslant \mathsf{H}(G)}} |c_{\boldsymbol{\ell}}||c_{\boldsymbol{k}}| \sqrt{\boldsymbol{\ell}!\boldsymbol{k}!} \hat{\Theta}^{\frac{|\boldsymbol{\ell}|+|\boldsymbol{k}|}{2}} (2n-1)^{\frac{|\boldsymbol{\ell}|+|\boldsymbol{k}|}{2}} \int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} \Big(1 \wedge |u-v|^{-\mathsf{H}(G)\beta} \Big)\, du\, dv \lesssim \frac{t}{\varepsilon},$$

since $\mathsf{H}(G) > \beta^{-1}$ and $\sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} |c_{\boldsymbol{\ell}}| \hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}} (2n-1)^{\frac{|\boldsymbol{\ell}|}{2}} \sqrt{\boldsymbol{\ell}!} < \infty$ by assumption. This concludes the proof. □

The following estimate will be used in Sec. 3.4.

Lemma 3.10 (L^p-Hölder bound) *Let $0 \leqslant s \leqslant t \leqslant T$. Let $G \in L^2(\mathbf{R}^n, N(0,\Sigma))$. If $\mathsf{H}(G) > \beta^{-1}$ and G satisfies the fast chaos decay assumption with parameter $(2n-1)(p-1)\hat{\Theta} + 1$, where $p > 2$, then*

$$\left\| \sqrt{\varepsilon} \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} G(y_r)\, dr \right\|_{L^p} \lesssim |t-s|^{\frac{1}{2}}.$$

Proof. As in the proof of Lemma 3.9, we may assume $s = 0$ without any loss of generality. Recall that, by Gaussian hypercontractivity, $\|X\|_{L^p} \leqslant (p-1)^{\frac{|\boldsymbol{\ell}|}{2}} \|X\|_{L^2}$ for any scalar random variable in the $\boldsymbol{\ell}^{\text{th}}$ Wiener chaos. It follows that

$$\begin{aligned} \left\| \int_0^{\frac{t}{\varepsilon}} G(y_r)\, dr \right\|_{L^p} &\leqslant \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} |c_{\boldsymbol{\ell}}| (p-1)^{\frac{|\boldsymbol{\ell}|}{2}} \left\| \int_0^{\frac{t}{\varepsilon}} H_{\boldsymbol{\ell}}(z_r)\, dr \right\|_{L^2} \\ &\lesssim \frac{t}{\varepsilon} \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n,} |c_{\boldsymbol{\ell}}| (p-1)^{\frac{|\boldsymbol{\ell}|}{2}} \sqrt{\boldsymbol{\ell}!} \hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}} (2n-1)^{\frac{|\boldsymbol{\ell}|}{2}}. \end{aligned}$$

The statement follows since the sum is finite by the fast chaos decay assumption, c.f. Definition 2.3 2. □

3.3 Lifted Functional Central Limit Theorem

The aim of this section is to prove the convergence of the iterated integrals. We begin with a definition inspired by [JS03]. Let us write $\mathcal{F}_t \overset{\text{def}}{=} \sigma(W_s, s \leqslant t)$ for the filtration generated by the Wiener process driving y through (2.1).

Definition 3.11 We say that a function $G : \mathbf{R}^n \to \mathbf{R}$ with $\mathbf{E}[G(y_0)] = 0$ satisfies the *conditional decay condition* (with respect to y) if

$$\int_0^\infty \left\| \mathbf{E}[G(y_s)|\mathcal{F}_0] \right\|_{L^2(\Omega)} ds < \infty.$$

As we shall see in the sequel, if $y \in \mathcal{V}_n(\beta, \Theta)$, then any centered $G \in L^2(\mathbf{R}^n, N(0,\Sigma))$ with $\mathsf{H}(G) > 2\beta^{-1}$ falls in the regime of Definition 3.11.

Lemma 3.12 *Let y be any stationary stochastic process which is ergodic under the canonical time-shift. For each $k \in \{1, \dots, N\}$, fix a function $G_k : \mathbf{R}^n \to \mathbf{R}$ satisfying the conditional decay condition with respect to y and set*

$$X_t^\varepsilon = \left(\sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_1(y_s)\, ds, \dots, \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_N(y_s)\, ds \right).$$

Then we have the decomposition

$$X_t^\varepsilon = M_t^\varepsilon + (Z_0^\varepsilon - Z_t^\varepsilon),$$

their respective k^{th}-components are:

$$M_t^{k,\varepsilon} \overset{\text{def}}{=} \sqrt{\varepsilon} \int_0^\infty \mathbf{E}\Big[G_k(y_s) | \mathcal{F}_{\frac{t}{\varepsilon}} \Big] ds - \sqrt{\varepsilon} \int_0^\infty \mathbf{E}[G_k(y_s)|\mathcal{F}_0]\, ds$$

is an L^2-bounded martingale with respect to the rescaled filtration $\left(\mathcal{F}_{\frac{t}{\varepsilon}}\right)_{t \geqslant 0}$ and

$$Z_t^{k,\varepsilon} \overset{\text{def}}{=} \sqrt{\varepsilon} \int_{\frac{t}{\varepsilon}}^\infty \mathbf{E}\Big[G_k(y_s) | \mathcal{F}_{\frac{t}{\varepsilon}} \Big]\, ds.$$

Proof. Due to the conditional decay condition, the processes M^ε and Z^ε are well defined and, in particular, M^ε is uniformly bounded in L^2. The rest of the lemma follows by an easy computation. □

Proposition 3.13 *Let y be a stationary ergodic stochastic process and X^ε be as in Lemma 3.12, with the conditional decay condition with respect to y in place. Let W be a standard Wiener process and let*

$$\Upsilon_{i,j}^2 = \int_0^\infty \mathbf{E}[G_i(y_0)G_j(y_r) + G_i(y_r)G_j(y_0)] dr.$$

If $X^\varepsilon \Rightarrow \Upsilon W$ in $\mathcal{C}\big([0,1], \mathbf{R}^N\big)$, as $\varepsilon \to 0$, then

$$\mathbf{X}^\varepsilon = \big(X^\varepsilon, \mathbb{X}^\varepsilon\big) \Rightarrow \mathbf{X}$$

in $\mathcal{C}\big([0,1], \mathbf{R}^N \oplus \mathbf{R}^N \otimes \mathbf{R}^N\big)$ where

$$\mathbf{X}_{s,t} = \big(\Upsilon W_{s,t}, \quad (\Upsilon \otimes \Upsilon)\mathbb{W}_{s,t} + \Xi(t-s)\big) = \big(\Upsilon W_{s,t}, \quad (\Upsilon \otimes \Upsilon)\mathbb{W}_{s,t}^{Itô} + \Lambda(t-s)\big)$$

with $\mathbb{W}_{s,t}$ denoting the Stratonovich integral, Υ the non-negative symmetric square root of Υ^2, and

$$\Xi_{i,j} = \frac{1}{2} \int_0^\infty \mathbf{E}[G_i(y_0)G_j(y_r) - G_i(y_r)G_j(y_0)] dr$$

Proof. Without loss of generality, we may assume $s = 0$ by stationarity. By Lemma 3.12, we have the decomposition

$$\int_0^t X_s^\varepsilon \otimes dX_s^\varepsilon = \int_0^t M_s^\varepsilon \otimes dX_s^\varepsilon + \int_0^t (Z_0^\varepsilon - Z_s^\varepsilon) \otimes dX_s^\varepsilon$$
$$= M_t^\varepsilon \otimes X_t^\varepsilon - \left(\int_0^t X_s^\varepsilon \otimes dM_s^\varepsilon \right)^T + \int_0^t (Z_0^\varepsilon - Z_s^\varepsilon) \otimes dX_s^\varepsilon,$$

Note that $M^\varepsilon \otimes X^\varepsilon = X^\varepsilon \otimes X^\varepsilon + \mathfrak{o}(\varepsilon)$ for an error term satisfying $\|\mathfrak{o}(\varepsilon)\|_{L^2} = \mathcal{O}(\sqrt{\varepsilon})$. Consequently, by a standard stability result (see e.g. [KP91, Theorem 2.2]) and an integration by parts, we see that, jointly with X^ε, the first two terms converge weakly to

$$(\Upsilon \otimes \Upsilon)(W_t \otimes W_t) - \left(\int_0^t (\Upsilon W)_s \otimes d(\Upsilon W)_s \right)^T = \int_0^t (\Upsilon W)_s \otimes d(\Upsilon W)_s + t\Upsilon^2 = (\Upsilon \otimes \Upsilon)\mathbb{W}_{0,t} + \frac{t}{2}\Upsilon^2.$$

By a standard estimate, see Lemma 3.14 below, this convergence not only holds for fixed times $t \in [0,1]$, but actually weakly in $\mathcal{C}\big([0,1], \mathbf{R}^N\big)$.

