CONSTRUCTION OF Φ^4_3 DIAGRAMS FOR PEDESTRIANS

JEAN-CHRISTOPHE MOURRAT, HENDRIK WEBER, AND WEIJUN XU

ABSTRACT. We aim to give a pedagogic and essentially self-contained presentation of the construction of various stochastic objects appearing in the dynamical Φ^4_3 model. The construction presented here is based on the use of paraproducts. The emphasis is on describing the stochastic objects themselves rather than introducing a solution theory for the equation.

1. INTRODUCTION

The purpose of this note is to give a pedagogic presentation of the construction of the various stochastic "basis" terms entering the construction of the dynamic Φ^4 theory in three space dimensions (Φ^4_3 for short). Formally, the dynamic Φ^4 model on the torus $\mathbb{T}^3 = [0,1]^3$ is the solution $X(t,x)$ to the stochastic partial differential equation

$$\begin{cases} \partial_t X = \Delta X - X^3 + mX + \xi, & \text{on } \mathbb{R}_+ \times [0,1]^3, \\ X(0,\cdot) = X_0, \end{cases} \tag{1.1}$$

where ξ denotes a white noise over $\mathbb{R} \times \mathbb{T}^3$, and m is a real parameter. Equation (1.1) describes the natural reversible dynamics for the "static" Φ^4_3 Euclidean field theory, which is formally given by the expression

$$\mu \propto \exp\left(-2 \int_{\mathbb{T}^3} \left[\frac{1}{2}|\nabla X|^2 + \frac{1}{4}X^4 - \frac{m}{2}X^2\right]\right) \prod_{x \in \mathbb{T}^3} \mathrm{d}X(x). \tag{1.2}$$

The quartic potential X^4 in this energy gives the model its name (replacing X by Φ). Mathematically, neither (1.1) nor (1.2) make sense as they stand. While this problem is the main concern of this note, we postpone its discussion and first proceed heuristically. (Alternatively, we temporarily replace the continuous space $\mathbb{T}^3$ by a finite approximation, with a suitable interpretation of the gradient.)

The potential function $x \mapsto \frac{1}{4}x^4 - \frac{m}{2}x^2$ has a single minimum at $x = 0$ for $m \le 0$. As m becomes positive, a pitchfork bifurcation occurs, with the appearance of two minima at $x = \pm\sqrt{m}$, while the point $x = 0$ becomes a local maximum. In the energy between square brackets appearing in (1.2), the part consisting of

$$\int_{\mathbb{T}^3} \left[\frac{1}{4}X^4 - \frac{m}{2}X^2\right]$$

favors fields X that take values close to those minima, while the part

$$\int_{\mathbb{T}^3} |\nabla X|^2$$

2010 *Mathematics Subject Classification.* 81T18, 81T08, 60H15.

Key words and phrases. Feynman diagrams, Singular stochastic PDEs, stochastic quantisation, paraproducts.

favors some agreement between nearby values of the field X. This description is highly reminiscent of that of the Ising model. Indeed, these two models are believed to have comparable phase transitions and large-scale properties. In one and two space dimensions (when $\mathbb{T}^3$ is replaced by $\mathbb{T}^d$, $d \in \{1, 2\}$), a precise link between the Glauber dynamics of an Ising model with long-range interactions and the dynamic Φ^4 model was obtained [2, 10, 22], and a similar result is conjectured to hold in our present three-dimensional setting.

Starting from the 60's, the Φ^4 model was the subject of an intense research effort from the perspective of quantum field theory. From this point of view, the construction of the so-called Euclidean Φ^4 measure (1.2) is a first step towards building the corresponding quantum field theory. This requires the verification of certain properties known as the Osterwalder-Schrader axioms [24, 25], among which the reflection positivity and the invariance under Euclidean isometry are the most important (we refer to [3] for a review on reflection positivity — in particular, the Ising measure is reflection positive, see [3, Corollary 5.4]). The whole endeavour was viewed as a test-bed for more complicated (and more physically relevant) quantum field theories. We stress that from this point of view, one of the directions of space becomes the time variable in the quantum field theory. The time variable appearing in (1.1) is then seen as an additional, physically fictitious variable, sometimes called the "stochastic time" in the literature. The construction of a quantum field theory based on the invariant measure of a random process is called "stochastic quantisation", and was proposed by Parisi and Wu [26]. We refer to [12, Section 20.1] for references and more precise explanations.

We now return to the problem that (1.1) and (1.2) do not actually make sense mathematically. In (1.1), the roughness of the noise requires X to be distribution-valued, and therefore the interpretation of the cubic power is unclear. In (1.2), one could interpret

$$\exp\Big(-\int_{\mathbb{T}^3} |\nabla X|^2\Big) \prod_{x \in \mathbb{T}^3} \mathrm{d}X(x)$$

as a formal notation to denote the law of a Gaussian free field. Again, the Gaussian free field is distribution-valued, and there is no canonical interpretation for X^4 or X^2.

From now on, we focus on making sense of (1.1). A naive attempt would consist of regularising the noise, defining the corresponding solution, and trying to pass to the limit. However, the progressive blow-up of the non-linearity forces the limit to be identically zero (see [17] for a rigorous justification in the case of two space dimensions). Thus, we need to take a step back and modify the original equation (1.1) in a way that is faithful to the intended "physics" of the model, as sketched above.

A formal scaling argument (see e.g. [7, Section 1.1]) shows that the non-linearity should become less and less relevant as we zoom in on the solution: the equation is said to be *subcritical*, or super-renormalisable. The basic idea for making sense of the equation is therefore to postulate a first-order expansion of X of the form $X = \varpi + Y$, where ϖ is the stationary solution to the linear equation

$$(\partial_t - \Delta + 1)\varpi = \xi. \tag{1.3}$$

In other words, letting $\{P_t = e^{t(\Delta-1)}\}_{t\geq 0}$ denote the heat semigroup on $\mathbb{T}^3$, we have

$$\text{•}(t) = \int_{-\infty}^{t} P_{t-s}(\xi(s))\,\mathrm{d}s. \tag{1.4}$$

The "+1" in (1.3) serves to prevent the divergence of the low-frequency part of • in the long-time limit (and to allow us to talk about a stationary solution over the whole time line $\mathbb{R}$).

Making the ansatz $X = \text{•} + Y$ and formally rewriting (1.1) in terms of Y leads to the equation

$$\partial_t Y = \Delta Y - (Y + \text{•})^3 + m(Y + \text{•}). \tag{1.5}$$

Solving this equation for Y requires us to make sense of quantities such as $(\text{•})^2$ or $(\text{•})^3$.

These are again ill-defined. We may regularise the noise, on scale $1/n$, and define the corresponding solution •_n. While $(\text{•}_n)^2$ and $(\text{•}_n)^3$ still diverge as we remove the regularisation, the very explicit and simple structure of • allows us now to identify a constant $\mathbf{c}_n$ such that $(\text{•}_n)^2 - \mathbf{c}_n$ and $(\text{•}_n)^3 - 3\mathbf{c}_n\,\text{•}_n$ converge to non-trivial limits as n tends to infinity, which we denote by  and respectively.

At this point, we can rewrite the equation (1.5) for $Y = X - \text{•}$ using •, and . In two space dimensions, this equation has been solved in [8] with classical methods, without further recourse to the probabilistic structure of the problem. Note that this approach shares the philosophy of rough path theory (see e.g. [11] for an introduction), in that one first constructs a few fundamental objects (here •, and) by relying on the probabilistic structure of the problem, and then one builds the solution as a deterministic and continuous map of the enriched datum (•, ,).

In three space dimensions, the equation one obtains for the remainder Y is still ill-defined, and we need to proceed further in the postulated "Taylor expansion" of the solution X. The procedure becomes more intricate, and was solved only recently by Hairer [15] within his ground-breaking theory of *regularity structures* (see also [14] for the treatment of the KPZ equation with rough paths). Catellier and Chouk [5] then showed how to recover the results of Hairer for the Φ^4_3 model, using the alternative theory of *para-controlled distributions* set up by Gubinelli, Imkeller and Perkowski [13]. We refer to [21, Section 1] for a presentation of the latter approach with notation consistent with the one we use here. Yet another approach based on Wilsonian *renormalisation group* analysis was given by Kupiainen in [19].

We work here in the para-controlled framework, as in [13, 5, 21]. As it turns out, six "basis" elements, that is, non-linear objects based on the solution to the linear equation (1.3) and built using the probabilistic structure, are required to define a solution to the Φ^4_3 equation. We call these processes "the diagrams". They are listed in Table 1.1; their precise meaning will be explained shortly. The purpose of this note is to review their construction.

Minor variants of these diagrams were built in [15] in the context of regularity structures. There, a very convenient graphical notation akin to Feynman diagrams was introduced to derive the bounds required for the construction of these diagrams (see [27] for an elementary introduction to Feynman diagrams. An earlier version of a graph-based method to bound stochastic terms using diagrams in the context

τ						
$\lvert\tau\rvert$	$-\frac{1}{2}-\varepsilon$	$-1-\varepsilon$	$\frac{1}{2}-\varepsilon$	$-\varepsilon$	$-\frac{1}{2}-\varepsilon$	$-\varepsilon$

TABLE 1.1. The list of relevant diagrams, together with their regularity exponent, where $\varepsilon > 0$ is arbitrary.

of the KPZ equation was developed in [14]). In the context of para-controlled distributions, the exact same diagrams as those we consider here were also built in [5], albeit with a possibly less transparent notation. More recently, a remarkable machinery was developped in the context of regularity structures, which ensures the convergence of diagrams for a large class of models under extremely general assumptions; see [16, Theorem A.3] and [6].

This note is mostly expository: we aim to provide a gentle introduction to this graphical notation, and to make clear that it applies equally well in the para-controlled setting. We do not strive to capture the deep results in [16, 6], but rather to give a "pedestrian" exposition of the calculations involved.

We now introduce some notation. Let $\mathcal{S}'$ denote the space of Schwartz distributions over the torus $\mathbb{T}^3$, and define

$$I(f) : \begin{cases} \mathbb{R} & \to & \mathcal{S}' \\ t & \mapsto & \int_{-\infty}^{t} P_{t-s}(f(s))\,\mathrm{d}s, \end{cases} \tag{1.6}$$

for every $f \in C(\mathbb{R}, \mathcal{S}')$ for which the integral makes sense. In other words, $I(f)$ is the "ancient" solution to the heat equation with right-hand side f, that is, the one "started at time $-\infty$". We measure the regularity of distributions on $\mathbb{T}^3$ via a scale of function spaces which we denote by $\mathcal{C}^\alpha$, $\alpha \in \mathbb{R}$. The precise definition is recalled below, and is a natural extension of the notion of α-Hölder regular functions. We also recall below the definition of the resonant product $\varolessthan$. Our goal is to identify suitable deterministic constants $\mathbf{c}_n, \mathbf{c}'_n \in \mathbb{R}$ and show the convergence as n tends to infinity of the following five processes:

$$\begin{aligned} &_n := (_n)^2 - \mathbf{c}_n, \\ &_n := I\left((_n)^3 - 3\mathbf{c}_n _n\right), \\ &_n := _n \varolessthan _n, \\ &_n := I\left(_n\right) \varolessthan _n - 2\mathbf{c}'_n, \\ &_n := _n \varolessthan _n - 6\mathbf{c}'_n _n. \end{aligned} \tag{1.7}$$

The interested reader is referred to the discussion in [21, Section 1.1] to see how these diagrams arise in the construction of solutions to (1.1). The stationarity in space and time of the white noise ξ as well as the fact that I defines ancient solutions to the inhomogeneous heat equation imply that these processes are stationary in space and time. Here is the main result on which we will focus.

Theorem 1.1 ([15, 5]). *Fix*

$$\mathbf{c}_n := \mathbb{E}\left[(_n(t))^2\right] \quad \textit{and} \quad \mathbf{c}'_n := \mathbb{E}\left[I\left(_n\right) \varolessthan _n(t)\right]. \tag{1.8}$$

For each pair $(\tau, |\tau|)$ as in Table 1.1, let τ_n be defined as in (1.7). There exists a stochastic process, denoted by τ and taking values in $C(\mathbb{R}, \mathcal{C}^{|\tau|})$, such that for every $p \in [1, +\infty)$, we have

$$\sup_{t\in\mathbb{R}} \mathbb{E}\left[\|\tau_n(t) - \tau(t)\|^p_{\mathcal{C}^{|\tau|}}\right] \xrightarrow[n\to+\infty]{} 0. \tag{1.9}$$

Moreover, for every $p \in [1, +\infty)$,

$$\sup_{t\in\mathbb{R}} \mathbb{E}\left[\|\tau(t)\|^p_{\mathcal{C}^{|\tau|}}\right] < +\infty, \tag{1.10}$$

and for every $p \in [1, \infty)$ and $\lambda \in [0, 1]$,

$$\sup_{s<t} \frac{\mathbb{E}\left[\|\tau(t) - \tau(s)\|^p_{\mathcal{C}^{|\tau|-2\lambda}}\right]}{|t-s|^\lambda} < +\infty. \tag{1.11}$$

Remark 1.2. Since the processes we consider are stationary in time, the constants in (1.8) do not depend on the time t. Moreover, the suprema in (1.9) and (1.10) are superfluous. We prefer to write them nonetheless, since the statements with the suprema are robust to perturbations of the stationarity property.

Remark 1.3. As will be seen below, the constants $\mathbf{c}_n$ and $\mathbf{c}'_n$ diverge at order n and $\log n$ respectively.

Remark 1.4. The bound (1.11) immediately implies the pathwise Hölder continuity of the symbols, by the Kolmogorov continuity test. In the construction of solutions to (1.1), this strong control on the Hölder regularity of τ is only needed for the symbol . For the remaining symbols, a weaker bound of the type

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|\tau(t)\|^p_{\mathcal{C}^{|\tau|}}\Big]$$

suffices. However, the proofs of (1.11) and (1.10) are relatively similar anyway, as we hope to convince the reader below.

This note is organised as follows. In Section 2, we introduce Besov spaces. The diagrams take values in these spaces. The definition of these function spaces is based on the Littlewood-Paley decomposition. This allows us to define paraproducts along the way, and to give relevant intuition for them. In Section 3, we introduce the white noise process and discuss the property of equivalence of moments. The latter is very convenient to reduce the bounds (1.10) and (1.11) to easy-to-check second moment computations. In Section 4, we construct the diagrams and prove the fixed time bound (1.10). In Section 5, we briefly discuss how the bound (1.11) for time differences follows easily from the fixed time one. Finally, in the appendix, we give an alternative proof of the equivalence of moments property exposed in Section 3.