To conclude, we observe that

$$\int_0^t (Z_0^\varepsilon - Z_s^\varepsilon) \otimes dX_s^\varepsilon = \varepsilon \int_0^{\frac{t}{\varepsilon}} \left(\int_0^\infty \mathbf{E}[G(y_r)|\mathcal{F}_0] \otimes G(y_s)\, dr - \int_s^\infty \mathbf{E}[G(y_r)|\mathcal{F}_s] \otimes G(y_s)\, dr \right) ds.$$

By Birkhoff's ergodic theorem and stationarity, for each fixed $t \in [0,1]$ the first term tends to 0 and the second term converges to

$$-t \int_0^\infty \mathbf{E}[G_i(y_r) G_j(y_0)] dr$$

almost surely as $\varepsilon \to 0$, which together with $\frac{t}{2}\Upsilon^2$ leads to the term $t\Xi$. This convergence actually takes place in $\mathcal{C}\big([0,1], \mathbf{R}^N\big)$, again by Lemma 3.14. Hence, as the limit is deterministic, the proof is complete. □

Lemma 3.14 *Let $f : \mathbf{R} \to \mathbf{R}$ be locally bounded. If $\lim_{\varepsilon \to 0} \varepsilon f\big(\frac{t}{\varepsilon}\big) = 0$ for any $t \in [0,1]$, then also*

$$\lim_{\varepsilon \to 0} \sup_{t \in [0,1]} \left| \varepsilon f\left(\frac{t}{\varepsilon}\right) \right| = 0.$$

We return to the representation $y_t = \int_{\mathbf{R}} K(t-u)\, dW_u$ where $K : \mathbf{R} \to \mathbb{L}(\mathbf{R}^d, \mathbf{R}^n)$ and set

$$\bar{y}_t^\tau \stackrel{\text{def}}{=} \int_{-\infty}^\tau K(t-u)\, dW_u, \qquad \tilde{y}_t^\tau \stackrel{\text{def}}{=} \int_\tau^t K(t-u)\, dW_u.$$

Note that, for every $\tau > 0$, $y_t = \bar{y}_t^\tau + \tilde{y}_t^\tau$. Moreover, $\bar{y}^\tau \in \mathcal{F}_\tau$, whereas $\tilde{y}^\tau$ is independent of $\mathcal{F}_\tau$. As the normalized process $z_t = D^{-\frac{1}{2}} O^\top y_t$ is nothing but a linear transformation of y, we also have the decomposition $z_t = \bar{z}_t^\tau + \tilde{z}_t^\tau$, where

$$\bar{z}_t^\tau = D^{-\frac{1}{2}} O^\top \bar{y}_t^\tau, \qquad \tilde{z}_t^\tau = D^{-\frac{1}{2}} O^\top \tilde{y}_t^\tau.$$

By independence, $\|\bar{z}_t^\tau\|_{L^2}^2 + \|\tilde{z}_t^\tau\|_{L^2}^2 = 1$ necessarily holds.

Lemma 3.15 *For $t \geqslant \tau$, we have that*

$$\|\bar{z}_t^\tau\|_{L^2} \leqslant \hat{\Theta}\big(1 \wedge |t-\tau|^{-\beta}\big).$$

Proof. For $j = 1, \dots, n$ let $\bar{z}^{j,\tau}$ be the j^{th} component of $\bar{z}^\tau$. We have that

$$\mathbf{E}\Big[\big(\bar{z}_t^{j,\tau}\big)^2\Big] = \int_{-\infty}^{-(t-\tau)} \langle \hat{K}_j(-u), \hat{K}_j(-u) \rangle\, du \leqslant \hat{\Theta}\big(1 \wedge |t-\tau|^{-2\beta}\big),$$

as required. □

Lemma 3.16 *Let $\boldsymbol{\ell} \in \mathbf{N}_0^n$ be a multi-index. Then, for each $t \geqslant \tau$,*

$$\big\|\mathbf{E}[H_{\boldsymbol{\ell}}(z_t) \,|\, \mathcal{F}_\tau]\big\|_{L^2} \leqslant \sqrt{\boldsymbol{\ell}!}\hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}}(2n-1)^{\frac{|\boldsymbol{\ell}|}{2}}\Big(1 \wedge |t-\tau|^{-\frac{|\boldsymbol{\ell}|}{2}\beta}\Big).$$

Proof. We use the well-known expansion formula for Hermite polynomial. For any $a, b \in \mathbf{R}$ with $a^2+b^2=1$:

$$H_m(ax+by) = \sum_{j=0}^{m} \binom{m}{j} a^j b^{m-j} H_j(x) H_{m-j}(y) \qquad \forall\, x, y \in \mathbf{R}.$$

Taking $a = \|\bar{z}_t^\tau\|_{L^2}$ and $b = \|\tilde{z}_t^\tau\|_{L^2}$, we obtain

$$\begin{aligned} H_{\boldsymbol{\ell}}(z_t) &= \prod_{k=1}^{n} H_{\boldsymbol{\ell}_k}(z_t^k) = \prod_{k=1}^{n}\sum_{j_k=0}^{\boldsymbol{\ell}_k} \binom{\boldsymbol{\ell}_k}{j_k} \big\|\bar{z}_t^{k,\tau}\big\|_{L^2}^{j_k} \big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}^{\boldsymbol{\ell}_k - j_k} H_{j_k}\left(\frac{\bar{z}_t^{k,\tau}}{\big\|\bar{z}_t^{k,\tau}\big\|_{L^2}}\right) H_{\boldsymbol{\ell}_k - j_k}\left(\frac{\tilde{z}_t^{k,\tau}}{\big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}}\right) \\ &= \sum_{j_1=0}^{\boldsymbol{\ell}_1} \cdots \sum_{j_n=0}^{\boldsymbol{\ell}_n} \prod_{k=1}^{n} \binom{\boldsymbol{\ell}_k}{j_k} \big\|\bar{z}_t^{k,\tau}\big\|_{L^2}^{j_k} \big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}^{\boldsymbol{\ell}_k - j_k} H_{j_k}\left(\frac{\bar{z}_t^{k,\tau}}{\big\|\bar{z}_t^{k,\tau}\big\|_{L^2}}\right) H_{\boldsymbol{\ell}_k - j_k}\left(\frac{\tilde{z}_t^{k,\tau}}{\big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}}\right). \end{aligned}$$

Since $\tilde{z}_t^{k,\tau}$ is independent of $\mathcal{F}_\tau$, after taking the conditional expectation all term vanish except those with $j_k = \boldsymbol{\ell}_k$, thus, the sum reduces to one term:

$$\begin{aligned} &\mathbf{E}[H_{\boldsymbol{\ell}}(z_t) \,|\, \mathcal{F}_\tau] \\ &= \sum_{j_1=0}^{\boldsymbol{\ell}_1} \cdots \sum_{j_n=0}^{\boldsymbol{\ell}_n} \prod_{k=1}^{n} \binom{\boldsymbol{\ell}_k}{j_k} \big\|\bar{z}_t^{k,\tau}\big\|_{L^2}^{j_k} \big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}^{\boldsymbol{\ell}_k - j_k} H_{j_k}\left(\frac{\bar{z}_t^{k,\tau}}{\big\|\bar{z}_t^{k,\tau}\big\|_{L^2}}\right) \mathbf{E}\left[H_{\boldsymbol{\ell}_k - j_k}\left(\frac{\tilde{z}_t^{k,\tau}}{\big\|\tilde{z}_t^{k,\tau}\big\|_{L^2}}\right) \,\middle|\, \mathcal{F}_\tau\right] \\ &= \prod_{k=1}^{n} \big\|\bar{z}_t^{k,\tau}\big\|_{L^2}^{\boldsymbol{\ell}_k} H_{\boldsymbol{\ell}_k}\left(\frac{\bar{z}_t^{k,\tau}}{\big\|\bar{z}_t^{k,\tau}\big\|_{L^2}}\right). \end{aligned}$$

The asserted estimate follows by an application of Proposition 2.7 and, after pulling out the multiplication factors, of the diagram formula to $\mathbf{E}[H_{\boldsymbol{\ell}}(\frac{\bar{z}_t^{k,\tau}}{\|\bar{z}_t^{k,\tau}\|_{L^2}})]$. Then the multiplicative factors just cancel. The requires estimate follows from Lemma 3.15 and Cauchy Schwarz, □

Proposition 3.17 *Let $y \in \mathcal{V}_n(\beta, \Theta)$. Then any function $G \in L^2(\mathbf{R}^n, N(0,\Sigma))$, which satisfies $\mathsf{H}(G) > 2\beta^{-1}$ and the fast chaos decay condition with parameter $\hat{\Theta}(2n-1)+1$, also satisfies the conditional decay condition of Definition 3.11.*

Proof. Using the Hermite expansion of G and Lemma 3.16, we compute

$$\big\|\mathbf{E}[G(y_s)|\mathcal{F}_0]\big\|_{L^2} \leqslant \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n, |\boldsymbol{\ell}|\geqslant \mathsf{H}(G)} |c_{\boldsymbol{\ell}}| \big\|\mathbf{E}[H_{\boldsymbol{\ell}}(z_s)|\mathcal{F}_0]\big\|_{L^2} \leqslant \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n, |\boldsymbol{\ell}|\geqslant \mathsf{H}(G)} |c_{\boldsymbol{\ell}}| \sqrt{\boldsymbol{\ell}!}\hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}}(2n-1)^{\frac{|\boldsymbol{\ell}|}{2}}\Big(1 \wedge s^{-\frac{|\boldsymbol{\ell}|}{2}\beta}\Big).$$

Since $\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n} |c_{\boldsymbol{\ell}}|\sqrt{\boldsymbol{\ell}!}\hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}}(2n-1)^{\frac{|\boldsymbol{\ell}|}{2}} < \infty$,

$$\int_0^\infty \big\|\mathbf{E}[G(y_s)|\mathcal{F}_0]\big\|_{L^2}\, ds \leqslant \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n, |\boldsymbol{\ell}|\geqslant \mathsf{H}(G)} |c_{\boldsymbol{\ell}}| \sqrt{\boldsymbol{\ell}!}\hat{\Theta}^{\frac{|\boldsymbol{\ell}|}{2}}(2n-1)^{\frac{|\boldsymbol{\ell}|}{2}} \sup_{|\boldsymbol{\ell}|\geqslant \mathsf{H}(G)} \int_0^\infty \Big(1 \wedge s^{-\frac{|\boldsymbol{\ell}|}{2}\beta}\Big)\, ds < \infty.$$

The finiteness of the integral follows from the assumption that $\mathsf{H}(G) > 2\beta^{-1}$. □

Lemma 3.18 *Let $\boldsymbol{\ell}, \boldsymbol{k} \in \mathbf{N}_0^n$ be multi-indices such that $|\boldsymbol{\ell}| \wedge |\boldsymbol{k}| > \beta^{-1}$. Then*

$$\mathbf{E}\left[\left(\varepsilon \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_{\frac{s}{\varepsilon}}^{r} H_{\boldsymbol{\ell}}(z_r) H_{\boldsymbol{k}}(z_u)\, du\, dr\right)^2\right] \lesssim \boldsymbol{\ell}!\boldsymbol{k}!\hat{\Theta}^{|\boldsymbol{\ell}|+|\boldsymbol{k}|}(4n-1)^{|\boldsymbol{\ell}|+|\boldsymbol{k}|}|t-s|^2$$