2. Function spaces and paraproducts

In this section, we introduce the function spaces we will use, denoted by $\mathcal{C}^\alpha$, for $\alpha \in \mathbb{R}$. When $\alpha \in (0, 1)$, they are (a separable version of) the usual Hölder spaces. For general α, they belong to the larger class of Besov spaces, and enjoy remarkable stability properties under multiplication. We choose to also give an informal presentation of these properties, although we will not refer to these in

our actual construction of the diagrams. Our choice is motivated by the fact that the question of defining products of distributions is central to making sense of the Φ^4 model. It is therefore useful to survey first what can be achieved with purely deterministic methods (and what can be ultimately used to show well-posedness of the Φ^4 model). Moreover, exploring this question naturally leads to the introduction of paraproducts. In this section, the space dimension d is arbitrary. For most results, we only provide a sketch of proof. A much more detailed treatment of the topics discussed in this section can be found in [1, Chapter 2].

We wish to extend the notion of α-Hölder regularity of a distribution f on $\mathbb{T}^d$ to exponents $\alpha < 0$. Roughly speaking, this should mean that for every $x \in \mathbb{T}^d$ and every test function $\varphi \in C_c^\infty(\mathbb{R}^d)$,

$$\left\langle f, \varepsilon^{-d}\varphi(\varepsilon^{-1}(\cdot - x))\right\rangle \lesssim \varepsilon^\alpha \quad \text{uniformly in } x \in \mathbb{T}^d \text{ and } \varepsilon \to 0, \tag{2.1}$$

where we interpret f as a periodic distribution on $\mathbb{R}^d$ in the duality pairing above[1]. A precise definition can be built along these lines (the interested reader can find it in [15, Definition 3.7]). However, we prefer to adopt here a point of view based on Fourier analysis, which allows for a more direct understanding of the stability of the spaces under multiplication.

Remark 2.1. For *positive* α, the condition (2.1) should be replaced by

$$\left\langle f - p_x(\cdot), \varepsilon^{-d}\varphi(\varepsilon^{-1}(\cdot - x))\right\rangle \lesssim \varepsilon^\alpha \quad \text{uniformly in } x \in \mathbb{T}^d \text{ and } \varepsilon \to 0, \tag{2.2}$$

where p_x is the Taylor approximation of order $\lfloor \alpha \rfloor$ of f. For negative α, there is no such polynomial, and this recentering is unnecessary.

For every $f \in L^1(\mathbb{T}^d)$ and $\omega \in \mathbb{Z}^d$, we write

$$\mathscr{F} f(\omega) = \widehat{f}(\omega) := \int_{\mathbb{T}^d} f(x) e^{-2i\pi\omega\cdot x}\,\mathrm{d}x, \tag{2.3}$$

for the Fourier coefficient of f with frequency ω, so that

$$f(x) = (\mathscr{F}^{-1}\widehat{f})(x) := \sum_{\omega\in\mathbb{Z}^d} \widehat{f}(\omega) e^{2i\pi\omega\cdot x},$$

where $\mathscr{F}^{-1}$ denotes the inverse Fourier transform.

The definition of Besov spaces rests on a decomposition of the Fourier series of a function along dyadic annuli, an idea due to Littlewood and Paley. More precisely, we think of splitting $\widehat{f}$ into

$$\widehat{f}\,\mathbf{1}_{B(0,1)} + \sum_{k=0}^{+\infty} \widehat{f}\,\mathbf{1}_{B(0,2^{k+1})\setminus B(0,2^k)}, \tag{2.4}$$

where $B(0,r) := \{\omega \in \mathbb{Z}^d \,:\, |\omega| < r\}$. In (2.4), the terms of the series associated with large k's measure the fast oscillations of the function; the general Besov norm can be thought of as a weighted average of the L^p norm of these summands. This type of decomposition enjoys better analytical properties if we replace the

[1]Here and below we write $A \lesssim B$ to mean that there exists a constant C, which is independent of the quantities of interest, such that $A \le CB$.

indicator functions by smoothened versions thereof. More precisely, we can find functions $\widetilde{\chi}, \chi \in C_c^\infty(\mathbb{R}^d)$ both taking values in $[0,1]$, with supports

$$\mathrm{Supp}\tilde{\chi} \subseteq B\left(0, \frac{4}{3}\right), \qquad \mathrm{Supp}\chi \subseteq B\left(0, \frac{8}{3}\right) \setminus B\left(0, \frac{3}{4}\right),$$

and such that

$$\widetilde{\chi}(\zeta) + \sum_{k=0}^{+\infty} \chi(\zeta/2^k) = 1, \quad \forall \zeta \in \mathbb{R}^d. \tag{2.5}$$

We may furthermore choose these functions to be radially symmetric. We write

$$\chi_{-1} := \widetilde{\chi}, \qquad \chi_k(\cdot) := \chi(\cdot/2^k) \quad k \geq 0. \tag{2.6}$$

The supports of $\tilde{\chi}$ and χ ensure that χ_k and $\chi_{k'}$ overlap only if $|k-k'| \leq 1$ (see [1, Proposition 2.10] for more details). For every $f \in C^\infty(\mathbb{T})$ and $k \geq -1$, we let

$$\delta_k f := \mathscr{F}^{-1}\left(\chi_k \widehat{f}\right),$$

so that $\widehat{f} = \sum_{k\geq -1} \chi_k \widehat{f}$ (compare with (2.4)) and $f = \sum_{k\geq -1} \delta_k f$. Let

$$\eta_k = \mathscr{F}^{-1}(\chi_k), \qquad \eta = \eta_0. \tag{2.7}$$

For $k \geq 0$, we have

$$\eta_k \simeq 2^{dk} \eta(2^k \cdot), \tag{2.8}$$

up to a small error due to the fact that our phase space $\mathbb{Z}^d$ is discrete (since our state space $\mathbb{T}^d$ is compact)[2]. For every $k \geq -1$, we have

$$\delta_k f = \eta_k \star f, \tag{2.9}$$

where $\star$ denotes the convolution on the torus $\mathbb{T}^d$. In agreement with (2.1), (2.8) and (2.9), we define the $\mathcal{C}^\alpha$ norm by

$$\|f\|_{\mathcal{C}^\alpha} := \sup_{k\geq -1} 2^{\alpha k} \|\delta_k f\|_{L^\infty}. \tag{2.10}$$

One can check that this quantity is finite if $f \in C^\infty(\mathbb{T}^d)$. The space $\mathcal{C}^\alpha$ is the completion of $C^\infty(\mathbb{T}^d)$ with respect to this norm. This space can be realised as a subspace of the space of Schwartz distributions.

Remark 2.2. We depart slightly from the convention to define $\mathcal{C}^\alpha$ as the space of distributions with finite $\|\cdot\|_{\mathcal{C}^\alpha}$ norm. Our definition makes the space separable and allows us below to define products of distributions and functions via approximation. Moreover, one can check that if a distribution f satisfies $\|f\|_{\mathcal{C}^\alpha} < \infty$, then $f \in \mathcal{C}^\beta$ for every $\beta < \alpha$.

The most important property of Besov spaces for our purpose is the following multiplicative structure.

Proposition 2.3. *Let $\alpha < 0 < \beta$ be such that $\alpha + \beta > 0$. The multiplication $(f,g) \mapsto fg$ extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$.*

[2]One may estimate the error in (2.8) and show that it is negligible for our purpose, but the simplest way around this technical point is probably to interpret each periodic function on $\mathbb{T}^d$ as a periodic function on $\mathbb{R}^d$, and then use $L^p(\mathbb{R}^d)$ norms and the continuous Fourier transform throughout, so that (2.8) becomes exact. The continuous Fourier transform of any Schwartz distribution is well-defined, by duality. For a periodic $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, the Fourier transform is a sum of Dirac masses at every $\omega \in \mathbb{Z}^d$, each carrying a mass $\widehat{f}(\omega)$ as defined in (2.3).

The proof of this proposition rests on the decomposition

$$fg = \sum_{k<l-1} \delta_k f\, \delta_l g + \sum_{|k-l|\le 1} \delta_k f\, \delta_l g + \sum_{k>l+1} \delta_k f\, \delta_l g,$$

which we will write suggestively in the form

$$fg = f \varolessthan g + f \circ g + f \varogreaterthan g.$$

This is often called *Bony's decomposition* into the *paraproducts* $f \varolessthan g$, $f \varogreaterthan g = g \varolessthan f$, and the *resonant product* $f \circ g$.

In order to prove Proposition 2.3, it suffices to show that each of these terms extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$. However, it is very important for our more general goal of making sense of the Φ^4 model to be precise about the behaviour of each term separately. We thus simply assume that $\alpha < 0 < \beta$ and $f \in \mathcal{C}^\alpha$, $g \in \mathcal{C}^\beta$ to begin with (but do not yet prescribe the sign of $\alpha + \beta$), and see whether and how we can estimate each of the terms in Bony's decomposition. We start with

$$f \varolessthan g = \sum_{k<l-1} \delta_k f\, \delta_l g. \tag{2.11}$$

In view of (2.10), in order to see which Hölder class $f \varolessthan g$ belongs to, we need to estimate

$$\|\delta_j (f \varolessthan g)\|_{L^\infty} = \Big\| \sum_{l=-1}^{+\infty} \delta_j \Big(\sum_{k=-1}^{l-2} \delta_k f \delta_l g \Big) \Big\|_{L^\infty}. \tag{2.12}$$

Recall that the Fourier transform of $\delta_k f$ (resp. $\delta_l g$) is supported on an annulus of both inner and outer radius of size about 2^k (resp. 2^l). The important observation is that as we sum over $k \le l-2$, the Fourier transform of $\delta_k f \delta_l g$ is still supported in an annulus of inner and outer radius both proportional to 2^l. Hence, only a finite number of terms $l \simeq j$ contribute to the outer sum on the right hand side above. Now, since

$$\|\delta_k f\|_{L^\infty} \le 2^{-\alpha k} \|f\|_{\mathcal{C}^\alpha}, \qquad \|\delta_l g\|_{L^\infty} \le 2^{-\beta l} \|g\|_{\mathcal{C}^\alpha},$$

the right hand side of (2.12) can then be bounded by

$$\| \sum_{k=-1}^{j-2} \delta_k f \delta_j g \|_{L^\infty} \le \sum_{k=-1}^{j-2} 2^{-\alpha k - \beta j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta} < C 2^{-(\alpha+\beta)j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta},$$

where we used that $\alpha < 0$, and C does not depend on f or g. This shows that

$$\|f \varolessthan g\|_{\mathcal{C}^{\alpha+\beta}} \le C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

The same analysis applies for the term $g \varolessthan f$, except that since we have $\beta > 0$, we get

$$\sum_{l=-1}^{j-2} 2^{-\alpha j - \beta l} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta} \le C 2^{-\alpha j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta},$$

which implies that

$$\|g \varolessthan f\|_{\mathcal{C}^\alpha} \le C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

Note that we have made no assumption on the sign of $\alpha + \beta$ so far. The term $f \varolessthan g$ inherits essential features of the small scale behaviour of g "modulated" by the low frequency modes of f. This is in agreement with the fact that $f \varolessthan g$ is more regular than $g \varolessthan f$ under our hypothesis $\alpha < \beta$.

	$f \circledcirc g$	$g \circledcirc f$	$f \circleddash g$	fg
Regularity	$\alpha+\beta$	α	$\alpha+\beta$	α
Needs $\alpha+\beta>0$	No	No	Yes	Yes

TABLE 2.1. Summary of the regularity properties of paraproducts for $\alpha < 0 < \beta$, $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$.

We now turn to the resonant term $f \circleddash g$, which for simplicity we think of as being

$$\sum_{k=-1}^{+\infty} \delta_k f \, \delta_k g.$$

The crucial difference in the analysis of this term, compared with the previous computations, is that the support of the Fourier transform of the summand indexed by k, which is the convolution of the annulus of radius about 2^k by itself, results in a *ball* of radius 2^k, as opposed to an *annulus*. Therefore, every summand indexed by $k \geq l$ contributes the the l-th Littlewood-Paley block $\delta_l(f \circleddash g)$. The L^∞ norm of each summand is bounded by

$$2^{-(\alpha+\beta)k} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

If we want this to be summable over $k \geq l$, we need to assume $\alpha+\beta>0$. In this case, the sum is of order $2^{-(\alpha+\beta)l}\|f\|_{\mathcal{C}^\alpha}\|g\|_{\mathcal{C}^\beta}$, which suggests that

$$\|f \circleddash g\|_{\mathcal{C}^{\alpha+\beta}} \leq C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

These computations can all be made rigorous (see e.g. [1]), and are summarised in Table 2.1. Note that these relations are relevant in the construction of Φ^4 since there, one needs to characterise the products between $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$ (though sometimes necessarily $\alpha+\beta<0$, and renormalisations are then needed for the resonant term).

We point out that the regularising effect of the heat kernel can be conveniently measured using the spaces $\mathcal{C}^\alpha$. While we will not use this proposition in itself here, it is a useful guide to the intuition. In particular, the time singularity in (2.13) is integrable as long as the difference of regularity exponents is less than 2. In other words, the integration operator I in (1.6) brings a gain of 2 units of regularity.

Proposition 2.4. *If $\alpha \leq \beta \in \mathbb{R}$, then there exists $C<\infty$ such that for every $t>0$, we have*

$$\|e^{t\Delta} f\|_{\mathcal{C}^\beta} \leq C\, t^{\frac{\alpha-\beta}{2}} \|f\|_{\mathcal{C}^\alpha}. \tag{2.13}$$

Sketch of proof. The Laplacian Δ is a multiplication operator in Fourier space. As a consequence, we have $\delta_k(e^{t\Delta}f) = e^{t\Delta}(\delta_k f)$, and since $\widehat{\Delta}(\omega) = -|\omega|^2$, roughly speaking, we have $e^{t\Delta}(\delta_k f) \simeq e^{-t2^{2k}}\delta_k f$. This suggests that

$$\|\delta_k(e^{t\Delta}f)\|_{L^\infty} \leq C \exp\left(-t2^{2k}\right) \|\delta_k f\|_{L^\infty} \leq C \exp\left(-t2^{2k}\right) 2^{-\alpha k} \|f\|_{\mathcal{C}^\alpha},$$

and therefore

$$2^{\beta k}\|\delta_k(e^{t\Delta}f)\|_{L^\infty} \leq C\left[(2^{2k}t)^{\frac{\beta-\alpha}{2}} \exp\left(-2^{2k}t\right)\right] t^{\frac{\alpha-\beta}{2}} \|f\|_{\mathcal{C}^\beta}.$$

Since the term between square brackets is bounded uniformly over k and t, the heuristic argument is complete. A rigorous proof can be derived using techniques similar to those exposed for Lemma 2.6 below (see also e.g. [20, Proposition 3.11]). □

In view of (2.8), we expect that for every $p, p' \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{p'} = 1$,

$$\sup_{k \geq -1} 2^{-\frac{dk}{p}} \|\eta_k\|_{L^{p'}} < \infty. \tag{2.14}$$

Indeed, the relation (2.8), and therefore the inequality (2.14), are immediate by scaling if the torus $\mathbb{T}^d$ is replaced by the full space $\mathbb{R}^d$ (and therefore the discrete Fourier series is replaced by the continuous Fourier transform). A rigorous proof of (2.14) can be found e.g. in [22, Lemma B.1]. By (2.9) and Hölder's inequality, we deduce the following lemma.

Lemma 2.5. *Let $p \in [1, \infty]$. There exists $C < \infty$ such that*

$$\|\delta_k f\|_{L^\infty} \leq C 2^{\frac{dk}{p}} \|f\|_{L^p} \tag{2.15}$$

for every $f \in C^\infty(\mathbb{T}^d)$ and $k \geq -1$.

If we had chosen to define the decomposition $f = \sum_{k \geq -1} \delta_k f$ from the Fourier series decomposition displayed in (2.4) based on indicator functions, then we would have $\delta_k \delta_k f = \delta_k f$, and we could therefore replace f by $\delta_k f$ on the right side of (2.15). With our actual definition of δ_k, this replacement is also possible: instead of using that

$$\mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)} \mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)} = \mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)},$$

we choose a smooth function $\chi' \in C_c^\infty(\mathbb{R}^d)$ with support in an annulus and such that $\chi' \equiv 1$ on the support of χ, so that

$$\chi' \chi = \chi.$$

Setting $\delta_k' f := \mathscr{F}^{-1}(\chi'(\cdot/2^k)\widehat{f})$, the identity above translates into

$$\delta_k' \delta_k f = \delta_k f. \tag{2.16}$$

Next, we verify that the argument leading to Lemma 2.5 also applies if $\delta_k f$ is replaced by $\delta_k' f$ on the left side of (2.15). Using (2.16), we thus obtain the following lemma.

Lemma 2.6 (Bernstein inequality)**.** *Let $p \in [1, \infty]$. There exists $C < \infty$ such that*

$$\|\delta_k f\|_{L^\infty} \leq C 2^{\frac{dk}{p}} \|\delta_k f\|_{L^p} \tag{2.17}$$

for every $f \in C^\infty(\mathbb{T}^d)$ and $k \geq -1$.

The following proposition uses the previous lemma to bound p-th moments of the Hölder norm of a random distribution in terms of the p-th moments of its decomposition in Fourier space.

Proposition 2.7 (Boundedness criterion)**.** *Let $\beta < \alpha - \frac{d}{p}$. There exists $C < \infty$ such that for every random distribution f on $\mathbb{T}^d$, we have*

$$\mathbb{E}\left[\|f\|_{\mathcal{C}^\beta}^p\right] \leq C \sup_{k \geq -1} 2^{\alpha k p} \mathbb{E}\left[\|\delta_k f\|_{L^p}^p\right]. \tag{2.18}$$

Proof. By definition of the $\mathcal{C}^\beta$ norm and then Lemma 2.6, we have

$$\|f\|^p_{\mathcal{C}^\beta} = \sup_{k\geq -1} 2^{\beta kp}\|\delta_k f\|^p_{L^\infty} \leq C \sup_{k\geq -1} 2^{k(\beta p+d)}\|\delta_k f\|^p_{L^p}.$$

In order to take the expectation of $\|\delta_k f\|^p_{L^p}$ directly, we enlarge the supremum on the right side above to a sum, and get

$$\mathbb{E}\|f\|^p_{\mathcal{C}^\beta} \leq C \sum_{k\geq -1} 2^{k(\beta p+d)}\mathbb{E}\|\delta_k f\|^p_{L^p} = C \sum_{k\geq -1} 2^{kp(\beta+\frac{d}{p}-\alpha)}2^{\alpha kp}\mathbb{E}\|\delta_k f\|^p_{L^p}.$$

The announced estimate then follows, since $\alpha > \beta + \frac{d}{p}$. □

Remark 2.8. With our definition of the space $\mathcal{C}^\beta$ as a completion, the fact that a distribution f satisfies $\|f\|_{\mathcal{C}^\beta} < \infty$ does not imply that $f \in \mathcal{C}^\beta$. However, in the context of Proposition 2.7, when the right side of (2.18) is finite, we do have $f \in \mathcal{C}^\beta$ with probability one. Indeed, we can always pick $\beta' \in (\beta, \alpha - \frac{d}{p})$, deduce that $\|f\|_{\mathcal{C}^{\beta'}}$ is finite with probability one, and conclude by Remark 2.2.

3. White noise and Nelson's estimate

Proposition 2.7 gives a criterion for determining whether a random distribution belongs to $\mathcal{C}^\beta$ by looking at the p-th moment of its Paley-Littlewood blocks. However, it is often difficult to get sharp bounds of high moments of a random distribution. On the other hand, fortunately, the objects we encounter in the construction of Φ^4_3 (and many other Gaussian models) are all built from multiplications of finitely many Gaussian random variables. These objects belong to a Wiener chaos of finite order, and we can therefore leverage on the equivalence of moments property, also often called Nelson's estimate, to deduce high-moment estimates from second-moment calculations. The purpose of this section is to present these arguments, and state the implied simpler criterion for belonging to $\mathcal{C}^\beta$. In this section, we continue to work in arbitrary space dimension d. We only sketch some well-known arguments concerning iterated integrals and Wiener chaos, and refer the interested reader to [18, Chapter 9] for a more detailed exposition of these topics.

We start by introducing the space-time white noise. Formally, the space-time white noise ξ is a centred Gaussian distribution on $\mathbb{R}\times\mathbb{T}^d$ with covariance

$$\mathbb{E}\xi(t,x)\,\xi(t',x') = \delta(t-t')\,\delta^d(x-x')\,, \tag{3.1}$$

where $\delta(\cdot)$ and $\delta^d(\cdot)$ denote Dirac delta functions over $\mathbb{R}$ and $\mathbb{T}^d$ respectively. Testing (3.1) against a function $\varphi : \mathbb{R}\times\mathbb{T}^d \to \mathbb{R}$ leads us to postulate that

$$\mathbb{E}\left[\xi(\varphi)^2\right] = \|\varphi\|^2_{L^2(\mathbb{R}\times\mathbb{T}^d)}\,. \tag{3.2}$$

Definition 3.1. *A space-time white noise over $\mathbb{R}\times\mathbb{T}^d$ is a family of centred Gaussian random variables $\{\xi(\varphi),\, \varphi \in L^2(\mathbb{R}\times\mathbb{T}^d)\}$ such that (3.2) holds.*

The existence of a space-time white noise follows from Kolmogorov's extension theorem. We prefer here to take a more constructive approach, based on Fourier analysis.

Let $(W(\cdot,\omega))_{\omega\in\mathbb{Z}^d}$ be a family of complex-valued Brownian motions over $\mathbb{R}$. These Brownian motions are independent for different values of ω except for

the constraint $W(t,\omega) = \overline{W(t,-\omega)}$ (so that the white noise we are building is real-valued). The magnitude of the variance is fixed by the condition

$$\mathbb{E}\left[W(t,\omega)W(t,-\omega')\right] = \begin{cases} |t| & \text{if } \omega = \omega', \\ 0 & \text{otherwise.} \end{cases}$$

We then set ξ to be the time derivative of the cylindrical Wiener process

$$(t,x) \mapsto \sum_{\omega\in\mathbb{Z}^d} W(t,\omega)e^{2i\pi\omega\cdot x}.$$

More precisely, for every $\varphi \in L^2(\mathbb{R}\times\mathbb{T}^d)$, we set

$$\xi(\varphi) := \sum_{\omega\in\mathbb{Z}^d} \int_{t=-\infty}^{+\infty} \widehat{\varphi}(t,\omega)\,\mathrm{d}W(t,\omega),$$

where the integral is interpreted in Itô's sense, and the notation $\widehat{\varphi}(t,\omega)$ stands for $\widehat{\varphi(t,\cdot)}(\omega)$. By Itô's and Fourier's isometries, this expression is well-defined and the relation (3.2) is satisfied. Since $\xi(\varphi)$ is also a centered Gaussian, this provides us with a construction of white noise. We use the somewhat informal notation

$$\xi(\varphi) := \int_{\mathbb{R}\times\mathbb{T}^d} \varphi(z)\,\xi(\mathrm{d}z)\,,$$

although ξ is almost surely not a measure. In particular, for a given $\varphi \in L^2(\mathbb{R}\times\mathbb{T}^d)$, the random variable $\xi(\varphi)$ is only defined outside of a set of measure zero, and a priori this set depends on the choice of φ.

As explained for example in [18, Chapter 9] or [23, Section 1.1.2], we can define iterated Wiener-Itô integrals based on ξ. For each $k \geq 1$ and $\varphi \in L^2((\mathbb{R}\times\mathbb{T}^d)^k)$, we denote the iterated integral of φ by

$$\xi^{\otimes k}(\varphi) = \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \varphi(z_1,\dots,z_k)\,\xi(\mathrm{d}z_1)\cdots\xi(\mathrm{d}z_k).$$

Denoting by $\widetilde{\varphi}$ the symmetrized function obtained from φ:

$$\widetilde{\varphi}(z_1,\dots,z_k) := \frac{1}{k!}\sum_{\sigma\in S_k} \varphi(z_{\sigma(1)},\dots,z_{\sigma(k)}), \tag{3.3}$$

where S_k denotes the permutation group over $\{1,\dots,k\}$, we have

$$\xi^{\otimes k}(\varphi) = \xi^{\otimes k}(\widetilde{\varphi}) = k! \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}(z_1,\dots,z_k)\,\mathbf{1}_{\{t_1<\cdots<t_k\}}\,\xi(\mathrm{d}z_1)\cdots\xi(\mathrm{d}z_k), \tag{3.4}$$

where t_i is the time component of z_i. Moreover, we have the isometry property

$$\mathbb{E}\left[\xi^{\otimes k}(\varphi)^2\right] = \mathbb{E}\left[\xi^{\otimes k}(\widetilde{\varphi})^2\right] = \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}^2(z_1,\dots,z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k, \tag{3.5}$$

and by Jensen's inequality,

$$\int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}^2(z_1,\dots,z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k \leq \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \varphi^2(z_1,\dots,z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k. \tag{3.6}$$

Assuming now for notational convenience that φ is a symmetric function, that is, $\varphi = \widetilde{\varphi}$, we may rewrite the expression (3.4) for the iterated integral $\xi^{\otimes k}(\varphi)$ as a series of finite-dimensional iterated Wiener-Itô integrals:

$$\xi^{\otimes k}(\varphi) = k! \sum_{\omega_1,\dots,\omega_k\in\mathbb{Z}^d} \int_{t_1<\cdots<t_k} \widehat{\varphi}(t_1,\omega_1,\dots,t_k,\omega_k)\,\mathrm{d}W(t_1,\omega_1)\cdots\mathrm{d}W(t_k,\omega_k), \tag{3.7}$$

where

$$\widehat{\varphi}(t_1,\omega_1,\dots,t_k,\omega_k) := \int_{(\mathbb{T}^d)^k} \varphi(t_1,x_1,\dots,t_k,x_k) e^{-2i\pi(\omega_1\cdot x_1+\dots+\omega_k\cdot x_k)}\,\mathrm{d}x_1\cdots\mathrm{d}x_k.$$

See also [18, Section 9.6] for a definition of iterated integrals. We let

$$\mathcal{H}_k := \left\{\xi^{\otimes k}(\varphi),\ \varphi\in L^2((\mathbb{R}\times\mathbb{T}^d)^k)\right\}$$

denote the k-th *Wiener chaos*, with $\mathcal{H}_0=\mathbb{R}$. Denote by $(\Omega,\mathcal{F},\mathbb{P})$ the probability space on which ξ is defined. The spaces $\mathcal{H}_k$ are orthogonal in $L^2(\Omega,\mathcal{F},\mathbb{P})$. Moreover, although we will not use it, we recall that if $\mathcal{F}$ is the σ-algebra generated by $\mathcal{H}_1$, then

$$L^2(\Omega,\mathcal{F},\mathbb{P}) = \bigoplus_{k=0}^{+\infty}\mathcal{H}_k$$

(the interested reader may find this result in [18, Section 9.5]). The more important property for our purpose is the following.

Lemma 3.2. *For each $n\in\mathbb{N}$, the closure in $L^2(\Omega,\mathcal{F},\mathbb{P})$ of the linear span of the set*

$$\left\{\xi(\varphi_1)\cdots\xi(\varphi_k),\ k\le n,\ \varphi_1,\dots,\varphi_k\in L^2(\mathbb{R}\times\mathbb{T}^d)\right\}. \tag{3.8}$$

coincides with

$$\mathcal{H}_{\le n} := \bigoplus_{k=0}^{n}\mathcal{H}_k. \tag{3.9}$$

Sketch of proof. Let $H_n = H_n(X,T)$ be the Hermite polynomials, defined recursively by

$$\begin{cases} H_0 = 1,\\ H_n = XH_{n-1} - T\,\partial_X H_{n-1} \qquad (n\in\mathbb{N}), \end{cases} \tag{3.10}$$

so that $H_1=X$, $H_2=X^2-T$, $H_3=X^3-3XT$, etc. By a recursive application of Itô's formula, see [18, Theorem 9.6.9], we have, for every $\varphi\in L^2(\mathbb{R}\times\mathbb{T}^d)$ and $n\in\mathbb{N}$,

$$\xi^{\otimes n}(\varphi^{\otimes n}) = H_n(\xi(\varphi),\|\varphi\|^2_{L^2(\mathbb{R}\times\mathbb{T}^d)}). \tag{3.11}$$

This relation already shows that every n-fold iterated integral is a linear combination of elements of (3.8). Conversely, it also shows that every Hermite polynomial in $\xi(\varphi)$ of degree at most n — and therefore every polynomial in $\xi(\varphi)$ of degree at most n — belongs to $\mathcal{H}_{\le n}$. The full proof of Lemma 3.2 can be derived from [18, Theorem 9.5.4]. □

By extension, we say that a stochastic process $\tau:\mathbb{R}\to\mathcal{S}'(\mathbb{T}^d)$ belongs to $\mathcal{H}_n$ (resp. $\mathcal{H}_{\le n}$) if for every t and smooth test function ϕ, we have

$$\langle\tau(t),\phi\rangle\in\mathcal{H}_n \text{ (resp. } \mathcal{H}_{\le n}).$$

If $\tau(t)$ is in fact a continuous function of the space variable, this boils down to asking that $\tau(t,x)\in\mathcal{H}_n$ (resp. $\mathcal{H}_{\le n}$) for each $x\in\mathbb{T}^d$. By Lemma 3.2, the approximations to the diagrams we want to construct, see (1.7), all belong to $\mathcal{H}_{\le 5}$. Since Wiener chaoses are closed, this remains true of their candidate limits (see also Section 4 for explicit representations).