Proof. By stationarity we may assume $s = 0$. We certainly have

$$\mathbf{E}\left[\left(\int_0^{\frac{t}{\varepsilon}}\int_0^r H_{\boldsymbol{\ell}}(z_r)H_{\boldsymbol{k}}(z_u)\,du\,dr\right)^2\right] \leqslant \int_{[0,\frac{t}{\varepsilon}]^4}\Big|\mathbf{E}[H_{\boldsymbol{\ell}}(z_{s_1})H_{\boldsymbol{\ell}}(z_{s_2})H_{\boldsymbol{k}}(z_{s_3})H_{\boldsymbol{k}}(z_{s_4})]\Big|\,d^4s. \tag{3.8}$$

Recall that $z_s \sim N(0,\mathrm{id})$ for any $s \geqslant 0$. Consequently, Proposition 2.7 furnishes the estimate

$$\Big|\mathbf{E}[H_{\boldsymbol{\ell}}(z_{s_1})H_{\boldsymbol{\ell}}(z_{s_2})H_{\boldsymbol{k}}(z_{s_3})H_{\boldsymbol{k}}(z_{s_4})]\Big| \leqslant \hat{\Theta}^{|\boldsymbol{\ell}|+|\boldsymbol{k}|}\sum_{\mathbb{G}\in\Gamma_{\boldsymbol{m}}}\prod_{\substack{i,j=1\\ i<j}}^4\Big(1\wedge|s_i-s_j|^{-\beta\gamma_{i,j}(\mathbb{G})}\Big),$$

where $\boldsymbol{m} \stackrel{\text{def}}{=} (\boldsymbol{\ell},\boldsymbol{\ell},\boldsymbol{k},\boldsymbol{k}) \in \mathbf{N}_0^{4n}$ is obtained by simple concatenation of $\boldsymbol{\ell}$ and $\boldsymbol{k}$. Recall, $\Gamma_{\boldsymbol{\ell}}$ denotes the set of graphs of of complete parings with the kth node having exactly $\boldsymbol{\ell}_k$ edges. For a graph $\mathbb{G} \in \Gamma_{\boldsymbol{\ell}}$, we write $\gamma_{i,j}(\mathbb{G})$ for the number of edges between the nodes i and j. For any given $\mathbb{G} \in \Gamma_{\boldsymbol{m}}$, $\gamma_{i,j}(\mathbb{G}) \in \mathbf{N}_0$ satisfy

$$\begin{aligned}
&\gamma_{1,2}(\mathbb{G})+\gamma_{1,3}(\mathbb{G})+\gamma_{1,4}(\mathbb{G}) = |\boldsymbol{\ell}|, \qquad &&\gamma_{1,2}(\mathbb{G})+\gamma_{2,3}(\mathbb{G})+\gamma_{2,4}(\mathbb{G}) = |\boldsymbol{\ell}|,\\
&\gamma_{1,3}(\mathbb{G})+\gamma_{2,3}(\mathbb{G})+\gamma_{3,4}(\mathbb{G}) = |\boldsymbol{k}|, \qquad &&\gamma_{1,4}(\mathbb{G})+\gamma_{2,4}(\mathbb{G})+\gamma_{3,4}(\mathbb{G}) = |\boldsymbol{k}|.
\end{aligned}$$

Owing to (2.11), the proof is concluded upon verifying that

$$\int_{[0,\frac{t}{\varepsilon}]^4}\prod_{\substack{i,j=1\\ i<j}}^4\Big(1\wedge|s_i-s_j|^{-\beta\gamma_{i,j}(\mathbb{G})}\Big)\,d^4s \lesssim \left(\frac{t}{\varepsilon}\right)^2, \tag{3.9}$$

uniformly in $\mathbb{G} \in \Gamma_{\boldsymbol{m}}$. To this end, let us fix such a graph and henceforth suppress it in our notation. We abbreviate $f(s) \stackrel{\text{def}}{=} 1\wedge|s|^{-\beta}$, $\nu_k \stackrel{\text{def}}{=} \sum_{i=1}^{k-1}\gamma_{i,k}$, and $\tau_k \stackrel{\text{def}}{=} \sum_{i=k+1}^4\gamma_{k,i}$. Then

$$\int_{[0,\frac{t}{\varepsilon}]^4}\prod_{i<j}f(s_i-s_j)^{\gamma_{i,j}}\,d^4s = \int_{[0,\frac{t}{\varepsilon}]^3}\left(\int_0^{\frac{t}{\varepsilon}}\prod_{j=2}^4 f(s_1-s_j)^{\gamma_{1,j}}\,ds_1\right)\prod_{\substack{i=2\\ i<j}}^4 f(s_i-s_j)^{\gamma_{i,j}}\,d^3s.$$

An application of Hölder's inequality shows that

$$\int_0^{\frac{t}{\varepsilon}}\prod_{j=2}^4 f(s_1-s_j)^{\gamma_{1,j}}\,ds_1 \leqslant \sup_{u\in[0,\frac{t}{\varepsilon}]}\int_0^{\frac{t}{\varepsilon}} f(s_1-u)^{\tau_1}\,ds_1 \leqslant 2\int_0^{\frac{t}{\varepsilon}} f(s_1)^{\tau_1}\,ds_1.$$

Iterating this argument, we find

$$\int_{[0,\frac{t}{\varepsilon}]^4}\prod_{i<j}f(s_i-s_j)^{\gamma_{i,j}}\,d^4s \leqslant 8\prod_{k=1}^4\int_0^{\frac{t}{\varepsilon}} f(s)^{\tau_k}\,ds.$$

Similarly, it holds that

$$\int_{[0,\frac{t}{\varepsilon}]^4}\prod_{i<j}f(s_i-s_j)^{\gamma_{i,j}}\,d^4s \leqslant 8\prod_{k=1}^4\int_0^{\frac{t}{\varepsilon}} f(s)^{\nu_k}\,ds,$$

whence

$$\left(\int_{[0,\frac{t}{\varepsilon}]^4}\prod_{i<j}f(s_i-s_j)^{\gamma_{i,j}}\,d^4s\right)^2 \lesssim \prod_{k=1}^4\int_0^{\frac{t}{\varepsilon}} f(s)^{\tau_k}\,ds\int_0^{\frac{t}{\varepsilon}} f(s)^{\nu_k}\,ds.$$

Since $\tau_k+\nu_k \geqslant |\boldsymbol{\ell}|\wedge|\boldsymbol{k}| > \beta^{-1}$ for each $k = 1,\dots,4$, it is easy to see that the right-hand side is $\lesssim (t/\varepsilon)^4$. This verifies (3.9) and the proof is complete. ∎

Proposition 3.19 *Let $p \geqslant 2$. Let $G_1, G_2 \in L^2(\mathbf{R}^n, N(0,\Sigma))$ satisfy the fast chaos decay condition with parameter $\hat{\Theta}(4n-1)(p-1)+1$. Then*

$$\left\| \varepsilon \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_{\frac{s}{\varepsilon}}^{r} G_1(y_r) G_2(y_u)\, du\, dr \right\|_{L^p} \lesssim |t-s|$$

Proof. We may gain assume $s = 0$ by stationarity. We make use of (2.4) and the hypercontractivity estimate already employed in Lemma 3.10:

$$\begin{aligned}
\left\| \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_{\frac{s}{\varepsilon}}^{r} G_1(y_r) G_2(y_u)\, du\, dr \right\|_{L^p} &\leqslant \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} \sum_{\boldsymbol{k} \in \mathbf{N}_0^n} \left| c^1_{\boldsymbol{\ell}} c^2_{\boldsymbol{k}} \right| \left\| \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_{\frac{s}{\varepsilon}}^{r} H_{\boldsymbol{\ell}}(z_r) H_{\boldsymbol{k}}(z_u)\, du\, dr \right\|_{L^p} \\
&\leqslant \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} \sum_{\boldsymbol{k} \in \mathbf{N}_0^n} \left| c^1_{\boldsymbol{\ell}} c^2_{\boldsymbol{k}} \right| (p-1)^{\frac{|\boldsymbol{\ell}|+|\boldsymbol{k}|}{2}} \left\| \int_{\frac{s}{\varepsilon}}^{\frac{t}{\varepsilon}} \int_{\frac{s}{\varepsilon}}^{r} H_{\boldsymbol{\ell}}(z_r) H_{\boldsymbol{k}}(z_u)\, du\, dr \right\|_{L^2} \\
&\lesssim \frac{t}{\varepsilon} \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} \sum_{\boldsymbol{k} \in \mathbf{N}_0^n} \left| c^1_{\boldsymbol{\ell}} c^2_{\boldsymbol{k}} \right| (p-1)^{\frac{|\boldsymbol{\ell}|+|\boldsymbol{k}|}{2}} \sqrt{\boldsymbol{\ell}!\boldsymbol{k}!}\, \hat{\Theta}^{\frac{|\boldsymbol{k}|+|\boldsymbol{\ell}|}{2}} (4n-1)^{\frac{|\boldsymbol{k}|+|\boldsymbol{\ell}|}{2}}.
\end{aligned}$$

Here, the last estimate relies on Lemma 3.18. The fast chaos decay assumption precisely states that the double sum on the right-hand side is finite. This proves the claim. □

3.4 Proof of Theorem 3.1

With the preparation work earlier, together with Proposition 4.9 below concerning convergence in the rough path topology, we can now prove Theorem 3.1. Let X^ε be as in (3.5).

Proof of Theorem 3.1. We have seen, in Remark 3.8, the convergence of $X^\varepsilon \to \Upsilon W$ in finite dimensional distributions if G_k are polynomials. Let $M \in \mathbf{N}$ and $G \in L^2(\mathbf{R}^n, N(0,\Sigma))$. We denote the function obtained by truncating the Hermite expansion of G at level M by

$$G^M(y_s) \stackrel{\text{def}}{=} \sum_{\substack{\boldsymbol{\ell} \in \mathbf{N}_0^n \\ |\boldsymbol{\ell}| \leqslant M}} c_{\boldsymbol{\ell}} H_{\boldsymbol{\ell}}(z_s).$$

Combining Lemma 3.3 and Proposition 3.7, it follows that, for each $M \in \mathbf{N}$,

$$X^{\varepsilon,M} \stackrel{\text{def}}{=} \left(\sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_1^M(y_s) ds, \dots, \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_N^M(y_s) ds \right)_{t \in [0,1]} \xrightarrow{\text{f.d.d.}} \Upsilon_M W,$$

where now

$$(\Upsilon_M^2)_{i,j} \stackrel{\text{def}}{=} \int_0^\infty \mathbf{E}\big[G_i^M(y_s) G_j^M(y_0) + G_i^M(y_0) G_j^M(y_s) \big]\, ds.$$

The convergence in the finite dimensional distribution follows from Portemanteau theorem, see e.g. [Bil99, Theorem 3.2]: we only need to verify the convergence of each component in probability (since we also have $\Upsilon_M \to \Upsilon$ and therefore $\Upsilon_M W \Rightarrow \Xi W$).