Since $\delta_k f$ is linear in f for every k, the fact that f belongs to some Wiener chaos implies that $\delta_k f$ belongs to the Wiener chaos of the same order. Thus, in

view of Proposition 2.7, the possibility to estimate arbitrarily high moments of elements of a fixed Wiener chaos from their L^2 moments will be very convenient.

Proposition 3.3 (Nelson's estimate)**.** *For every $n \geq 1$ and $p \in [2, \infty)$, there exists a constant $C < \infty$ such that for every $X \in \mathcal{H}_{\leq n}$,*

$$\mathbb{E}\left[|X|^p\right]^{\frac{1}{p}} \leq C\mathbb{E}\left[X^2\right]^{\frac{1}{2}}.$$

If $X \in \mathcal{H}_n$, then we can take $C = (p-1)^{\frac{n}{2}}$.

We now give a proof of Proposition 3.3 based on the Burkholder-Davis-Gundy inequality ([28, Section IV.4]). The appendix contains an alternative argument based on the logarithmic Sobolev inequality.

Lemma 3.4 (BDG inequality)**.** *Let $p \in (0, \infty)$. There exists $C < \infty$ such that if $(M_t)_{t \geq 0}$ is a continuous local martingale starting from 0, then*

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} |M_s|^p\right] \leq C\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right],$$

where $\langle M \rangle$ denotes the quadratic variation of M.

First proof of Proposition 3.3. We show the result by induction on n. For $n = 0$, the space $\mathcal{H}_0$ only contains constants, so the result is obvious. We now fix $n \geq 1$. We need to verify the property for random variables of the form $\xi^{\otimes n}(\varphi)$, $\varphi \in L^2((\mathbb{R} \times \mathbb{T}^d)^n)$. By (3.4), we may assume that φ is symmetric in its variables. Let $\mathcal{F}_t$ be the σ-algebra generated by

$$\left\{\xi(\varphi),\ \operatorname{Supp}\varphi \subseteq (-\infty, t] \times \mathbb{T}^d\right\}.$$

The process

$$M_t := \int_{(\mathbb{R}\times\mathbb{T}^d)^n} \varphi(z_1, \dots, z_n)\, \mathbf{1}_{\{t_1 < \dots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_n) \qquad (t \in \mathbb{R})$$

is an $(\mathcal{F}_t)$-martingale. This can be justified either by approximation of φ by elementary functions which vanish on the diagonal (see [23, (1.10)], and take the A_i there of product form), or by appealing to the representation (3.7). Moreover,

$$\langle M \rangle_t = \int_{\mathbb{R}\times\mathbb{T}^d} \left(\int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi(z_1, \dots, z_n)\, \mathbf{1}_{\{t_1 < \dots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_{n-1}) \right)^2 \mathrm{d}z_n.$$

By Minkowski's triangle inequality for the exponent $\frac{p}{2} \geq 1$,

$$\begin{aligned} &\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right]^{\frac{2}{p}} \\ &\quad \leq \int_{\mathbb{R}\times\mathbb{T}^d} \mathbb{E}\left[\left(\int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi(z_1, \dots, z_n)\, \mathbf{1}_{\{t_1 < \dots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_{n-1}) \right)^p \right]^{\frac{2}{p}} \mathrm{d}z_n. \end{aligned}$$

By symmetrizing φ as in (3.3), the induction hypothesis and (3.5), we infer that

$$\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right]^{\frac{2}{p}} \leq C \int_{\mathbb{R}\times\mathbb{T}^d} \int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi^2(z_1, \dots, z_n)\, \mathbf{1}_{\{t_1 < \dots < t_n < t\}}\, \mathrm{d}z_1 \cdots \mathrm{d}z_{n-1}\, \mathrm{d}z_n.$$

The conclusion then follows from the Lemma 3.4, used with $t = +\infty$. □

Remark 3.5. As hinted at in the introduction, certain Ising-type Markov processes converge (or are expected to converge) to the Φ^4 model (see [22]). The proof of Nelson's estimate based on the BDG inequality is sufficiently robust to allow for a generalization to these Markov processes. Indeed, loosely speaking, a Markov process can be thought of as an evolution equation with a random forcing that is white in time. In more precise words, a Markov process comes with a martingale structure indexed by time, and the possibly surprisingly special role played by the time variable in the proof we presented above becomes very natural in this context. Using versions of Itô's formula and the BDG inequality for processes with jumps ([22, Appendix C]), one can show that (3.11) still holds approximately ([22, Proposition 5.3]) and prove a version of Nelson's estimate ([22, Lemma 4.1]) by following essentially the same reasoning as above.

With Nelson's estimate, we are now ready to provide a simple criterion to check the main convergence result in Theorem 1.1. Since for most of the processes defined in (1.7), their limits can be characterised explicitly without referring to the limiting procedure as $n \to +\infty$ (and in the cases when the limiting procedure is necessary, it is also obvious what the limit should be), we only give detailed characterisations of the limiting processes τ's themselves. Once all the properties of the limits are well understood, the convergence does not pose any further problem.

Proposition 3.6. *Let $n \in \mathbb{N}$, and let $\tau : \mathbb{R} \to \mathcal{S}'(\mathbb{T}^d)$ be a random process in $\mathcal{H}_{\leq n}$ which is stationary in space, in the sense that for every $x \in \mathbb{T}^d$,*

$$\text{the processes } (\tau(t,\cdot))_{t\in\mathbb{R}} \text{ and } (\tau(t,x+\cdot))_{t\in\mathbb{R}} \text{ have the same law.} \tag{3.12}$$

Let $(\widehat{\tau}(t,\omega))_{\omega\in\mathbb{Z}^d}$ denote the Fourier coefficients of $\tau(t)$. If for some $t \in \mathbb{R}$, there exists $C < \infty$ and $\alpha \in \mathbb{R}$ such that for every $\omega \in \mathbb{Z}^d$,

$$\mathbb{E}\big[|\widehat{\tau}(t,\omega)|^2\big] \leq C(1+|\omega|)^{-d-2\alpha}, \tag{3.13}$$

then for every $\beta < \alpha$, we have $\tau(t) \in \mathcal{C}^\beta(\mathbb{T}^3)$, and moreover,

$$\mathbb{E}\big[\|\tau(t)\|^p_{\mathcal{C}^\beta}\big] < +\infty. \tag{3.14}$$

If, in addition to (3.13), there exists $\lambda \in (0,1)$ such that

$$\mathbb{E}\big[|\widehat{\tau}(t,\omega)-\widehat{\tau}(s,\omega)|^2\big] \leq C|t-s|^\lambda(1+|\omega|)^{-d-2\alpha+2\lambda} \tag{3.15}$$

uniformly in $0 < |t-s| < 1$ and $\omega \in \mathbb{Z}^d$, then for every $\beta < \alpha - \lambda$, we have $\tau \in C(\mathbb{R}, \mathcal{C}^\beta(\mathbb{T}^3))$, and moreover,

$$\sup_{0<|t-s|<1} \frac{\mathbb{E}\big[\|\tau(t)-\tau(s)\|^p_{\mathcal{C}^\beta}\big]}{|t-s|^{\frac{\lambda p}{2}}} < +\infty. \tag{3.16}$$

Proof. Step 1. We first show that by the stationarity assumption (3.12), for every $\omega, \omega' \in \mathbb{Z}^d$,

$$\omega + \omega' \neq 0 \quad \Longrightarrow \quad \mathbb{E}\big[\widehat{\tau}(s,\omega)\widehat{\tau}(t,\omega')\big] = 0. \tag{3.17}$$

Indeed, we have, using a slightly informal integral-sign notation,

$$\mathbb{E}\left[\widehat{\tau}(s,\omega)\widehat{\tau}(t,\omega')\right] = \iint_{(\mathbb{T}^d)^2} \mathbb{E}[\tau(s,x)\tau(t,y)]e^{-2\pi i(\omega\cdot x+\omega'\cdot y)}\,dx\,dy$$
$$= \iint_{(\mathbb{T}^d)^2} \mathbb{E}[\tau(s,x)\tau(t,y)]e^{-2\pi i[(\omega+\omega')\cdot x+\omega'\cdot(y-x)]}\,dx\,dy.$$

By the stationarity assumption, the expectation above is a function of $(y-x)$ only. Integrating in y first and then in x, we therefore obtain (3.17).

Step 2. We now focus on the proof of (3.16); the proof of (3.14) is only simpler. Let $\tau_{s,t} := \tau(t) - \tau(s)$. We have

$$(\delta_k\tau_{s,t})(x) = \sum_{\omega} \chi_k(\omega)\widehat{\tau}_{s,t}(\omega)e^{2\pi i\omega\cdot x}.$$

We implicitly assume here that the processes under consideration are real-valued, and therefore that for every $\omega \in \mathbb{Z}^d$,

$$\widehat{\tau}(t,-\omega) = \overline{\widehat{\tau}(t,\omega)}.$$

Since χ_k is an even function, we deduce that

$$\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right] = \sum_{\omega,\omega'\in\mathbb{Z}^d} \chi_k(\omega)\chi_k(\omega')\mathbb{E}\left[\widehat{\tau}_{s,t}(\omega)\widehat{\tau}_{s,t}(\omega')\right] e^{2\pi i(\omega+\omega')\cdot x}.$$

Using (3.17) and the bound (3.15), we get

$$\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right] = \sum_{\omega\in\mathbb{Z}^d} |\chi_k(\omega)|^2\mathbb{E}\left[|\widehat{\tau}_{s,t}(\omega)|^2\right] \lesssim |t-s|^{\lambda}2^{-2k(\alpha-\lambda)}.$$

The above bound holds uniformly in $k \geq -1$, $0 < |t-s| < 1$ and $x \in \mathbb{T}^d$. Also, since τ belongs to $\mathcal{H}_{\leq n}$, Nelson's estimate implies

$$\mathbb{E}\left[\|\delta_k\tau_{s,t}\|_{L^p}^p\right] \leq \sup_{x\in\mathbb{T}^d} \mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^p\right]$$
$$\lesssim \sup_{x\in\mathbb{T}^d} \left(\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right]\right)^{\frac{p}{2}}$$
$$\lesssim |t-s|^{\frac{\lambda p}{2}}2^{-kp(\alpha-\lambda)}.$$

By Proposition 2.7, we deduce that for each $\beta < \alpha - \lambda - \frac{d}{p}$,

$$\mathbb{E}\left[\|\tau_{s,t}\|_{\mathcal{C}^\beta}^p\right] \lesssim |t-s|^{\frac{\lambda p}{2}},$$

uniformly over all $0 < |t-s| < 1$. □

4. Construction of the diagrams

We are now ready to construct the diagrams listed in Table 1.1, and prove the relevant bounds appearing in Theorem 1.1. We will focus on the bound (1.10) for fixed times, and will only briefly discuss in the next section how the continuity in time follows from there. We also omit the detailed proof of the convergence (1.9), since once the bounds (1.10) for the limit diagrams are established, the convergence of approximations follows in essentially the same way. Since all our processes τ belong to $\mathcal{H}_{\leq 5}$ and are stationary both in space and time, we can invoke Proposition 3.6 and reduce the proof of (1.10) to showing the second moment bound (3.13) for each τ.

The derivation of these bounds involves the estimation of several nested integrals and sums. We use a graphical notation to represent these operations. This has the advantage of making the manipulations with potentially very long expressions shorter and more transparent. Naturally, the price to pay is to get used to the notation. One of the aims of these notes is to convince the reader that this investment is worth their while.

The graphical notation is heavily inspired by the treatment of the stochastic terms in [15, Section 10]. One difference is that there the calculations are performed in "real space", while we prefer to work with the spatial Fourier transform, and to keep the time variable fixed. It turns out that despite this change, the graphs we encounter in our approach are very similar to the ones in [15, Section 10] — only the interpretation changes slightly. Another difference is that we work with resonant parts of products, while Hairer considers increments of processes. We comment on this difference below (see also [14, Section 5] for an earlier graphical approach to bounding stochastic quantities which are represented using the spatial Fourier transform).

The presentation is separated into two parts. We first show how to represent the various processes as iterated stochastic integrals, and then derive the bounds on these.

4.1. **Iterated integral representation.** We start by showing how all of the stochastic terms τ in Table 1.1 can be represented as sums of iterated stochastic integrals. As stated above, each of the terms τ is an element of the inhomogeneous Wiener chaos $\mathcal{H}_{\leq 5}$, and this sum yields the explicit decomposition into its components in the homogeneous Wiener chaoses $\mathcal{H}_0, \dots, \mathcal{H}_5$. This representation as iterated integrals reduces the proof of the required moment bounds to an application of the isometry property (3.5). As will be shown below, this representation also makes the choice of infinite renormalisation constants transparent.

Case $\tau = \bullet$. We take the simplest process $\tau = \bullet$, the solution to the stochastic heat equation (1.3), as the starting point of our discussion. For each $t \in \mathbb{R}$ and $\omega \in \mathbb{Z}^3$, we can write

$$\widehat{\bullet}(t,\omega) = \int_{u=-\infty}^{t} \widehat{P}_{t-u}(\omega)\, \mathrm{d}W(u,\omega), \tag{4.1}$$

where $(W(\cdot,\omega))_\omega$ is the family of complex valued Brownian motions introduced in Section 3, and, for $t \geq 0$, $\widehat{P}_t$ is the Fourier transform of the heat kernel for $P_t = e^{t(\Delta-1)}$, that is,

$$\widehat{P}_t(\omega) = e^{-t(1+4\pi^2|\omega|^2)} = e^{-t\langle\omega\rangle^2}, \tag{4.2}$$

where we set

$$\langle\omega\rangle := \sqrt{1+4\pi^2|\omega|^2}$$

for concision. It is convenient to extend P_t and $\widehat{P}_t$ to every time $t \in \mathbb{R}$, by setting

$$\text{for every } t<0, \qquad P_t \equiv 0 \quad \text{and} \quad \widehat{P}_t \equiv 0, \tag{4.3}$$

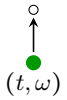
so that (4.1) can be rewritten as

$$\widehat{\bullet}(t,\omega) = \int_{u\in\mathbb{R}} \widehat{P}_{t-u}(\omega)\, \mathrm{d}W(u,\omega).$$

In future expressions of integrals against $\mathrm{d}W(u,\omega)$, we always understand that the variable ω is fixed, and that the variable u is the variable of integration. For instance, we simply write

$$\widehat{\Xi}(t,\omega) = \int_{\mathbb{R}} \widehat{P}_{t-u}(\omega)\,\mathrm{d}W(u,\omega). \tag{4.4}$$

The graphical version of (4.1) or (4.4) is

$\widehat{\Xi}(t,\omega) =$  . (4.5)

Here the root ● represents the pair (t,ω), i.e., the time and frequency at which we seek to evaluate $\widehat{\Xi}$. The leaf ○ represents an instance of the noise $dW(u,\omega)$, and the line connecting them is the kernel $\widehat{P}_{t-u}$. The time variable u associated to the node ○ is integrated out.