It remains to send $M \to \infty$. For this we observe that, for each $t \in [0,1]$ and each $i = 1, \dots, N$,

$$\begin{aligned}
\mathbf{E}\left[\left(\sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_i(y_s)\, ds - \sqrt{\varepsilon} \int_0^{\frac{t}{\varepsilon}} G_i^M(y_s)\, ds \right)^2 \right] &= \varepsilon \sum_{\substack{\boldsymbol{k}, \boldsymbol{\ell} \in \mathbf{N}_0^n \\ |\boldsymbol{k}|, |\boldsymbol{\ell}| > M}} c_{\boldsymbol{\ell}} c_{\boldsymbol{k}} \int_0^{\frac{t}{\varepsilon}} \int_0^{\frac{t}{\varepsilon}} \mathbf{E}[H_{\boldsymbol{\ell}}(y_u) H_{\boldsymbol{k}}(y_v)]\, du\, dv \\
&\lesssim \sum_{\substack{\boldsymbol{\ell}, \boldsymbol{\ell} \in \mathbf{N}_0^{n_k} \\ |\boldsymbol{k}|, |\boldsymbol{\ell}| > M}} \delta_{|\boldsymbol{\ell}|, |\boldsymbol{k}|} c_{\boldsymbol{\ell}} c_{\boldsymbol{k}} \sqrt{\boldsymbol{\ell}!\boldsymbol{k}!}\, \hat{\Theta}^{|\boldsymbol{\ell}|} (2n-1)^{|\boldsymbol{\ell}|} \to 0
\end{aligned}$$

uniformly in $\varepsilon > 0$ as $M \to \infty$.

Fix $\gamma < \frac{1}{2} - \big(\min_k p_k\big)^{-1}$. The weak convergence of X^ε to ΥW in $\mathcal{C}^\gamma([0,T],\mathbf{R}^N)$ follows from the tightness of the left-hand side of (3.1) in $\mathcal{C}^\gamma([0,T],\mathbf{R}^N)$ (upon slightly decreasing γ). Together with the fact that finite-dimensional distributions uniquely determine Radon measures on $\mathcal{C}^\gamma([0,T],\mathbf{R}^N)$, c.f. [Bog07, Ex. 7.14.79] this follows similarly as in Proposition 4.9 below. We conclude the asserted weak convergence.

It is only left to lift this limit theorem to the iterated integrals. By Proposition 3.17 and the convergence of X^ε to ΥW, the process X^ε, with $y \in \mathcal{V}_n(\beta,\Theta)$ and with $G_k \in L^2(\mathbf{R}^n, N(0,\Sigma))$ satisfying $\mathsf{H}(G) > 2\beta^{-1}$ and the fast chaos decay condition with parameter $\hat{\Theta}(2n-1)+1$, also satisfies the conditions of Proposition 3.13. Applying the latter we conclude the convergence

$$\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon) \xrightarrow{\text{f.d.d.}} \big(\Upsilon W, (\Upsilon\otimes\Upsilon)\mathbb{W}_{s,t} + \Xi(t-s)\big).$$

Furthermore, $\|X^\varepsilon_{s,t}\|_p \lesssim |t-s|^{\frac{1}{2}}$ by Lemma 3.10 and $\|\mathbf{X}^\varepsilon_{s,t}\|_p \lesssim |t-s|^{\frac{1}{2}}$ by Proposition 3.19 for every $\gamma < \frac{1}{2} - (\min_k p_k)^{-1}$, with bounds independent of ε. Thus, $\sup_{\varepsilon\in(0,1]} \mathbf{E}[|\mathbf{X}^\varepsilon|_{\mathcal{C}^\gamma\oplus\mathcal{C}^{2\gamma}}] < \infty$, and an application of Proposition 4.9 concludes the proof. □

4 Rough Homogenization

Consider the families of rough differential equations (RDE)

$$dx^\varepsilon_t = \frac{1}{\sqrt{\varepsilon}} f\big(x^\varepsilon_t, y_{\frac{t}{\varepsilon}}\big)\,dt + g(x^\varepsilon_t)\,d\mathbf{B}_t, \qquad x^\varepsilon_0 = x_0, \qquad t\in[0,T], \tag{4.1}$$

depending on a scale parameter $\varepsilon > 0$. We refer to Section 2.4 for further details on the solution theory to (4.13). Let $D_\ell f$ denote the derivative of f in the x-variable, of order $\boldsymbol{\ell} = (\ell_1,\dots,\ell_d)$. We also write $D_x f$ for the full Jacobian in the x-variable.

To illustrate the method, take the case where $\mathbf{B} \equiv 0$ in (4.13). Even in this case, we interpret the resulting ordinary differential equation (ODE)

$$\dot{x}^\varepsilon_t = \frac{1}{\sqrt{\varepsilon}} f\big(x^\varepsilon_t, y_{\frac{t}{\varepsilon}}\big) \tag{4.2}$$

as a Banach space-valued *rough differential equation* (RDE).

4.1 Casting the Homogenization Problem as RDE

Let us explain how to cast the ODE (4.2) as an infinite-dimensional RDE. To this end, we set

$$X^\varepsilon : [0,T] \to \mathcal{C}^\alpha_b(\mathbf{R}^d,\mathbf{R}^d), \qquad t \mapsto \frac{1}{\sqrt{\varepsilon}}\int_0^t f(\cdot, y_{\frac{s}{\varepsilon}})ds. \tag{4.3}$$

Thus, X^ε is a curve in $\mathcal{C}^\alpha_b(\mathbf{R}^d,\mathbf{R}^d)$. We denote its canonical rough path lift by $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon)$, where

$$\mathbb{X}^\varepsilon_{s,t} = \int_s^t \big(X^\varepsilon_r - X^\varepsilon_s\big)\otimes dX^\varepsilon_r = \frac{1}{\varepsilon}\int_s^t \left(\int_s^r f(\cdot, y_{\frac{u}{\varepsilon}}).\,du\right)\otimes f(\cdot\cdot, y_{\frac{r}{\varepsilon}}))\,dr.$$

This notation has to be understood in terms of the embedding $\mathcal{C}^\alpha_b(\mathbf{R}^d,\mathbf{R}^d)\otimes\mathcal{C}^\alpha_b(\mathbf{R}^d,\mathbf{R}^d) \hookrightarrow \mathcal{C}^\alpha_b(\mathbf{R}^d\times\mathbf{R}^d, \mathbf{R}^d\otimes\mathbf{R}^d)$, see Lemma 2.8. The function $\mathbb{X}^\varepsilon_{s,t} : \mathbf{R}^d\times\mathbf{R}^d \to \mathbf{R}^d\otimes\mathbf{R}^d$ is given by

$$\mathbb{X}^\varepsilon_{s,t}(x_1,x_2) = \frac{1}{\varepsilon}\int_s^t \left(\int_s^r f(x_1, y_{\frac{u}{\varepsilon}}))\,du\right)\otimes f(x_2, y_{\frac{r}{\varepsilon}}))\,dr.$$

Furthermore, define

$$\delta : \mathbf{R}^d \to \mathbb{L}\big(\mathcal{C}^\alpha_b(\mathbf{R}^d,\mathbf{R}^d),\mathbf{R}^d\big), \qquad \text{with } \delta_x(f) = f(x). \tag{4.4}$$

Then δ is sufficiently regular.

Lemma 4.1 *The operator (4.4) inherits the regularity of the image space, that is,*

$$\delta \in \mathcal{C}_b^\alpha\Big(\mathbf{R}^d, \mathbb{L}\big(\mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d), \mathbf{R}^d\big)\Big).$$

We relate (4.12) to the original equation (4.13):

Lemma 4.2 *Let $f \in \mathcal{C}_b^{3,0+}(\mathbf{R}^d \times \mathbf{R}^n, \mathbf{R}^d)$ and $y \in \mathcal{C}^{\bar\gamma}\big([0, \frac{T}{\varepsilon}], \mathbf{R}^n\big)$. Then, for any $\varepsilon > 0$, the RDE below driven by a rough path $\mathbf{X}^\varepsilon$ on $\mathcal{X}$ is well posed,*

$$dx_t^\varepsilon = \delta(x_t^\varepsilon)\, d\mathbf{X}_t^\varepsilon. \tag{4.5}$$

Furthermore, the unique global solution to (4.2) solves the RDE (4.5).

Proof. Well-posedness of (4.5) follows immediately from Proposition 2.9 and Lemma 4.1. For the last part set

$$\Xi_{s,t}^\varepsilon \stackrel{\text{def}}{=} \delta(x_s^\varepsilon)(X_t^\varepsilon - X_s^\varepsilon) + \big(\delta(x_s^\varepsilon) \odot \mathcal{D}\delta(x_s^\varepsilon)\big)\mathbb{X}_{s,t}^\varepsilon,$$
$$\bar\Xi_{s,t}^\varepsilon \stackrel{\text{def}}{=} f(x_s^\varepsilon, y_{\frac{s}{\varepsilon}})(t-s).$$

By the Sewing Lemma, it suffices to show that, for some $\eta > 1$,

$$\big|\Xi_{s,t}^\varepsilon - \bar\Xi_{s,t}^\varepsilon\big| \lesssim |t-s|^\eta,$$

as it implies the integrals obtained by $\Xi_{s,t}^\varepsilon$ and $\bar\Xi_{s,t}^\varepsilon$, are the same, see [Gubo4, FdLPo6, FH14]. Note that, since $\varepsilon > 0$ is fixed, the prefactor in this bound is allowed to diverge as $\varepsilon \to 0$.