Case $\tau = \Xi\Xi$. We now proceed to represent the process $\tau = \Xi\Xi$, the limit of $\Xi\Xi_n := (\Xi_n)^2 - \mathbf{c}_n$. We start with the product Ξ_n^2, writing

$$\begin{aligned}
\widehat{\Xi_n^2}(t,\omega) &= \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\le n}} \widehat{\Xi}(t,\omega_1)\widehat{\Xi}(t,\omega_2)\\
&= \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\le n}} \Big(\int_{-\infty}^t \widehat{P}_{t-u_1}(\omega_1)\,\mathrm{d}W(u_1,\omega_1)\Big)\Big(\int_{-\infty}^t \widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\Big)\\
&= \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\le n}} \Big(2\int_{-\infty}^t \Big[\int_{-\infty}^{u_1} \widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\Big]\mathrm{d}W(u_1,\omega_1)\\
&\qquad + \mathbf{1}_{\{\omega_1=-\omega_2\}}\int_{-\infty}^t \widehat{P}_{t-u}(\omega_1)\widehat{P}_{t-u}(\omega_2)\,\mathrm{d}u\Big),
\end{aligned} \tag{4.6}$$

where the last equality follows from Itô's formula. The last term on the right side vanishes for $\omega \neq 0$, because in this case the conditions $\omega_1+\omega_2=\omega$ and $\omega_1=-\omega_2$ are incompatible. For $\omega = 0$, the sum of these terms can be rewritten as

$$\sum_{|\omega_1|\le n}\int_{-\infty}^t |\widehat{P}_{t-u}(\omega_1)|^2\,\mathrm{d}u = \sum_{|\omega_1|\le n}\frac{1}{2\langle\omega_1\rangle^2}.$$

This is precisely the term $\mathbf{c}_n := \mathbb{E}\big[(\Xi_n(t))^2\big]$, which is of the order of n as n goes to infinity, and is removed in the renormalisation procedure. Below we will show that we can pass to the limit $n\to\infty$ in the first term on the right side of (4.6). We denote the limit by $\widehat{\Xi\Xi}(t,\omega)$. It is instructive to translate the expressions back into "real space", and to check that in the notation introduced in Section 3, we have

$$\widehat{\Xi\Xi}(t,\omega) = \sum_{\omega_1+\omega_2=\omega} 2\Big(\int_{-\infty}^t \Big[\int_{-\infty}^{u_1} \widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\Big]\mathrm{d}W(u_1,\omega_1)\Big) \tag{4.7}$$

$$= \int_{(\mathbb{R}\times\mathbb{T}^3)^2}\int_{\mathbb{T}^3} P_{t-u_1}(y-x_1)P_{t-u_2}(y-x_2)e^{-i2\pi\omega\cdot y}\,\mathrm{d}y\,\xi(\mathrm{d}u_1,\mathrm{d}x_1)\xi(\mathrm{d}u_2,\mathrm{d}x_2),$$

where we identify the operator P_t with its kernel, which we interpret as being null for $t \le 0$. This expression shows in particular that $\widehat{\Vee}(t,\omega)$ is an element of the homogeneous Wiener chaos $\mathcal{H}_2$ as defined in Section 3.

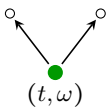
By the definition of iterated stochastic integrals, we may rewrite the identity in (4.6) as

$$\widehat{\bullet^2_n}(t,\omega) = \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\le n}} \Big(\int_{-\infty}^t \int_{-\infty}^t \widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\,\mathrm{d}W(u_1,\omega_1) + \mathbf{1}_{\{\omega_1=-\omega_2\}} \int_{-\infty}^t \widehat{P}_{t-u}(\omega_1)\widehat{P}_{t-u}(\omega_2)\,\mathrm{d}u\Big),$$

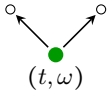
or, with our convention (4.3),

$$\widehat{\bullet^2_n}(t,\omega) = \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\le n}} \Big(\int_{\mathbb{R}^2} \widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\,\mathrm{d}W(u_1,\omega_1) + \mathbf{1}_{\{\omega_1=-\omega_2\}} \int_{-\infty}^t \widehat{P}_{t-u}(\omega_1)\widehat{P}_{t-u}(\omega_2)\,\mathrm{d}u\Big). \quad (4.8)$$

As the reader can see, this expression is already quite bulky — and much worse is to come. This motivates to introduce a graphical notation which encodes these expressions in a much more transparent way, and which we will be able to manipulate directly. At this stage, we disregard the truncation, and perform all calculations in the formal limit $n = \infty$. The expression (4.8) then becomes

$\widehat{\bullet^2}(t,\omega) =$  $+$  .

As before the root ● of the graph represents the pair (t,ω), and each of the leaves ○ represents one occurrence of the white noise and carries a pair (u_i,ω_i) itself. The kernel $\widehat{P}_t$ is represented by the arrow connecting the nodes evaluated at time $(t-u_i)$, and the arrow points towards the node whose time variable is "earlier". Then all time variables u_i except for the one t at the node are integrated out and the ω_i are summed over, subject to $\omega_1+\omega_2=\omega$. We will see below that this last rule corresponds to Kirchhoff's law that the "ingoing" variable ω must coincide with the sum of all "outgoing" ω_i's. The second graph on the right side is obtained by "contracting" the two nodes of the first graph. In this second graph, we may associate a frequency ω_1 to the left arrow, and a frequency ω_2 to the right arrow. Kirchhoff's law for the bottom node then imposes $\omega=\omega_1+\omega_2$, and for the top node, $\omega_1+\omega_2=0$, and we recover that this term is zero unless $\omega=0$. This contracted graph is removed in the renormalisation procedure, so that

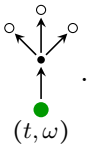
$\widehat{\Vee}(t,\omega) =$  . (4.9)

Case $\tau = \Vee\bullet$. We now discuss the next term $\Vee\bullet$, which as announced arises as the limit of $\Vee\bullet_n := I\big((\bullet_n)^3 - 3\mathbf{c}_n\bullet\big)$. As above in (4.6) (see also (3.11) with $n=3$)

we can use Itô's formula to obtain an iterated integral representation for $\widehat{\bullet^3_n}$, which takes the form

$$\widehat{\bullet^3_n}(t,\omega) = \sum_{\substack{\omega_1+\omega_2+\omega_3=\omega \\ |\omega_i|\le n}} 6 \int_{-\infty}^{t} \int_{-\infty}^{u_1} \int_{-\infty}^{u_2} \widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_2}(\omega_2)\widehat{P}_{t-u_3}(\omega_3) \, \mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_2,\omega_2)\,\mathrm{d}W(u_1,\omega_1) + 3\mathbf{c}_n \widehat{\bullet}_n(t,\omega). \tag{4.10}$$

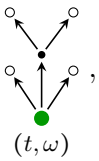
We had already seen above that $\mathbf{c}_n$ diverges as n goes to ∞, which motivates to remove the second term in the renormalisation procedure. The reason why we choose to work with the integrated object $\Ivee$ rather than $\Vee$ is that (as we will see below) although the latter can be defined as a space time distribution, it cannot be evaluated at any fixed t. (This is similar to temporal white noise which can also only be interpreted when tested against a function of time and space, but never pointwise in t). In fact, this is an instance of the well-known fact that Wick powers of order ≥ 3 over the three dimensional Gaussian free field do not exist (since the covariance function would not be integrable; see e.g. [9, Section 2.7]), and for readers familiar with this fact, it may be surprising that Wick powers up to order 4 can be constructed as space-time objects. Using a similar graphical notation to the one used for $\Vee$, we arrive at

$\widehat{\Ivee}(t,\omega) =$ 

.

Case $\tau = \Ydiagram$. At this point we want to start to think more systematically about the derivation of the diagrammatic expressions and their interpretation, and we illustrate this with the diagram $\Ydiagram$. This particular expansion follows from an iterated application of Itô's formula, see [23, Propositions 1.1.2 and 1.1.3] and [15, Section 10].

For the moment we ignore the additional complexity introduced by the resonant products $\circleddash$ in (1.7), and give a graphical representation for $\Ydiagram$ defined as in (1.7) with $\circleddash$ replaced by the usual product. To begin with, we give the graphical representation of this symbol without taking into account the renormalisation procedure, i.e. we work with the random function $I(\widehat{\bullet^2_n})\widehat{\bullet^2_n}$. This random function takes values in $\mathcal{H}_{\le 4}$, with components in $\mathcal{H}_4$, $\mathcal{H}_2$ and $\mathcal{H}_0$. The component in the highest Wiener chaos $\mathcal{H}_4$ is represented by the graph



, (4.11)

i.e. precisely the graph we use as a symbol to represent this object. This graph can be interpreted as random variable either in "real space" coordinates or in "Fourier coordinates". Both interpretations are equivalent and the former is closer in spirit to [15, Section 10] and also (4.7) above, while the latter is closer to the spirit of the present notes. Here we present both interpretations, starting with

the "real space" interpretation because it is slightly easier to explain. The Fourier interpretation then follows by turning multiplication into convolution in the space coordinates: For the "real space" interpretation we assign a space-time point to each of the vertices of this graph (e.g. $(u_1, x_1), \dots, (u_4, x_4)$ to the four leaves, (u_5, x_5) to the internal vertex, and (t, x_6) to the root), an instance of the heat kernel P evaluated at the difference of the variables associated to the adjacent vertices and the arrow pointing towards the "earlier" time variable to each of the arrows (e.g. $P_{u_5-u_1}(x_5 - x_1)$ to the upper right arrow), and multiply all of these kernels. Finally, the variable corresponding to the internal variable (u_5, x_5) is integrated out over space-time, the variables (u_i, x_i) for the leaves are integrated against the white noise ξ and we take the spatial Fourier transform with respect to the variable x_6 at the root, yielding the expression[3]

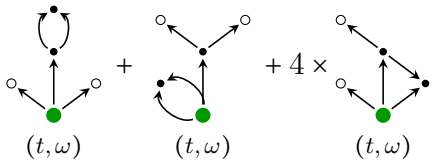
$$\int_{(\mathbb{R}\times\mathbb{T}^3)^4} \Big(\int_{\mathbb{T}^3} \mathrm{d}x_6 \int_{\mathbb{R}\times\mathbb{T}^3} \mathrm{d}u_5\, \mathrm{d}x_5 P_{u_5-u_1}(x_5 - x_1) P_{u_5-u_2}(x_5 - x_2) \\ \times P_{t-u_5}(x_6 - x_5) P_{t-u_3}(x_6 - x_3) P_{t-u_4}(x_6 - x_4) e^{-i2\pi\omega\cdot x_6} \Big) \\ \xi(\mathrm{d}u_1, \mathrm{d}x_1)\xi(\mathrm{d}u_2, \mathrm{d}x_2)\xi(\mathrm{d}u_3, \mathrm{d}x_3)\xi(\mathrm{d}u_4, \mathrm{d}x_4). \tag{4.12}$$

Translating the previous expression and interpretation into Fourier variables can be done as follows: each of the vertices is equipped with a time-frequency variable in $\mathbb{R} \times \mathbb{Z}^3$, say $(u_1, \omega_1), \dots, (u_4, \omega_4)$ for the leaves, (u_5, ω_5) for the internal vertex, and (t, ω) for the root. The formula (4.12) then becomes

$$\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3 \\ \omega_1+\omega_2=\omega_5 \\ \omega_3+\omega_4+\omega_5=\omega}} \int_{\mathbb{R}^4} \Big(\int_{\mathbb{R}} \mathrm{d}u_5 \widehat{P}_{u_5-u_1}(\omega_1) \widehat{P}_{u_5-u_2}(\omega_2) \widehat{P}_{t-u_5}(\omega_5) \widehat{P}_{t-u_3}(\omega_3) \widehat{P}_{t-u_4}(\omega_4) \Big) \\ dW(u_1, \omega_1)\, \mathrm{d}W(u_2, \omega_2)\, \mathrm{d}W(u_3, \omega_3)\, \mathrm{d}W(u_4, \omega_4), \tag{4.13}$$

that is, each arrow now corresponds to an instance of $\widehat{P}$, where the time variables stay the same as before, but the difference of the space variables is replaced by the frequency variable corresponding to the top of the arrow. The fact that the product turns into convolution under the Fourier transform is reflected in the "Kirchhoff rule" that at each internal vertex, the sum of "incoming" frequency variables equals the sum of "outgoing" frequency variables. The same rule applies at the root, with the understanding that ω is an "ingoing" frequency.

The terms in the lower order Wiener chaoses $\mathcal{H}_2$ and $\mathcal{H}_0$ arise from Itô's formula, in the same spirit to our discussion of 𝔙 above. For $\mathcal{H}_2$ we get



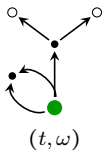
These are precisely the graphs that can be obtained by picking a pair of leaves in (4.11) and "gluing" them together. The pre-factor "4" is combinatorial and

[3]Actually, for finite n the heat kernels connecting to the leaves, i.e. $P_{u_5-u_1}$, $P_{u_5-u_2}$, P_{t-u_3} and P_{t-u_4} (but not P_{t-u_5}) should be replaced by the regularised heat kernel $(t, x) \mapsto \sum_{|\omega|\le n} \widehat{P}_t(\omega) e^{i2\pi\omega\cdot x}$. Similarly, in (4.13) and after we will leave implicit the constraint $|\omega_1|, \dots, |\omega_4| \le n$. Here and below we drop the regularisations for convenience.

corresponds to the fact that there are four different ways of picking one vertex from the "top level" and one from the "bottom level" of the graph, each of which giving rise to the same iterated stochastic integral. The interpretation for these graphs is the same as before. For instance, for the first of these graphs corresponds to the expression

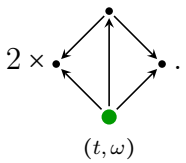
$$\begin{aligned}
&\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5\\ \omega_3+\omega_4+\omega_5=\omega\\ \omega_1+\omega_2=0}} \int_{\mathbb{R}^2}\Big(\int_{\mathbb{R}^2} \mathrm{d}u_5\,\mathrm{d}u_1\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_1}(\omega_2)\widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_3}(\omega_3)\widehat{P}_{t-u_4}(\omega_4)\Big)\\
&\qquad\qquad \mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_4,\omega_4)\\
&= \mathbf{c}_n \sum_{\substack{\omega_3,\omega_4\in\mathbb{Z}^3\\ \omega_3+\omega_4=\omega}} \int_{\mathbb{R}^2} \widehat{P}_{t-u_3}(\omega_3)\widehat{P}_{t-u_4}(\omega_4)\,\mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_4,\omega_4)\\
&= \mathbf{c}_n \widehat{\Xi}(t,\omega),
\end{aligned} \tag{4.14}$$

which arises from (4.13) by replacing the white noises $\mathrm{d}W(du_1,\omega_1)\,\mathrm{d}W(du_2,\omega_2)$ from the vertices which are "glued together" by $\delta_0(u_1-u_2)\mathbf{1}_{\omega_1=-\omega_2}$. In the same way we get the identity



$$= \mathbf{c}_n \times \widehat{I(\Xi)}(t,\omega). \tag{4.15}$$

Finally, the term in the zero-th Wiener chaos (that is, the constant) is given by the only graph that can be obtained from two contractions, namely



(4.16)

We now move on to discussing the renormalisation of these terms. As the reader can verify, working with $I(\Xi)\Xi$ instead of $I(\xi^2)\xi^2$ (i.e. performing the Wick renormalisation of the product ξ^2 as above) corresponds exactly to removing the graphs in (4.14) and (4.15). The logarithmic sub-divergence corresponding to $\mathbf{c}'_n$

arises in (4.16). Indeed, evaluating this expression yields

$$\begin{aligned}
&\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5\\ \omega_3+\omega_4+\omega_5=\omega\\ \omega_1+\omega_3=0\\ \omega_2+\omega_4=0}} \int_{\mathbb{R}^3} \mathrm{d}u_5\,\mathrm{d}u_1\,\mathrm{d}u_2\,\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_2}(\omega_2)\widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_1}(\omega_3)\widehat{P}_{t-u_2}(\omega_4)\\
=&\mathbf{1}_{\{\omega=0\}} \sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \int_{\mathbb{R}^3} \mathrm{d}u_5\,\mathrm{d}u_1\,\mathrm{d}u_2\,\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_2}(\omega_2)\\
&\qquad\qquad \times \widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{t-u_2}(-\omega_2)\\
=&\mathbf{1}_{\{\omega=0\}} \sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \int_{\mathbb{R}} \mathrm{d}u_5 \widehat{P}_{t-u_5}(\omega_5) \frac{e^{-|t-u_5|\langle\omega_1\rangle^2}}{2\langle\omega_1\rangle^2} \frac{e^{-|t-u_5|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2}\\
=&\mathbf{1}_{\{\omega=0\}} \frac{1}{4} \sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^2}\frac{1}{\langle\omega_1\rangle^2+\langle\omega_2\rangle^2+\langle\omega_5\rangle^2},
\end{aligned} \tag{4.17}$$

In the first identity above, we have used the fact that the four restrictions on the ω_i are incompatible unless $\omega = 0$. The fact that the expression vanishes for non-zero ω could also be deduced from the fact that (4.16) represents the expectation of $I(\varv_n^2)\varv_n^2$, which is constant in space by stationarity. In the second identity, we have made use of the symmetry of $\widehat{P}$ in ω and of the fact that $\int_{\mathbb{R}} \widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{t-u_1}(\omega_1)\,\mathrm{d}u_1 = \frac{e^{-|t-u_5|\langle\omega_1\rangle^2}}{2\langle\omega_1\rangle^2}$, as well as the corresponding identity for u_2.

The sum in (4.17), with cutoffs $|\omega_i| \le n$, diverges logarithmically as n tends to infinity. It is therefore necessary to remove the diagram in (4.16) in the renormalisation procedure. We arrive at the expression $I(\varv_n)\varv_n - 2\mathbf{c}'_n$,[4] which is already very close to the definition of $\varv_n$ in (1.7).

It only remains to re-introduce the resonant product $\circledcirc$ in our construction, in place of the full product. We start by briefly explaining why this is actually necessary, going back to the discussion of product estimates in Section 2. In the solution theory of (1.1), the term $\varv$ plays the role of a product of $\varUpsilon = I(\varv)$ and $\varv$. Now, as we have already discussed at length, $\varUpsilon$ is a random function of class $\mathcal{C}^{1-}$, $\varv$ is a random distribution of class $\mathcal{C}^{-1-}$, and products (more specifically the resonant part of products) are not well defined in this regularity class. (Of course, the purpose of the present article is to explain how to define these products as probabilistic limits of renormalised approximations). As we have just seen, a renormalisation procedure can be used to define the products. Yet, it will not improve the regularity of the resulting object (predicted in Table 2.1), i.e. the product $\varUpsilon\varv$ will inherit the bad regularity $\mathcal{C}^{-1-}$ from $\varv$. It is however crucial, both in Hairer's theory of regularity structures and in the theory of paracontrolled distributions, to obtain a bound which reflects the "good" regularity of $\varUpsilon$, i.e. we need to get a bound of regularity $(-1-)+(1-) = 0-$. In Hairer's theory, this is

[4] In fact, the sum represented by the diagram (4.16) does not coincide exactly with the constant $\mathbf{c}'_n$ as defined in (1.8) because the latter is defined as the expectation of the resonant product $\mathbb{E}[I(\varv_n)\circledcirc\varv_n(t)]$ while the former coincides with $\mathbb{E}[I(\varv_n)\varv_n(t)]$. However, as the reader can check, the difference between these constants remains bounded as n tends to infinity.

accomplished by working with "increments": there the fundamental object is

$$(\Upsilon(y) - \Upsilon(x))v(y),$$

which indeed behaves like $(y-x)^{0-}$ as $y \to x$. One key observation in [13] was that the same effect can be obtained by working with the resonant product $\Upsilon \ominus v$, which is of class $\mathcal{C}^{0-}$.

After this short detour, it remains to incorporate the resonant product $\ominus$ into the graphical notation. For this we recall from Section 2 that for arbitrary f, g (e.g. $\in \mathcal{C}^\infty$) we have

$$f \ominus g(x) = \sum_{|k-l|\le 1} \delta_k f(x) \delta_l g(x) = \sum_{\omega_1,\omega_2 \in \mathbb{Z}^3} e^{2i\pi(\omega_1+\omega_2)\cdot x} \widehat{f}(\omega_1)\widehat{g}(\omega_2) \sum_{|k-l|\le 1} \chi_k(\omega_1)\chi_l(\omega_2).$$

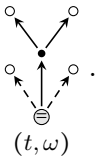
For the last sum appearing in this expression, we have

$$\sum_{|k-l|\le 1} \chi_k(\omega_1)\chi_l(\omega_2) \begin{cases} \in [0,1] & \text{for all } \omega_1, \omega_2 \\ = 0 & \text{if } (|\omega_1| > \tfrac{8}{3} \text{ or } |\omega_2| > \tfrac{8}{3}) \text{ and } \tfrac{|\omega_1|}{|\omega_2|} \notin [c, c^{-1}], \end{cases} \tag{4.18}$$

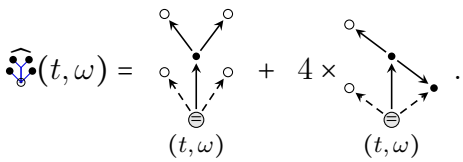
where $c = \frac{9}{64}$. Roughly speaking, and in agreement with the intuition in (2.4), this term acts as a smooth indicator function, which only selects pairs (ω_1, ω_2) for which $|\omega_1|$ and $|\omega_2|$ are close to one another on the logarithmic scale. This intuition justifies the slightly abusive notation

$$\widehat{f \ominus g}(\omega) = \sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1 \sim \omega_2}} \widehat{f}(\omega_1)\widehat{g}(\omega_2) := \sum_{\omega_1+\omega_2=\omega} \widehat{f}(\omega_1)\widehat{g}(\omega_2) \sum_{|k-l|\le 1} \chi_k(\omega_1)\chi_l(\omega_2). \tag{4.19}$$

We represent these "restricted convolutions" in the graphical notation by dotted lines. For example, when the product is replaced by $\ominus$, the first graph (4.11) becomes



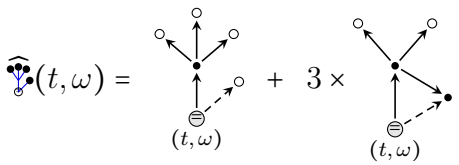
The interpretation of this diagram is the same as in (4.13), with the only exception that the additional restriction $\{\omega_3 + \omega_4 \sim \omega_5\}$ is enforced. The convention is that next to a node $\ominus$, the sum of the frequency variables corresponding to the dotted arrows is similar to the sum of frequency variables from the regular arrows. Summarising all of this discussion, we finally arrive at the graphical expression



(4.20)

Remaining two diagrams. The graphical representation/decomposition of its components in different Wiener chaoses for the remaining terms follows the

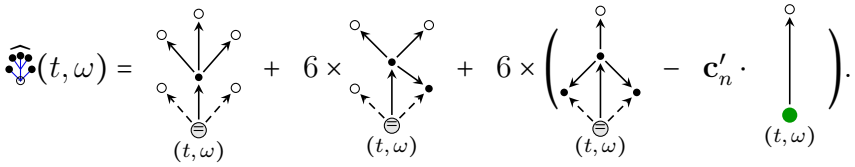
same line of reasoning, and we omit the details. For $\tau =$, we get



(4.21)

We only mention that as above in the discussion for , the Wick renormalisation (i.e. the fact that we work with $I($$)\ominus$ rather than $I($$^3)\ominus$) corresponds exactly to removing the graphs in the second Wiener chaos which arise by contracting two of the three leaves at the top of the graph. The logarithmic sub-divergence plays no role for this term.

Finally, for $\tau =$ , we have



It is worth pointing out here that for this diagram, the renormalisation with the logarithmically diverging constant $\mathbf{c}'_n$ does not result in the complete removal of a diagram, but only in its modification. In fact, when performing the renormalisation procedure for more complicated equations, it is common that the renormalisation of a graph results in the subtraction of divergent substructures, rather than the removal of the whole graph. It is rather a peculiarity that in the graphs (4.14) and (4.15) the removal of the divergent substructure amounts to removing the whole graph. See [6] for a discussion of this point in a much more general framework.

4.2. **Bounds on iterated integrals.** We now proceed to explain how to derive the bound (3.13) for the various stochastic integrals introduced in the previous subsection. The core ingredient is the isometry identity (3.5), which permits to bound the second moment of an iterated stochastic integral by the L^2 norm of the corresponding kernel.

The symbols $\tau =$ , , . As before, we first treat the symbol . For this symbol, equation (4.1) together with the standard Itô isometry yields

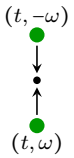
$$\mathbb{E}\big[\widehat{\bullet}(t,\omega)\widehat{\bullet}(t',-\omega)\big] = \int_{\mathbb{R}} \widehat{P}_{t-u}(\omega)\widehat{P}_{t'-u}(-\omega)\,\mathrm{d}u = \frac{e^{-|t-t'|\langle\omega\rangle^2}}{2\langle\omega\rangle^2}, \tag{4.22}$$

and in particular the bound (3.13) (for $\alpha = -\frac{1}{2}$) follows from the trivial bound $e^{-|t-t'|\langle\omega\rangle^2} \le 1$. For later use, we record the following immediate corollary of the previous bound: for any $\gamma \ge 0$,

$$\big|\mathbb{E}\big[\widehat{\bullet}(t,\omega)\widehat{\bullet}(t',-\omega)\big]\big| \lesssim \frac{1}{\langle\omega\rangle^2}\Big(\frac{1}{|t-t'|\langle\omega\rangle^2}\Big)^{\gamma}, \tag{4.23}$$

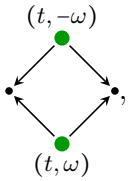
where the implicit constant depends only the choice of γ. In graphical notation, the previous calculation with $t = t'$ becomes

$$\mathbb{E}\big[|\widehat{\text{•}}(t,\omega)|^2\big] \;=\; \lesssim \langle\omega\rangle^{-2}, \tag{4.24}$$



which can be obtained from (4.5) by "doubling" the graph and by contracting the leaves $\circ$ to a vertex $\bullet$. The interpretation of the resulting graph then remains the same as in the previous section, i.e. the time variable is integrated out. This algorithm for producing the diagram is very natural: first, taking the square corresponds to the doubling of the graph; and then taking the expectation results in collapsing each pair of leaves represented by a vertex $\circ$ to a single vertex $\bullet$, for each possible pairing; this being due to the trivial covariance structure of the instances of white noise represented by the leaves $\circ$ (in other words, this being due to Itô's isometry). In the same "graphical" way, we obtain the formula

$$\mathbb{E}\big[|\widehat{\text{••}}(t,\omega)|^2\big] \;=\; 2 \;\; , \tag{4.25}$$



which should be read as

$$\begin{aligned}\mathbb{E}\big[|\widehat{\text{••}}(t,\omega)|^2\big] &= 2\sum_{\omega_1+\omega_2=\omega}\int_{\mathbb{R}^2}\Big(\widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{t-u_2}(\omega_2)\widehat{P}_{t-u_2}(-\omega_2)\Big)\,\mathrm{d}u_1\,\mathrm{d}u_2 \\ &= 2\sum_{\omega_1+\omega_2=\omega}\frac{1}{2\langle\omega_1\rangle^2}\frac{1}{2\langle\omega_2\rangle^2}.\end{aligned} \tag{4.26}$$

Of course, this expression could now be bounded directly, but we prefer a slightly more general approach which will allow us to systematise the calculations to come. The following lemma, which gives a bound on discrete convolutions, is essentially contained in [15, Lemma 10.14]. We formulate it in arbitrary space dimension d, although we are only interested in the case $d = 3$ here.