We have that

$$D\delta(x) \in \mathbb{L}\Big(\mathbf{R}^d, \mathbb{L}\big(\mathcal{X}, \mathbf{R}^d\big)\Big), \qquad D\delta_x(h)(f) = Df(x)\cdot h.$$

Indeed,

$$\big\|\delta(x+h) - \delta(x) - Df(x)(h)\big\|_{\mathbb{L}(\mathcal{X}, \mathbf{R}^d)} = \sup_{\|f\|_{\mathcal{C}_b^\alpha(\mathbf{R}^d, \mathbf{R}^d)} \leqslant 1} \big|f(x+h) - f(x) - Df(x)(h)\big| = o(|h|)$$

as $|h| \to 0$. Remember that, by assumption, we have $y \in \mathcal{C}^{\bar\gamma}\big([0, \frac{T}{\varepsilon}], \mathbf{R}^m\big)$ with probability 1 and $f \in \mathcal{C}_b^{3,0+}$. It follows that, for some $\eta > 0$,

$$\begin{aligned}\big|\Xi_{s,t}^\varepsilon - \bar\Xi_{s,t}^\varepsilon\big| &\leqslant \frac{1}{\sqrt{\varepsilon}} \int_s^t \big|f\big(x_s^\varepsilon, y_{\frac{r}{\varepsilon}}\big) - f\big(x_s^\varepsilon, y_{\frac{s}{\varepsilon}}\big)\big|\, dr + \frac{1}{\varepsilon}\int_s^t \left(\int_s^r \big|D_x f\big(x_s^\varepsilon, y_{\frac{r}{\varepsilon}}\big)\big|\big|f\big(x_s^\varepsilon, y_{\frac{u}{\varepsilon}}\big)\big|\, du\right) dr \\ &\lesssim |t-s|^{1+\eta},\end{aligned}$$

as required. □

4.2 Statement of the result

Lemma 4.3 *Let $f : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}^d$ be differentiable such that $f(x,\cdot) \in L^2\big(\mathbf{R}^n, N(0,\Sigma); \mathbf{R}^n\big)$ for every $x \in \mathbf{R}^d$. Then, if $\sup_{|x|\leqslant R} \|D_x f(x,\cdot)\|_{L^2} < \infty$ for each $R > 0$,*

$$\inf_x \mathsf{H}\big(D_x f(x,\cdot)\big) \geqslant \inf_x \mathsf{H}\big(f(x,\cdot)\big).$$

Proof. If $\langle f(x,\cdot), H_\ell\rangle = 0$ for some multi-index $\ell \in \mathbf{N}_0^n$, then also $\langle D_x f(x,\cdot), H_\ell\rangle = D_x\langle f(x,\cdot), H_\ell\rangle = 0$. □

We define the Hermite rank of $f(x,\cdot)$ to be

$$\mathsf{H}(f(x,\cdot)) = \inf_{i\in\{1,\dots,n\}} \mathsf{H}(f_i(x,\cdot)),$$

Letting $c_{\boldsymbol{\ell}} = (c^1_{\boldsymbol{\ell}},\dots,c^n_{\boldsymbol{\ell}})$, $c^{\boldsymbol{k}}_{\boldsymbol{\ell}} = (c^{1,\boldsymbol{k}}_{\boldsymbol{\ell}},\dots,c^{n,\boldsymbol{k}}_{\boldsymbol{\ell}})$ where for $i=1,\dots,n$,

$$f_i(x,\cdot) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n,|\boldsymbol{\ell}|\geq \mathsf{H}(f(x,\cdot))} c^i_{\boldsymbol{\ell}}(x) H_{\boldsymbol{\ell}}, \qquad D_{\boldsymbol{k}} f_i(x,\cdot) = \sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n,|\boldsymbol{\ell}|\geq \mathsf{H}(f(x,\cdot))} c^{i,\boldsymbol{k}}_{\boldsymbol{\ell}}(x) H_{\boldsymbol{\ell}}.$$

Condition 4.4 *Let $\gamma \in \left(\frac{1}{3},\frac{1}{2}\right)$ and $p > \max\left(\frac{2}{1-2\gamma},2\right)$. We impose the following conditions on the data in (4.13):*

(i) There is a $\bar\gamma > 0$ such that fast process y has almost sure sample paths in $\mathcal{C}^{\bar\gamma}\big([0,T],\mathbf{R}^n\big)$. Moreover, there are $\beta,\Theta>0$ such that $y\in\mathcal{V}_n(\beta,\Theta)$.

(ii) The process $\mathbf{B}$ is a random rough path of Hölder regularity γ with values in a Banach space $\mathcal{X}$. Moreover, we assume that $\mathbf{B}$ is independent of y.

(iii) $f\in\mathcal{C}_b^{3,0+}\big(\mathbf{R}^d\times\mathbf{R}^n,\mathbf{R}^d\big)$. For each $j=1,\dots,d$, $0\leq|\boldsymbol{k}|\leq 3$, we have $\inf_x \mathsf{H}\big(f_j(x,\cdot)\big) > \frac{2}{\beta}$, $D_{\boldsymbol{k}} f_j(x,\cdot)\in L^2\big(\mathbf{R}^n,N(0,\Sigma)\big)$, and

$$\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}\sup_{x\in B_R}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|(4n-1)^{\frac{|\boldsymbol{\ell}|}{2}}(p-1)^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}\hat\Theta^{\frac{|\boldsymbol{\ell}|}{2}}<\infty \qquad \forall\, R>0 \tag{4.6}$$

where $c^{j,\boldsymbol{k}}$ are the coefficients in the Hermite expansion of $D_{\boldsymbol{k}} f_j$.

(iv) $g\in\mathcal{C}_b^{\gamma^{-1}+}\big(\mathbf{R}^d,\mathbb{L}(\mathcal{X},\mathbf{R}^d)\big)$.

Lemma 4.5 *The condition (4.6) holds if for some $\theta>\hat\Theta$ and every $j=1,\dots,d$, $|\boldsymbol{k}|\leqslant 3$,*

$$\sup_{x\in B_R}\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\,\theta^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}<\infty. \tag{4.7}$$

Proof. Firstly for any $\theta>0, a>1$,

$$\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}\sup_{x\in B_R}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\theta^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}\leq\sup_{x\in B_R}\sup_{\boldsymbol{\ell}\in\mathbf{N}_0^n}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\theta^{\frac{|\boldsymbol{\ell}|}{2}}a^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}a^{-\frac{|\boldsymbol{\ell}|}{2}}.$$

Since the composition number of $|\boldsymbol{\ell}|$ into n blocks is at most of the order $|2\boldsymbol{\ell}|^n$, we have

$$\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}a^{-\frac{|\boldsymbol{\ell}|}{2}}\lesssim\sum_{|\boldsymbol{\ell}|\in\mathbf{N}_0^n}a^{-\frac{|\boldsymbol{\ell}|}{2}}(2|\boldsymbol{\ell}|)^n<\infty.$$

Hence (4.6) follows from $\sup_{x\in B_R}\sup_{\boldsymbol{\ell}\in\mathbf{N}_0^n}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\;\hat\Theta^{\frac{|\boldsymbol{\ell}|}{2}}a^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}<\infty$, the latter clearly follows from the following $\sup_{x\in B_R}\sum_{\boldsymbol{\ell}\in\mathbf{N}_0^n}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\,\theta^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}<\infty$.

$$\sup_{x\in B_R}\sup_{\boldsymbol{\ell}\in\mathbf{N}_0^n}|c^{j,\boldsymbol{k}}_{\boldsymbol{\ell}}(x)|\theta^{\frac{|\boldsymbol{\ell}|}{2}}\sqrt{\boldsymbol{\ell}!}<\infty, \tag{4.8}$$

concluding the proof. □

We have the following limit theorem on the solution to (4.13):

Theorem 4.6 *Consider the RDE (4.13) with Condition 4.4 in place. Then, for each $\varepsilon > 0$, there is a unique pathwise solution $(x_t^\varepsilon)_{t\in[0,T]}$. For $i, j = 1, \dots, d$, let*

$$\sigma_{i,j}(x,z) \stackrel{\text{def}}{=} \int_0^\infty \mathbf{E}[f_i(x,y_r)f_j(z,y_0) + f_i(x,y_0)f_j(z,y_r)]\,dr, \tag{4.9}$$

$$\Lambda_{i,j}(x,z) \stackrel{\text{def}}{=} \int_0^\infty \mathbf{E}[f_i(x,y_0) \otimes f_j(z,y_r)]\,dr \tag{4.10}$$

$$\Gamma_i(x) \stackrel{\text{def}}{=} \int_0^\infty \mathbf{E}[(Df(\cdot,y_r)_x(f(x,y_0))]\,dr. \tag{4.11}$$

Suppose furthermore that that $x_0 \in L^\infty$ is independent of $(y_t)_{t\geqslant 0}$ and $\mathbf{B}$. Then the following hold:

1. *Suppose that $f(\cdot,y)$ and g have common compact support for every $y \in \mathbf{R}^n$. Then there is a limiting rough path $\mathbf{X} \in \mathscr{C}^{\frac{1}{2}^-}\big([0,T], \mathcal{C}_b^3(\mathbf{R}^d,\mathbf{R}^d)\big)$ such that $(\mathbf{X}^\varepsilon, B)$ converges weakly to $(\mathbf{X}, B)$ where the latter $\mathbf{B}$ is a copy of B independent of $\mathbf{X}$. Furthermore $\mathbf{X} = (X, \mathbb{X}_{s,t} + (t-s)\Lambda)$ where X is a Gaussian field with covariance $\sigma_{i,j}(x,z)(t\wedge s)$.*
2. **(Rough case)** *Suppose that $f(\cdot,y)$ and g have common compact support for every $y \in \mathbf{R}^n$. Then for any initial condition x_0, as $\varepsilon \to 0$, the solutions x^ε converge weakly in $\mathcal{C}^{\gamma\wedge(\frac{1}{2})^-}\big([0,T],\mathbf{R}^d\big)$ to the solution of the RDE*
$$dx_t = \delta(x_t)\,d\mathbf{X}_t + g(x_t)\,d\mathbf{B}_t. \tag{4.12}$$
The convergence is simultaneous for any finite number of initial conditions.
3. **(Wiener case)** *Let $\mathbf{B} = (B, \mathbb{B})$ denote the Itô lift of a standard Wiener process and $g_1 \in \mathcal{C}_b^{1+}(\mathbf{R}^d,\mathbf{R}^d)$. Then for any initial solution x_0, the solutions of the equation*
$$dx_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} f\big(x_t^\varepsilon, y_{\frac{t}{\varepsilon}}\big)\,dt + g(x_t^\varepsilon)\,d\mathbf{B}_t + g_1(x_t^\varepsilon)dt, \qquad x_0^\varepsilon = x_0, \qquad t\in[0,T], \tag{4.13}$$
converge weakly in $\mathcal{C}\big([0,T],\mathbf{R}^d\big)$ to the unique solution of the following Kunita type Itô SDE:
$$dx_t = \Gamma(x_t)dt + F(x_t,dt) + g(x_t)dB_t + g_1(x_t)dt.$$
Here $F(x,\cdot) \stackrel{\text{def}}{=} X(x)$ is a martingale with spatial parameters and with characteristics $a_{i,j}(x,z,t) = \sigma_{i,j}(x,z)t$ independent of B. Furthermore if φ^ε denote the solution flow the N-point motion $(\varphi_t^\varepsilon(x_1),\dots,\varphi_t^\varepsilon(x_N))$ converge to the N-point motion of the limiting equation.