Lemma 4.1. *Let $d \geq 1$ and $\alpha, \beta \in \mathbb{R}$ satisfy*

$$\alpha + \beta > d \quad \text{and} \quad \alpha, \beta < d. \tag{4.27}$$

We have, uniformly over $\omega \in \mathbb{Z}^d$,

$$\sum_{\substack{\omega_1,\omega_2\in\mathbb{Z}^d\\ \omega_1+\omega_2=\omega}} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta}. \tag{4.28}$$

Proof. We subdivide the index set $\mathcal{A} = \{(\omega_1, \omega_2) \in (\mathbb{Z}^d)^2 : \omega_1 + \omega_2 = \omega\}$ of the summation into

$$\mathcal{A}_1 := \big\{(\omega_1, \omega_2) \in \mathcal{A}: \quad |\omega_1| \geq 2|\omega|\big\}$$
$$\mathcal{A}_2 := \big\{(\omega_1, \omega_2) \in \mathcal{A}: \quad |\omega_1| \leq \frac{1}{2}|\omega|\big\}$$
$$\mathcal{A}_3 := \big\{(\omega_1, \omega_2) \in \mathcal{A}: \quad |\omega_2| \leq \frac{1}{2}|\omega|\big\}$$
$$\mathcal{A}_4 := \mathcal{A} \setminus \Big(\bigcup_{j=1}^{3} \mathcal{A}_j\Big),$$

and bound the sums over the individual $\mathcal{A}_j$ separately. For $(\omega_1, \omega_2) \in \mathcal{A}_1$, we make use of the triangle inequality in the form $|\omega_2| = |\omega - \omega_1| \geq |\omega_1| - |\omega| \geq \frac{1}{2}|\omega_1|$ to get

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_1} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \sum_{(\omega_1,\omega_2)\in\mathcal{A}_1} \langle\omega_1\rangle^{-\alpha}\big(\frac{1}{2}\langle\omega_1\rangle\big)^{-\beta} \lesssim \sum_{|\omega_1|\geq 2|\omega|} \langle\omega_1\rangle^{-\alpha-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

where we have used the first condition in (4.27). For $\mathcal{A}_2$, we use the triangle inequality in the form $|\omega_2| = |\omega - \omega_1| \geq |\omega| - |\omega_1| \geq \frac{1}{2}|\omega|$ to get

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_2} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \big(\frac{1}{2}\langle\omega\rangle\big)^{-\beta} \sum_{|\omega_1|\leq\frac{1}{2}|\omega|} \langle\omega_1\rangle^{-\alpha} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

where this time we have used the second assumption (on α) in (4.27). Exchanging the role of ω_1 and ω_2, the same bound follows for the sum over $\mathcal{A}_3$. Finally, on $\mathcal{A}_4$ we have $|\omega_1|, |\omega_2| \geq \frac{1}{2}|\omega|$, so that

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_4} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \big(\frac{1}{2}\langle\omega\rangle\big)^{-\alpha} \sum_{|\omega_2|\geq\frac{1}{2}|\omega|} \big(\frac{1}{2}\langle\omega_2\rangle\big)^{-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

and the statement follows. □

We briefly discuss the conditions (4.27) on the exponents α, β. The first condition $\alpha + \beta > d$ is necessary to obtain a bound of the type (4.28), because without it even the convergence of the sum cannot be guaranteed. The second condition $\alpha, \beta < d$ may seem more surprising. It states that a decay beyond summability for $\langle\omega_1\rangle^{-\alpha}$ or $\langle\omega_2\rangle^{-\beta}$ does not improve the behaviour of the convolution. We will see below that this restriction corresponds exactly to the fact that a (renormalised) product of a random function f and a random distribution g cannot have better regularity than g itself. We will show below how the use of the resonant product $\circleddash$ instead of the usual product translates into a Lemma 4.2, which can be understood as a variant of Lemma 4.1 for which this restriction is removed.

Applying this Lemma to the right side of (4.26) yields

$$\mathbb{E}\big[|\widehat{\Psi}(t,\omega)|^2\big] \lesssim \frac{1}{\langle\omega\rangle},$$

which is the desired bound (3.13) with $\alpha = -1$ for this symbol. This implies that this process belongs to C^β for every $\beta < -1$.

We move on to the symbol Ψ. We first discuss why the need for the extra integration against the heat kernel arises. If we tried to work with Ψ, that is to

say, to define the limit of $\Upsilon_n^3 - 3\mathbf{c}_n \Upsilon_n$ (see (4.10)), then the same calculation as (4.25) and (4.26) would become

$$\mathbb{E}\Big[|\widehat{\Upsilon}(t,\omega)|^2\Big] = 6 \times \quad (t,\omega) \bullet \longrightarrow \bullet \longleftarrow \bullet (t,-\omega) \quad = 6 \sum_{\omega_1+\omega_2+\omega_3=\omega} \frac{1}{2\langle\omega_1\rangle^2} \frac{1}{2\langle\omega_2\rangle^2} \frac{1}{2\langle\omega_3\rangle^2},$$

but then Lemma 4.1 does not apply to this situation (applying the Lemma once for the variables ω_1 and ω_2 would yields the bound $\lesssim \sum_{\tilde\omega+\omega_3=\omega} \langle\tilde\omega\rangle^{-1} \langle\omega_3\rangle^{-2}$, but then the resulting exponents $\alpha = 1$ and $\beta = 2$ just fail the summability condition $\alpha + \beta > 3$.) We leave it for the reader to check that indeed, this sum diverges logarithmically for every ω. However, this problem can be fixed by considering different times $t \neq t'$. Then recalling (4.23) we get for any $\gamma \geq 0$

$$\begin{aligned}\mathbb{E}\big[\widehat{\Upsilon}(t,\omega)\widehat{\Upsilon}(t',-\omega)\big] &= 6 \times \quad (t,\omega) \bullet \longrightarrow \bullet \longleftarrow \bullet (t',-\omega) \\ &\lesssim \frac{1}{|t-t'|^\gamma} \sum_{\omega_1+\omega_2+\omega_3=\omega} \frac{1}{\langle\omega_1\rangle^{2+2\gamma}} \frac{1}{\langle\omega_2\rangle^2} \frac{1}{\langle\omega_3\rangle^2} \\ &\lesssim \frac{1}{|t-t'|^\gamma \langle\omega\rangle^{2\gamma}},\end{aligned}$$

which (for $\gamma < 1$) can be taken as a basis for defining Υ as a space-time distribution. We prefer, to integrate it once more against a heat kernel (because this is the way it enters the solution theory for (1.1)), yielding an object which can be evaluated at fixed time. More precisely, for every $\gamma \in (0,1)$,

$$\begin{aligned}\mathbb{E}\Big[|\widehat{\Upsilon}(t,\omega)|^2\Big] &= 6 \quad \bullet \longrightarrow \bullet_u \bullet \longleftarrow \bullet_{u'} \bullet \\ &\lesssim \int_{\mathbb{R}^2} \widehat{P}_{t-u}(\omega) \widehat{P}_{t-u'}(\omega) \frac{1}{|u-u'|^\gamma \langle\omega\rangle^{2\gamma}} \,\mathrm{d}u \,\mathrm{d}u' \lesssim \frac{1}{\langle\omega\rangle^4}, \qquad (4.29)\end{aligned}$$

which proves that this symbol satisfies (3.13) for $\alpha = \frac{1}{2}$.

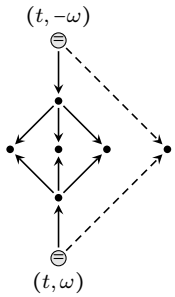
Case $\tau = \Upsilon\bullet$. We now turn to the case $\tau = \Upsilon\bullet$. Recall that its decomposition into components in homogeneous Wiener chaoses was given in (4.21). We write

$$\widehat{\Upsilon\bullet}(t,\omega) = \quad \circ \quad \circ \quad \circ \quad (t,\omega) \quad + \quad 3 \times \quad \circ \quad \circ \quad \circ \quad (t,\omega) \quad =: \widehat{\Upsilon\bullet}^{(4)}(t,\omega) + \widehat{\Upsilon\bullet}^{(2)}(t,\omega). \qquad (4.30)$$

It is then clear that $\widehat{\Upsilon\bullet} \in \mathcal{H}_n$, and as a consequence of the orthogonality of the $\mathcal{H}_n$'s in L^2, we have

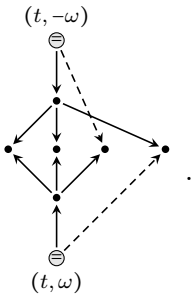
$$\mathbb{E}\Big[|\widehat{\Upsilon\bullet}(t,\omega)|^2\Big] = \mathbb{E}\Big[|\widehat{\Upsilon\bullet}^{(4)}(t,\omega)|^2\Big] + \mathbb{E}\Big[|\widehat{\Upsilon\bullet}^{(2)}(t,\omega)|^2\Big]. \qquad (4.31)$$

In our graphical notation, the first term above can be bounded by

$\mathbb{E}\Big[|$ $(t,\omega)|^2\Big] \lesssim$ 

. (4.32)

There is a slightly subtle point worth underlying here: standard Gaussian calculus (the Wick formula) yields an explicit identity for the quantity on the left hand side of this expression, in terms of contractions of all of the different leaves on the original diagram representing . This formula includes additional graphs such as



.

However, using (3.6) to bound the term on the left side of (4.32) greatly simplifies the ensuing argument, as opposed to relying on the exact formula involving the asymmetric trees. (This idea was first used in this context in [15, Section 10].)

Going back to bounding (4.32), we use (4.29) on the "left part" and (4.24) on the "right part" of the graph to obtain the bound

$$\mathbb{E}\Big[|\,\widehat{\Psi}^{(4)}(t,\omega)|^2\Big] \lesssim \sum_{\substack{\omega_1+\omega_2=\omega\\ \omega_1\sim\omega_2}} \frac{1}{\langle\omega_1\rangle^4}\frac{1}{\langle\omega_2\rangle^2}. \tag{4.33}$$

Lemma 4.1 on the decay of convolutions is not enough to bound the remaining sum. Indeed, this is precisely a case as discussed below Lemma 4.1, where the second condition in (4.27) fails (here, $4 \geq 3$). Of course, the same estimate could be used by "forgetting" some of the good decay of $\langle\omega_1\rangle^{-4}$, and replacing it by $\langle\omega_1\rangle^{-(3-)}$, but this would only yield a bound of order $\langle\omega\rangle^{-(2-)}$ corresponding to a regularity of index $-\frac{1}{2}-$ instead of $0-$. The following lemma shows that the additional condition $\omega_1 \sim \omega_2$, which arises from our use of the resonant product $\ominus$ in the definition of the diagram, resolves this problem.

Lemma 4.2. *Let $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha + \beta > d$. We have, uniformly over $\omega \in \mathbb{Z}^d$,*

$$\sum_{\substack{\omega_1+\omega_2=\omega\\ \omega_1\sim\omega_2}} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta}.$$

Proof. Recall (4.19) and the definition of the "smooth cut-off" (4.18). For small $|\omega|$ (say, $|\omega| \leq \frac{16}{3}$) there is only a bounded number of admissible ω_1, ω_2 with

$\omega_1+\omega_2=\omega$ for which $\sum_{|k-l|\le 1}\chi_k(\omega_1)\chi_l(\omega_2)\neq 0$. Hence, for such ω, we have

$$\sum_{\substack{\omega_1+\omega_2=\omega\\ \omega_1\sim\omega_2}}\langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta}\lesssim 1.$$

We can thus now assume that $|\omega|>\frac{16}{3}$. For such ω, the conditions $\omega_1+\omega_2=\omega$ and $\omega_1\sim\omega_2$ enforce that $\frac{|\omega_1|}{|\omega_2|}\in[c,c^{-1}]$ for $c=\frac{9}{64}$, by (4.18). Hence, on the one hand, we have

$$|\omega|\le|\omega_1|+|\omega_2|\le|\omega_1|+c^{-1}|\omega_1|,$$

that is, $|\omega_1|\ge\frac{1}{1+c^{-1}}|\omega|$, and on the other hand, $|\omega_2|\ge c|\omega_1|$. These considerations allow us to write

$$\sum_{\substack{\omega_1+\omega_2=\omega\\ \omega_1\sim\omega_2}}\langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta}\lesssim\sum_{|\omega_1|\ge\frac{1}{1+c^{-1}}|\omega|}\langle\omega_1\rangle^{-\alpha-\beta}\lesssim\langle\omega\rangle^{d-\alpha-\beta},$$

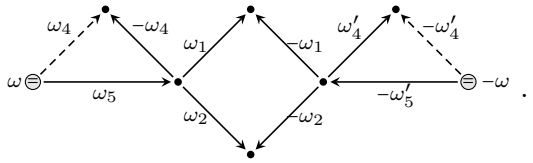
as desired. □

Applying this Lemma to (4.33) immediately yields

$$\mathbb{E}\Big[\big|\hat{\Xi}^{(4)}(t,\omega)\big|^2\Big]\lesssim\langle\omega\rangle^{-3}.$$

We now turn to the variance of $\hat{\Xi}^{(2)}$, which has the expression

$$\mathbb{E}\Big[\big|\hat{\Xi}^{(2)}(t,\omega)\big|^2\Big]\lesssim$$



.

For later reference, we have labelled all the edges with their frequency variables (but dropped the time variables) in the diagram above. In addition to the identities already implicit in this diagram, Kirchhoff's law enforces that we have to sum over the indices $\omega_1,\omega_2,\omega_4,\omega_5,\omega_4',\omega_5'$ satisfying

$$\omega_4+\omega_5=\omega,\qquad\omega_4'+\omega_5'=\omega,\qquad\omega_1+\omega_2=\omega,$$

so that we ultimately have to sum over three free variables. The fact that we have a resonant product $\ominus$ at the roots yields the additional constraints $\omega_4\sim\omega_5$ and $\omega_4'\sim\omega_5'$, but we will not rely on these additional constraints to bound this diagram. With this notation in place, we proceed to analyse each part of this diagram separately. The inner square corresponds to the integral

$$\sum_{\omega_1+\omega_2=\omega}\Big(\int_{\mathbb{R}}P_{u-u_1}(\omega_1)P_{u'-u_1}(-\omega_1)\,\mathrm{d}u_1\Big)\Big(\int_{\mathbb{R}}P_{u-u_2}(\omega_2)P_{u'-u_1}(-\omega_2)\,\mathrm{d}u_2\Big)$$
$$\lesssim\sum_{\omega_1+\omega_2=\omega}\frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^2}\lesssim\frac{1}{\langle\omega\rangle},$$

where we have used Lemma 4.1 in the last inequality. Note that this bound is slightly sub-optimal, because we have not used the fact that the time variables u and u' corresponding to the nodes at the left and right corners of the square are different. In principle, this would yield extra factors $\exp(-|u-u'|\langle\omega_i\rangle)$ for each

term, but we simply bound these factors by 1. Similarly, we get for the left-most triangle

$$\sum_{\omega_4+\omega_5=\omega} \int P_{t-u}(\omega_5)\Big(\int_{\mathbb{R}} P_{t-u_4}(\omega_4) P_{u-u_4}(-\omega_4)\,\mathrm{d}u_4\Big)\,\mathrm{d}u$$
$$\lesssim \sum_{\omega_4+\omega_5=\omega} \int P_{t-u}(\omega_5) \frac{1}{\langle\omega_4\rangle^2}\,\mathrm{d}u \lesssim \sum_{\omega_4+\omega_5=\omega} \frac{1}{\langle\omega_4\rangle^2}\frac{1}{\langle\omega_5\rangle^2} \lesssim \frac{1}{\langle\omega\rangle},$$

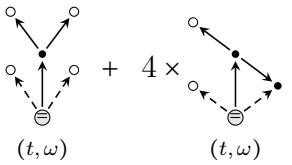
where again we have used Lemma 4.1. The right-most triangle in the diagram is bounded in the same way, resulting in the final bound

$$\mathbb{E}\Big[\big|\Xi^{(2)}(t,\omega)\big|^2\Big] \lesssim \langle\omega\rangle^{-3},$$

as desired.