Let us record a few remarks on Theorem 4.6:

Remark 4.7

- ▷ Theorem 4.6 extends the results of [GL20a] in several directions: non-product drift form, multi-dimensional environmental noise, and additional noise in the equation.
- ▷ We observe that—unlike the one-dimensional work of the first two authors of this article—the limiting rough path features a non-vanishing Lévy area. The reason for this is the non-reversibility of the Gaussian process $(y_t)_{t\geqslant 0}$ in dimension $d \geqslant 2$. Indeed, a *one-dimensional*, stationary Gaussian process is *always* reversible in the sense that, for each $T > 0$,
$$(y_t)_{t\in[0,T]} \stackrel{d}{=} (y_{T-t})_{t\in[0,T]}.$$
In higher dimensions, stationarity of a Gaussian process is a genuinely weaker requirement than reversibility. We also note that the presence of the non-trivial area term matches the findings of [DOP19].

4.3 Weak Convergence in Rough Path Spaces

This method of lifting the RDE on $\mathbf{R}^d$ to a Banach space of functions has previously been successfully employed in the very nice work of Kelly and Melbourne [KM17]. Since there is however a minor inaccuracy in that article (see Example B) below, we choose to recall an appropriate amount of detail of their approach. Without further notice, we shall assume the regularity assumptions of Condition 4.4 in the sequel. Let us first elaborate on the minor inaccuracy in the work of Kelly and Melbourne, then present results leading tightness of $\mathbf{X}^\varepsilon$ in the rough path spaces over $\mathcal{C}^3_R(\mathbf{R}^d, \mathbf{R}^d)$, the space of smooth functions with compact support in B_R, presented in Proposition 4.9 below which resolves a question raised in [KM17] and allow us to bypass the martingale method problem and obtaining a Kunita type SDE in the limit.

The example below shows that tightness in rough path spaces over infinite-dimensional Banach spaces is a touchy business.

Example B When proving tightness of the driving rough path $\mathbf{X}^\varepsilon = (X^\varepsilon, \mathbb{X}^\varepsilon)$ in [KM17, Corollary 5.9] (there denoted by $\mathbf{W}^\varepsilon$), the authors assert that the unit ball in $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}^\alpha_b(\mathbf{R}^d, \mathbf{R}^d)\big)$ is compact in the metric of $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}^\alpha_b(\mathbf{R}^d, \mathbf{R}^d)\big)$ for each $\gamma < \gamma'$. This, they claim, should follow by a standard Arzelà-Ascoli argument as in [FV10, Chapter 5]. In the latter, however, the authors only consider rough paths with values in a *finite*-dimensional Euclidean space. In fact, for any infinite-dimensional Banach space $\mathcal{X}$, the embedding $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{X}\big) \hookrightarrow \mathscr{C}^{\gamma}\big([0,T], \mathcal{X}\big)$ is not compact. To see this, let us take $T = 1$ and a sequence of points $\{x_n\}_{n\in\mathbf{N}} \subset B_{\mathcal{X}}$, given by Riesz's lemma, with $|x_m - x_n| \geqslant \frac{1}{2}$ for any $m \neq n$. Set $F^n_t \stackrel{\text{def}}{=} t x_n$ and $\mathbb{F}^n_{s,t} \stackrel{\text{def}}{=} \frac{(t-s)^2}{2}(x_n \otimes x_n)$. Then $\big\{\mathbf{F}_n = (F^n, \mathbb{F}^n)\big\}_{n\in\mathbf{N}}$ is bounded in $\mathscr{C}^\gamma\big([0,T], \mathcal{X}\big)$, but not compact, as $|\mathbf{F}_m - \mathbf{F}_n|_{\mathcal{C}^\gamma \oplus \mathcal{C}^{2\gamma}} \geqslant \frac{1}{2}$ for all $m \neq n$.

A remedy of this lack of compactness was presented in the recent preprint [CFK+20]. There, the authors worked in the p-variation setting, but the arguments of course transfer to Hölder rough paths. We recapitulate a streamlined version of the argument in the sequel; mainly for the reader's convenience, also to fix some notations.

Let $R > 0$. We let $\mathcal{C}^\alpha_R$ denote the Hölder functions supported in the ball of radius R,

$$\mathcal{C}^\alpha_R(\mathbf{R}^d, \mathbf{R}^d) \stackrel{\text{def}}{=} \big\{f \in \mathcal{X} : \operatorname{supp}(f) \subset B_R\big\}, \qquad \mathcal{X} = \mathcal{C}^\alpha_b(\mathbf{R}^d, \mathbf{R}^d)$$

Definition 4.8 Let $\mathbf{X} = (X, \mathbb{X})$ and $\mathbf{Y} = (Y, \mathbb{Y})$ be random variables with values in $\mathscr{C}^\gamma\big([0,T], \mathcal{X}\big)$. We say that $\mathbf{X}$ and $\mathbf{Y}$ are equal in *finite-dimensional (space-time) distributions* if, for each $n \in \mathbf{N}$, $0 \leqslant t_1 < \cdots < t_m \leqslant T$, and $x^i_1, \dots, x^i_n \in \mathbf{R}^d$ $(i = 1, 2)$,

$$\begin{pmatrix} \big(X_{t_1}(x^1_1), \mathbb{X}_{t_1}(x^1_1, x^2_1)\big) \\ \vdots \\ \big(X_{t_n}(x^1_n), \mathbb{X}_{t_1}(x^1_n, x^2_n)\big) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \big(Y_{t_1}(x^1_1), \mathbb{Y}_{t_1}(x^1_1, x^2_1)\big) \\ \vdots \\ \big(Y_{t_n}(x^1_n), \mathbb{Y}_{t_1}(x^1_n, x^2_n)\big) \end{pmatrix}$$

as $\Big(\mathbf{R}^d \oplus \big(\mathbf{R}^d \otimes \mathbf{R}^d\big)\Big)^n$-valued random variables.

Our main weak convergence result of the section is as follows:

Proposition 4.9 *Let $\gamma \in (0, 1]$, $\alpha > 0$, and $R > 0$. Let $(\mathbf{X}^n)_{n\in\mathbf{N}}$ be a sequence of random variables with values in $\mathscr{C}^\gamma\big([0,T], \mathcal{C}^\alpha_R(\mathbf{R}^d, \mathbf{R}^d)\big)$. If there is a $p \geqslant 1$ such that*

$$\sup_{n\in\mathbf{N}} \mathbf{E}\Big[|\mathbf{X}^n|^p_{\mathcal{C}^\gamma \oplus \mathcal{C}^{2\gamma}}\Big] < \infty$$

and if the finite-dimensional space-time distributions of any weak limit point of $(\mathbf{X}^n)_{n\in\mathbf{N}}$ coincide, then there is a random variable $\mathbf{X} \in \mathscr{C}^{\gamma-}\big([0,T], \mathcal{C}^{\alpha-}_R(\mathbf{R}^d, \mathbf{R}^d)\big)$ such that $\mathbf{X}^n \Rightarrow \mathbf{X}$ weakly.

Remark 4.10 Proposition 4.9 resolves the question raised in [KM17, Remark 5.14]. In fact, by employing it, one could bypass the invocation of the martingale problem used in [KM17, Section 6] in order to characterize the limiting equation.

Proof of Proposition 4.9. Fix $\gamma' < \gamma$ and $\alpha' < \alpha$. Let $\varepsilon > 0$. Then $(\mathbf{X}^n)_{n\in\mathbf{N}}$ is tight in the space $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$. Indeed, there is an $M > 0$ such that $\mathbf{P}\big(\mathbf{X}^n \notin K_M\big) \leqslant \varepsilon$ where

$$K_M \stackrel{\text{def}}{=} \big\{\mathbf{X} \in \mathscr{C}^{\gamma}\big([0,T], \mathcal{C}_R^{\alpha}(\mathbf{R}^d,\mathbf{R}^d)\big) : |\mathbf{X}|_{\mathcal{C}^\gamma\oplus\mathcal{C}^{2\gamma}} \leqslant M\big\}.$$

By Lemma 4.11 below this is relatively compact as a subset of $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$.

It remains to show that the finite-dimensional space-time distributions uniquely characterize the limit points of $(\mathbf{X}^n)_{n\in\mathbf{N}}$. Let $\mathbf{X}^1$ and $\mathbf{X}^2$ be limit points of $(\mathbf{X}^n)_{n\in\mathbf{N}}$. By the Portmanteau theorem, the laws of both $\mathbf{X}^1$ and $\mathbf{X}^2$ are Radon measures. A compactification argument (see e.g. [Bog07, Exercise 7.14.79]) shows that these two measures coincide, provided we can exhibit a test set $\mathcal{F}$ of bounded continuous functions $f : \mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big) \to \mathbf{C}$ such that

- $\mathcal{F}$ separates points on $\mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$,
- $\mathcal{F}$ is closed under pointwise multiplication,
- $\mathbf{1} \in \mathcal{F}$, where $\mathbf{1}$ is the function with constant value 1,
- $\mathbf{E}\big[f(\mathbf{X}^1)\big] = \mathbf{E}\big[f(\mathbf{X}^2)\big]$ for all $f \in \mathcal{F}$.

We choose $\mathcal{F}$ as the family of characteristic functionals furnished by the space-time evaluations of Definition 4.8:

$$\mathcal{F} \stackrel{\text{def}}{=} \Bigg\{(X,\mathbb{X}) \mapsto \exp\left(i\sum_{j=1}^n \pi_j\Big(X_{t_j}(x_j^1), \mathbb{X}_{t_j}(x_j^1,x_j^2)\Big)\right) : \\ n \in \mathbf{N},\, t_j \in [0,T],\, x_j^{1,2} \in \mathbf{R}^d,\, \pi_j \in \mathbb{L}\Big(\mathbf{R}^d \oplus \big(\mathbf{R}^d \otimes \mathbf{R}^d\big), \mathbf{R}\Big)\Bigg\}.$$

One verifies that $\mathcal{F}$ satisfies the requirements above, whence $\mathbf{X}^1 \stackrel{d}{=} \mathbf{X}^2$, as required. □

Lemma 4.11 *Let $R > 0$, $0 < \gamma' < \gamma \leqslant \frac{1}{2}$, and $0 < \alpha' < \alpha$. Then the embedding*

$$\mathscr{C}^{\gamma}\big([0,T], \mathcal{C}_R^{\alpha}(\mathbf{R}^d,\mathbf{R}^d)\big) \hookrightarrow \mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big) \tag{4.14}$$

is compact.

Proof. The embedding (4.14) is certainly continuous. To see that it is actually compact, it is enough to show that the set $K_1 \stackrel{\text{def}}{=} \big\{\mathbf{X} \in \mathscr{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big) : |\mathbf{X}|_{\mathcal{C}^\gamma\oplus\mathcal{C}^{2\gamma}} \leqslant 1\big\}$ is relatively compact.

We shall make use of a general version of the Arzelà-Ascoli theorem recalled in Proposition 4.12 after the proof. Let $(\mathbf{X}^n)_{n\in\mathbf{N}}$ be a sequence in K_1. We need to show that both $(X^n)_{n\in\mathbf{N}}$ and $(\mathbb{X}^n)_{n\in\mathbf{N}}$ are relatively compact in the spaces $\mathcal{C}^{\gamma'}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$ and $\mathcal{C}^{2\gamma'}\big(\Delta_T, \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d) \otimes \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$, respectively. Since the arguments are similar, we only detail the relative compactness of $(\mathbb{X}^n)_{n\in\mathbf{N}}$.

By the algebraic constraint (2.12) and a straight-forward interpolation estimate (see e.g. [FH14, Exercise 2.9]), it is enough to show that $(\mathbb{X}^n_{0,\cdot})_{n\in\mathbf{N}}$ is relatively compact in $\mathcal{C}\big([0,T], \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d) \otimes \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)\big)$. First note that this family is equicontinuous. Indeed, again by Chen's relation,

$$\|\mathbb{X}^n_{0,t} - \mathbb{X}^n_{0,s}\|_{\mathcal{C}^{\alpha'}\otimes\mathcal{C}^{\alpha'}} \leqslant \|\mathbb{X}^n_{s,t}\|_{\mathcal{C}^{\alpha'}\otimes\mathcal{C}^{\alpha'}} + |X^n_s|_{\mathcal{C}^{\alpha'}}|X^n_{s,t}|_{\mathcal{C}^{\alpha'}}$$

for all $0 \leqslant s \leqslant t \leqslant T$. We are thus left to show that, for each $t \in [0,T]$, the set $\big\{\mathbb{X}^n_{0,t} : n \in \mathbf{N}\big\}$ is relatively compact in $\mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d) \otimes \mathcal{C}_R^{\alpha'}(\mathbf{R}^d,\mathbf{R}^d)$ in order to conclude with Proposition 4.12 below. For this it is enough to note the compact embedding

$$\mathcal{C}^{\alpha}\big(\overline{B}_R \times \overline{B}_R, \mathbf{R}^d \otimes \mathbf{R}^d\big) \hookrightarrow \mathcal{C}^{\alpha'}\big(\overline{B}_R \times \overline{B}_R, \mathbf{R}^d \otimes \mathbf{R}^d\big),$$

which in turn again follows from the Arzelà-Ascoli theorem—but this time in the spatial coordinate. □

The following version of the Arzelà-Ascoli, employed in the previous proof, can be found in multiple places, see e.g. [Kel75, Theorem 7.17]:

Proposition 4.12 (Arzelà-Ascoli) *Let X be a compact metric space and Y be a metric space. Let $\mathcal{C}(X,Y)$ be the space of continuous mappings $f : X \to Y$, equipped with the uniform topology. Then a set $K \subset \mathcal{C}(X,Y)$ is relatively compact if and only if*

- ▷ *the set $\{f(x) : f \in K\}$ is relatively compact in Y for each $x \in X$ and*
- ▷ *the set K is equicontinuous, i.e., for all $\varepsilon > 0$, there is a $\delta > 0$ such that $d_Y\big(f(x), f(y)\big) \leqslant \varepsilon$ for all $f \in K$, provided that $d_X(x,y) \leqslant \delta$.*

4.4 Proof of Theorem 4.6

We can now conclude the proof our homogenization result:

Proof of Theorem 4.6. We first assume that $f(\cdot, y)$ and g are supported in $\overline{B}_R$ for each $y \in \mathbf{R}^n$. Owing to Proposition 2.9, it is enough to show that

$$(\mathbf{X}^\varepsilon, \mathbf{B}) \Rightarrow (\mathbf{X}, \mathbf{B}) \tag{4.15}$$

in $\mathscr{C}^\gamma\big([0,T], \mathcal{C}^3_R(\mathbf{R}^d, \mathbf{R}^d)\big) \times \mathscr{C}^\gamma\big([0,T], \mathcal{X}\big)$. By independence it is sufficient to show the weak convergence of the first marginal for which we wish to employ Proposition 4.9. To see that

$$\sup_{\varepsilon \in (0,1]} \mathbf{E}\big[|X^\varepsilon|^2_{\mathcal{C}^\gamma}\big] < \infty,$$

we argue as in Lemma 3.9: For any $i = 1, \dots, d$

$$\begin{aligned}\mathbf{E}\left[\sup_{x \in \overline{B}_R} \left| \frac{1}{\sqrt{\varepsilon}} \int_s^t f_i(x, y_{\frac{u}{\varepsilon}})\, du \right|^2\right] &\leqslant \frac{1}{\varepsilon} \sum_{\boldsymbol{k}, \boldsymbol{\ell} \in \mathbf{N}_0^n} \sup_{x \in \overline{B}_R} c^i_{\boldsymbol{k}}(x) c^i_{\boldsymbol{\ell}}(x) \int_s^t \int_s^t \mathbf{E}\big[H_{\boldsymbol{k}}(y_{\frac{u}{\varepsilon}}) H_{\boldsymbol{m}}(y_{\frac{v}{\varepsilon}})\big]\, du\, dv \\ &\lesssim |t-s| \sum_{\boldsymbol{\ell} \in \mathbf{N}_0^n} \sup_{x \in \overline{B}_R} c^i_{\boldsymbol{\ell}}(x)^2 \boldsymbol{\ell}! \hat{\Theta}^{|\boldsymbol{\ell}|} (2n-1)^{|\boldsymbol{\ell}|}.\end{aligned}$$

We exchanged summation with the supremum, and then the expectation with the supremum, valid by condition (iii). Here, we used that $\sup_x c_{\boldsymbol{\ell}}(x) = 0$ whenever $|\boldsymbol{\ell}| < \inf_x \mathsf{H}\big(f(x,\cdot)\big)$. In view of Lemma 4.3, higher order derivatives are similar. Arguing as in Lemma 3.18 also shows that $\sup_\varepsilon \mathbf{E}[|\mathbb{X}^\varepsilon|_{\mathcal{C}^{2\gamma}}] < \infty$. Then Theorem 3.1 and Remark 3.2 show that: $\mathbf{X}^\varepsilon$ converges to $\mathbf{X}$ with $\mathbb{X}_{s,t}(x,z) = (\Upsilon \otimes \Upsilon)\mathbb{W}^{\text{Itô}}_{s,t}(x,z) + \Lambda(x,z)(t-s)$ and $\mathbb{W}_{s,t}(x)$ is a Wiener process and the convergence of the solutions follows. Since the drivers converge, they do not know of the initial conditions, the convergence is simultaneous for any finite number fo initial conditions.

It remains to prove statement 3. We first study the proposed limit equation. Let B be a Brownian motion independent of $\mathbf{X}_t$. Observe that X is a Gaussian process with covariance $\sigma_{i,j}(x,z)(s \wedge t)$ and

$$G(x,t) \stackrel{\text{def}}{=} X_t(x) + g(x) B_t + \Gamma t$$

is a semi-martingale with spatial parameters and with the characteristics

$$A_{i,j}(x,z,t) \stackrel{\text{def}}{=} \sigma_{i,j}(x,z)t + g(x)g^T(z)t, \qquad \text{and} \qquad \Gamma(x)t.$$

The characteristics are sufficiently regular so the the Kunita type equation (with Itô integration)

$$dx_t = F(x_t, dt) + \Gamma(x_t) + g(x_t) dB_t \tag{4.16}$$

is well posed. The regularity of $A_{i,j}$ and Γ comes from the uniform correlation decay assumption $\alpha = \inf_x \mathsf{H}\big(f(x,\cdot) > \frac{2}{\beta}$. For example, the functions $\mathbf{E}[D(f_i(x,y_s)f_j(z,y_0) + f_i(x,y_0)f_j(z,y_s))]$ are absolutely integrable in s on $[0,\infty)$, consequently

$$\sigma_{i,j}(x,z) = \int_0^\infty \mathbf{E}[f_i(x,y_r)f_j(z,y_0) + f_i(x,y_0)f_j(z,y_r)]\, dr$$

is $\mathcal{C}_b^3$ in both variables and jointly continuous in (x,z). The same argument applies to Γ.

Furthermore, by the SDE theory drive by semi-martingale with spatial parameters, [Kun97], there is a unique Brownian flow $\varphi_{s,t}(x)$ to the equation (4.16) and for each s,x,

$$\varphi_{s,t}(x) - x - \int_s^t \Gamma(\varphi_{s,r})dr$$

is a square integrable martingale with $\lim_{h\to 0}\mathbf{E}[(\varphi_{t,t+h}(x)-x)(\varphi_{t,t+h}(y)-y)^T] = \sigma_{i,j}(x,y)t$. The RDE is 'equivalent' to the Kunita type SDE which can be seen from the Riemann sum approximation for integration with respect to the semi-martingale G :

$$\int_0^t G(x_s, ds) \sim \sum_{[u,v]\subset\mathcal{P}} G(x_u,u) - G(x_v,v)$$

where $\mathcal{P}$ denotes a partition of $[0,t]$. On the other hand, the rough integral,

$$\int_0^t \delta(x_s)d\mathbf{X}_s \sim \sum_{[u,v]\subset\mathcal{P}} \delta(x_u)(X_u - X_v) + \delta(x_u)'(\mathbb{X}^{\text{Itô}}_{u,v} + \Lambda(u-v)),$$

where the prime on $\delta(x_t)$ denotes its Gubinelli derivative, which is $(D\delta)_{x_u}(\delta(x_u)\cdot)(\cdot)$ which applied to $\Lambda(\cdot,\cdot)$ is:

$$\Gamma_j(x) = \int_0^\infty \mathbf{E}[(Df_j(\cdot,y_0))_x(f(x,y_r))]\,dr.$$

It is then routine to verify that $\varphi_t(x)$ is the solution to the rough differential equation

$$dx_t = \delta(x_t)d\mathbf{X}_t + \Lambda(x_t)dt + g(x_t)d\mathbf{B}_t. \tag{4.17}$$

The latter equation is also well posed.

We then fix $R>0$ let $\eta_R : \mathbf{R}^d \to \mathbf{R}_+$ be a non-negative smooth function with $\eta_R = 1$ on $\overline{B}_R$ and $\eta_R = 0$ on B_{2R}^c. Set $f_R(x,y) \overset{\text{def}}{=} \eta_R(x)f(x,y)$ and similar for g_R. By the first part of the theorem, we know that there is an $\mathbf{X}^R = (X^R, \mathbb{X}^R) \in \mathscr{C}\big([0,T], \mathcal{C}_R^3(\mathbf{R}^d,\mathbf{R}^d)\big)$ such that $x^{R,\varepsilon} \Rightarrow x^R$, where

$$\begin{aligned} dx_t^{R,\varepsilon} &= \frac{1}{\sqrt{\varepsilon}} f_R(x_t^{R,\varepsilon}, y_t)\,dt + g_R(x_t^{R,\varepsilon})\,d\mathbf{B}_t + g_{1,R}(x_t^\varepsilon)\,dt, \\ dx_t^R &= \delta(x_t^R)\,d\mathbf{X}_t^R + g_R(x_t^R)\,d\mathbf{B}_t + g_{1,R}(x_t^\varepsilon)\,dt. \end{aligned}$$

As before,the convergence is simultaneous for any finite number fo initial conditions. Mimicking that for G, let $G_R = \mathbf{X}^R + g_R B + (\Gamma_R + g_{1,R})t$ denote the spatial semi-martingale with characteristics $\sigma_{i,j}^R(x,y) + g_R(x)g(y)^T$ and $\Gamma_R + g_{1,R}\,dt$. Then for each initial condition x_0, the solutions of the Kunita type SDEs

$$dx_t^R = G_R(x_t^R, dt)\,dt$$

converge weakly to the solution of (4.17) with the same initial distribution. We have applied [Kun97, Thm. 5.2.1], it is trivial to verify the conditions there: $\sup_x |D^a D^b(\sigma_R(x,y) + g_R(x)g_R^T(y))|_{|_{y=x}}$ where $a,b\in\{0,1\}$ and $\sup_x |D_x\Gamma_R + g_{1,R}|$ are uniformly bounded, also the characteristics are in the required class $\tilde{C}_b^\alpha$ for some $\alpha>0$, c.f. [Kun97], and converging uniformly in (x,y) on compact sets. The convergence is in the sense of diffusions in the sense of given any initial data (x_0^i), $i=1,\dots,N$, the N-point motion converge.

Finally, we show the weak convergence theorem of $x^\varepsilon \to x$. By the Portemanteau theorem, it is equivalent to showing that $\operatorname{limsup}_{\varepsilon\to 0}\mathbf{P}(x^\varepsilon\in A) \leqslant \mathbf{P}(X\in A)$ for any closed set $A\subset\mathcal{C}\big([0,T],\mathbf{R}^d\big)$. Since $X^{R,\varepsilon}\Rightarrow X^R$ weakly in $\mathcal{C}\big([0,T],\mathbf{R}^d\big)$, Portemanteau theorem gives the estimate:

$$\begin{aligned} \operatorname*{limsup}_{\varepsilon\to 0}\mathbf{P}(x^\varepsilon\in A) &\leqslant \operatorname*{limsup}_{\varepsilon\to 0}\mathbf{P}(x^{R,\varepsilon}\in A) + \operatorname*{limsup}_{\varepsilon\to 0}\mathbf{P}\big(|x^{R,\varepsilon}|_\infty > R\big) \\ &\leqslant \mathbf{P}(X^R\in A) + \mathbf{P}(|X^R|_\infty \geqslant R). \end{aligned} \tag{4.18}$$

Hence, $\limsup_{R\to\infty} \mathbf{P}(X^R \in A) \leqslant \mathbf{P}(X \in A)$. Note also that $\mathbf{P}(|X^R|_\infty \geqslant R) = \mathbf{P}(|X|_\infty \geqslant R)$ for $R > \|x_0\|_{L^\infty}$. Markov's inequality gives that $\mathbf{P}(|X|_\infty \geqslant R) \to 0$ as $R \to \infty$. Consequently, sending $R \to \infty$ in (4.18), we have proven that $\limsup_{\varepsilon\to 0} \mathbf{P}(x^\varepsilon \in A) \leqslant \mathbf{P}(X \in A)$ which concludes the proof for the weak convergence of $x^\varepsilon \to x$. The proof for the convergence of the N-point motion is an easy adaption of the above, as we have the N-point convergence for both $x^{\varepsilon,R}$ as $\varepsilon \to 0$ and for x^R as $R \to \infty$. □

References

[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.

[Bog07] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[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+20] Ilya Chevyrev, Peter K. Friz, Alexey Korepanov, Ian Melbourne, and Huilin Zhang. Deterministic homogenization under optimal moment assumptions for fast-slow systems. part 2, 2020.

[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 nonstationarity. In Arxiv: arXiv:1807.09740, 2018.

[CR96] F. Comte and E. Renault. Long memory continuous time models. *J. Econometrics*, 73(1):101–149, 1996.

[DOP19] Jean-Dominique Deuschel, Tal Orenshtein, and Nicolas Perkowski. Additive functionals as rough paths, 2019.

[FdLP06] Denis Feyel and Arnaud de La Pradelle. Curvilinear integrals along enriched paths. *Electron. J. Probab.*, 11:no. 34, 860–892, 2006.

[FH14] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[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.

[FW12] Mark I. Freidlin and Alexander D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, third edition, 2012. Translated from the 1979 Russian original by Joseph Szücs.

[Geh20] Johann Gehringer. Functional limit theorems of moving averages of hermite processes and an application to homogenization, 2020.

[GL19] J. Gehringer and Xue-Mei Li. Homogenization with fractional random fields. arXiv:1911.12600. This is now improved and split into 'Functional limit theorem for fractional OU' and 'Diffusive and rough homogenisation in fractional noise field", 2019.

[GL20a] J. Gehringer and Xue-Mei Li. Diffusive and rough homogenisation in fractional noise field. Part 2 of an improved version of arXiv:1911.12600, 2020.

[GL20b] Johann Gehringer and Xue-Mei Li. Functional limit theorems for the fractional ornstein-uhlenbeck process. *J. Theoretical Probability*, 2020. https://doi.org/10.1007/s10959-020-01044-7, in Arxiv as part of arXiv:1911.12600.

[Gub04] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.

[ILRMS13] Alexander V. Ivanov, Nikolai Leonenko, María D. Ruiz-Medina, and Irina N. Savich. Limit theorems for weighted nonlinear transformations of Gaussian stationary processes with singular spectra. *Ann. Probab.*, 41(2):1088–1114, 2013.

[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.

[Kal21] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, third edition, 2021.

[Kel75] John L. Kelley. *General topology*. Springer-Verlag, New York-Berlin, 1975. Reprint of the 1955 edition [Van Nostrand, Toronto, Ont.], Graduate Texts in Mathematics, No. 27.

[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.

[KP78] H. Kesten and G. C. Papanicolaou. A limit theorem for turbulent diffusion. In *Mathematical problems in theoretical physics (Proc. Internat. Conf., Univ. Rome, Rome, 1977)*, volume 80 of *Lecture Notes in Phys.*, pages 371–374. Springer, Berlin-New York, 1978.

[KP91] Thomas G. Kurtz and Philip Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.

[Kun97] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. Reprint of the 1990 original.

[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.

[LC85] W. A. Light and E. W. Cheney. *Approximation theory in tensor product spaces*, volume 1169 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.

[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.

[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.

[NOL08] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.*, 118(4):614–628, 2008.

[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.

[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.

[PT05] Giovanni Peccati and Ciprian A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 247–262. Springer, Berlin, 2005.

[Rio17] Emmanuel Rio. *Asymptotic theory of weakly dependent random processes*, volume 80 of *Probability Theory and Stochastic Modelling*. Springer, Berlin, 2017. Translated from the 2000 French edition [MR2117923].

[Rob95] P. M. Robinson. Gaussian semiparametric estimation of long range dependence. *Ann. Statist.*, 23(5):1630–1661, 1995.

[Rya02] Raymond A. Ryan. *Introduction to tensor products of Banach spaces*. Springer Monographs in Mathematics. Springer-Verlag London, Ltd., London, 2002.

[Sam06] Gennady Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1(3):163–257, 2006.

[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.

[You36] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1):251–282, 1936.