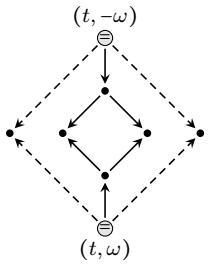
Case $\tau =$ Ξ. For this symbol, we recall from (4.20) the Wiener chaos decomposition

$$\Xi(t,\omega) = \quad + \quad 4\times \quad =: \Xi^{(4)}(t,\omega) + \Xi^{(2)}(t,\omega)\,.$$



We start with $\Xi^{(4)}$. Similarly to the case for Ξ, we have the bound

$$\mathbb{E}\Big[\big|\Xi^{(4)}(t,\omega)\big|^2\Big] \lesssim \quad . \tag{4.34}$$



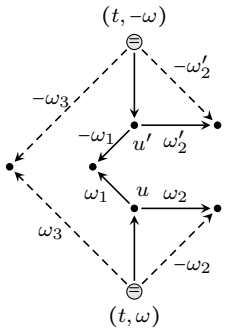
After all our preparation, this diagram poses no additional difficulty. First, the same calculation as in (4.29) allows to bound the "inner part" of the diagram by $\langle\omega_2\rangle^{-5}$, and the integrals corresponding to the "outer parts" can be bounded by $\langle\omega_1\rangle^{-2}$ and $\langle\omega_3\rangle^{-2}$ immediately, yielding the bound

$$\mathbb{E}\Big[\big|\Xi^{(4)}(t,\omega)\big|^2\Big] \lesssim \sum_{\substack{\omega_1+\omega_2+\omega_3=\omega\\ \omega_1+\omega_3\sim\omega_2}} \frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^5}\frac{1}{\langle\omega_3\rangle^2} \lesssim \sum_{\substack{\tilde\omega+\omega_3=\omega\\ \tilde\omega\sim\omega_2}} \frac{1}{\langle\tilde\omega\rangle}\frac{1}{\langle\omega_2\rangle^5} \lesssim \frac{1}{\langle\omega\rangle^3}, \tag{4.35}$$

where we have used Lemma 4.1 in the first and Lemma 4.2 in the second inequality.

We now turn to the bound for $\widehat{\Xi}^{(2)}$, for which we have

$$\mathbb{E}\Big[|\widehat{\Xi}^{(2)}(t,\omega)|^2\Big] \lesssim$$



. (4.36)

This graph is more complicated, and requires more careful treatment. We integrate first the innermost time variable, with two arrows pointing to it labelled ω_1 and $-\omega_1$ respectively. We bound this contribution by $\langle\omega_1\rangle^{-2}$, uniformly over u and u', as in (4.24). We are more careful with the integration of the time variable with incoming arrows labelled ω_2 and $-\omega_2$, and evaluate its contribution to be

$$\frac{e^{-|t-u|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2},$$

as in (4.22). We next compute the contribution of the triangle in the lower part of the diagram by integrating over u:

$$\int_{\mathbb{R}} \widehat{P}_{t-u}(\omega_1+\omega_2)\,\frac{e^{-|t-u|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2}\,\mathrm{d}u = \frac{1}{2\langle\omega_2\rangle^2}\,\frac{1}{\langle\omega_2\rangle^2+\langle\omega_1+\omega_2\rangle^2}.$$

A similar calculation applies to the upper part of the diagram, and we therefore obtain the bound

$$\mathbb{E}\Big[|\widehat{\Xi}^{(2)}(t,\omega)|^2\Big] \lesssim \sum\Big(\langle\omega_1\rangle\langle\omega_2\rangle\langle\omega_2'\rangle\langle\omega_3\rangle\big(\langle\omega_2\rangle+\langle\omega_1+\omega_2\rangle\big)\big(\langle\omega_2'\rangle+\langle\omega_2'-\omega_1\rangle\big)\Big)^{-2}, \quad (4.37)$$

where the sum is over all $(\omega_1,\omega_2,\omega_2',\omega_3)$ satisfying

$$\omega_1+\omega_3=\omega, \quad \omega_1+\omega_2\sim\omega_3-\omega_2, \quad -\omega_1+\omega_2'\sim-\omega_3-\omega_2'. \quad (4.38)$$

The first requirement comes from the "Kirchhoff" law in the bottom node, while the other two constraints come from the paraproduct in the bottom and uppermost nodes.

We proceed to estimate this sum. Note first that the first two conditions in (4.38) imply that

$$\omega_1+\omega_2\sim\omega-(\omega_1+\omega_2).$$

By (4.18), if $|\omega_1+\omega_2|>\frac{8}{3}$, we deduce that

$$|\omega_1+\omega_2|\geq c\,(|\omega|-|\omega_1+\omega_2|),$$

where $c=\frac{9}{64}$, and therefore that

$$|\omega_1+\omega_2|\geq\frac{c}{1+c}|\omega|.$$

After reducing the constant $c > 0$ as necessary, if follows that in every case, the first two conditions in (4.38) imply that

$$\langle \omega_1 + \omega_2 \rangle \geq 2c\langle \omega \rangle. \tag{4.39}$$

(The factor of 2 is of course a matter of convenience only.) The same argument also shows that under the conditions in (4.38), we have

$$\langle \omega_1 - \omega_2' \rangle \geq 2c\langle \omega \rangle. \tag{4.40}$$

Define the sets of indices $E_1(\omega)$ and $E_2(\omega)$ by

$$E_1(\omega) := \{(\omega_1, \omega_2, \omega_2', \omega_3) \ : \ (4.38) \text{ holds and } (|\omega_1| < c|\omega| \ \text{ or } \ |\omega_3| < c|\omega|)\},$$
$$E_2(\omega) := \{(\omega_1, \omega_2, \omega_2', \omega_3) \ : \ (4.38) \text{ holds and } |\omega_1| \geq c|\omega| \text{ and } |\omega_3| \geq c|\omega|\}.$$

We first estimate the sum on the right side of (4.37) over the set $E_1(\omega)$. By (4.39), (4.40) and the constraint $\omega_3 - \omega_2 \sim \omega_1 + \omega_2$, for variables in the set $E_1(\omega)$, we must have

$$\langle \omega_2 \rangle \geq \widetilde{c}\langle \omega \rangle \quad \text{ and } \quad \langle \omega_2' \rangle \geq \widetilde{c}\langle \omega \rangle,$$

for some $\widetilde{c} > 0$. Thus, using the simple bounds

$$\langle \omega_2 \rangle + \langle \omega_1 + \omega_2 \rangle \geq \langle \omega_2 \rangle, \qquad \langle \omega_2' \rangle + \langle \omega_2' - \omega_1 \rangle \geq \langle \omega_2' \rangle,$$

and summing over ω_1 and $\omega_3 = \omega - \omega_1$ using Lemma 4.1, we arrive at

$$\sum_{E_1(\omega)} \cdots \lesssim \langle \omega \rangle^{-1} \sum_{|\omega_2|, |\omega_2'| \gtrsim |\omega|} \Big(\langle \omega_2 \rangle^{-4} \cdot \langle \omega_2' \rangle^{-4} \Big) \lesssim \langle \omega \rangle^{-3}, \tag{4.41}$$

where the left side above stands for the sum on the right side of (4.37) restricted to the index set $E_1(\omega)$. As for the the index set $E_2(\omega)$, we start by summing over ω_2 to get

$$\sum_{\omega_2} \langle \omega_2 \rangle^{-2} \big(\langle \omega_2 \rangle + \langle \omega_1 + \omega_2 \rangle \big)^{-2} \lesssim \langle \omega_1 \rangle^{-2} \sum_{|\omega_2| \leq \frac{|\omega_1|}{2}} \langle \omega_2 \rangle^{-2} + \sum_{|\omega_2| \geq \frac{|\omega_1|}{2}} \langle \omega_2 \rangle^{-4} \lesssim \langle \omega_1 \rangle^{-1},$$

and similarly we have

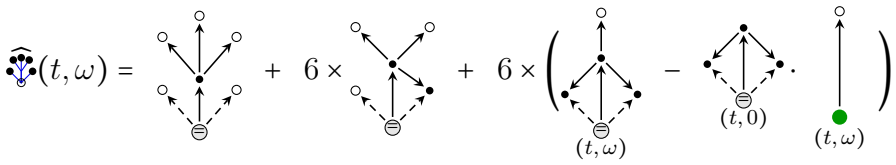
$$\sum_{\omega_2'} \langle \omega_2' \rangle^{-2} \big(\langle \omega_2' \rangle + \langle \omega_2' - \omega_1 \rangle \big)^{-2} \lesssim \langle \omega_1 \rangle^{-1}.$$

We deduce that

$$\sum_{E_2(\omega)} \cdots \lesssim \sum_{\substack{\omega_1 + \omega_3 = \omega \\ |\omega_1|, |\omega_3| \geq \frac{|\omega|}{16}}} \langle \omega_1 \rangle^{-4} \langle \omega_3 \rangle^{-2} \lesssim \langle \omega \rangle^{-2} \sum_{|\omega_1| \geq \frac{|\omega|}{16}} \langle \omega_1 \rangle^{-4} \lesssim \langle \omega \rangle^{-3},$$

with the same notational convention as in (4.41). Combining this with (4.37) and (4.41) gives the desired bound for .

Case $\tau =$. This is the last diagram. We have the Wiener chaos decomposition

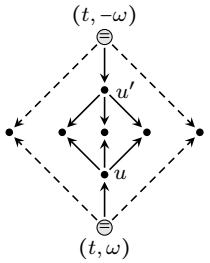


$$=: \widehat{\bullet}^{(5)}(t,\omega) + \widehat{\bullet}^{(3)}(t,\omega) + \widehat{\bullet}^{(1)}(t,\omega),$$

where the last term subtracted in the parenthesis above corresponds to the renormalisation $\mathbf{c}'_n$ in (1.7) (in the limit as $n \to +\infty$). As we will see, neither term in the parenthesis above makes sense separately: they both represent some divergent object, but their difference converges to a well-defined limit as the Fourier mode cut-off n goes to infinity. In addition, this limit can be characterised explicitly without referring to a limiting procedure, so this justifies the notation $\widehat{\bullet}^{(1)}(t,\omega)$.

We start with $\widehat{\bullet}^{(5)}$. Proceeding as in (4.34)-(4.35), using Lemma 4.1 and Lemma 4.2, we immediately have

$$\mathbb{E}\Big[|\widehat{\bullet}^{(5)}(t,\omega)|^2\Big] \lesssim$$

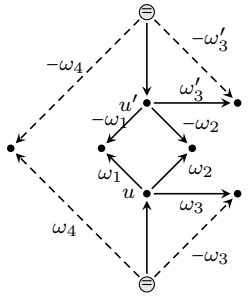


$$\lesssim \sum_{\substack{\omega_1+\omega_2+\omega_3=\omega \\ \omega_1+\omega_3 \sim \omega_2}} \frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^4}\frac{1}{\langle\omega_3\rangle^2} \lesssim \frac{1}{\langle\omega\rangle^2},$$

where the term $\frac{1}{\langle\omega_2\rangle^4}$ comes from the previous bound for the tree .

We now turn to the component in the third Wiener chaos, whose second moment is bounded by the graph

$$\mathbb{E}\Big[|\widehat{\bullet}^{(3)}(t,\omega)|^2\Big] \lesssim$$



.

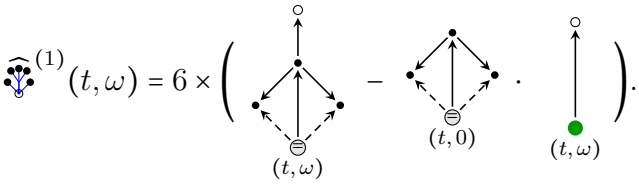
The bound for this graph is similar to the one in (4.36), and one can proceed essentially in the same way to get

$$\mathbb{E}\Big[|\widehat{\bullet}^{(3)}(t,\omega)|^2\Big] \lesssim \langle\omega\rangle^{-2},$$

which is the desired bound.

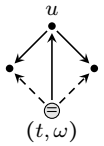
We now turn to the last term

$$\widehat{\Xi}^{(1)}(t,\omega) = 6 \times \Bigg(\quad - \quad \cdot \quad \Bigg).$$



As mentioned before, the two terms in the parenthesis are both ill-defined, but their difference is well-defined as the limit when the regularisation parameter n tends to infinity. To see this, we introduce a notation for the "lower square" which both of the expressions have in common, namely

$$K_{t-u}(\omega) := \quad = \sum_{\omega_1 \sim \omega_2} \widehat{P}_{t-u}(\omega+\omega_1+\omega_2)\Big(\int_{-\infty}^{u} \widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{u-u_1}(\omega_1)\mathrm{d}u_1 \Big)$$



$$\times \Big(\int_{-\infty}^{u} \widehat{P}_{t-u_2}(-\omega_2)\widehat{P}_{u-u_1}(\omega_2)\mathrm{d}u_2 \Big)$$

$$= \sum_{\omega_1 \sim \omega_2} \frac{e^{-(t-u)(\langle\omega_1\rangle^2+\langle\omega_2\rangle^2+\langle\omega+\omega_1+\omega_2\rangle^2)}}{4\langle\omega_1\rangle^2\langle\omega_2\rangle^2}.$$

Note that the divergent constant $\mathbf{c}'$ coincides with $\int_{-\infty}^{t} K_{t-u}(0)\,\mathrm{d}u$.[5] The kernel K_{t-u} is clearly well defined and controlled uniformly over the regularisation for any fixed $t-u>0$. We use this notation to represent $\widehat{\Xi}^{(1)}(t,\omega)$ as

$$\widehat{\Xi}^{(1)}(t,\omega) = \int_{-\infty}^{t} K_{t-u}(\omega)\big(\widehat{\Psi}(u,\omega) - \widehat{\Psi}(t,\omega)\big)\,\mathrm{d}u, \tag{4.42}$$

so that

$$\mathbb{E}\Big[|\widehat{\Xi}^{(1)}(t,\omega)|^2\Big] = \int_{-\infty}^{t}\int_{-\infty}^{t} K_{t-u}(\omega)K_{t-u'}(\omega)$$
$$\times \mathbb{E}\big[\big(\widehat{\Psi}(u,\omega) - \widehat{\Psi}(t,\omega)\big)\big(\widehat{\Psi}(u',-\omega) - \widehat{\Psi}(t,-\omega)\big)\big]\,\mathrm{d}u\,\mathrm{d}u'.$$

Now, since by (4.22) we have for any $\lambda>0$

$$\mathbb{E}\big[|\widehat{\Psi}(u,\omega) - \widehat{\Psi}(t,\omega)|^2\big]^{\frac{1}{2}} \lesssim (t-u)^{\lambda}\langle\omega\rangle^{-1+2\lambda},$$

an application of the Cauchy-Schwarz inequality yields

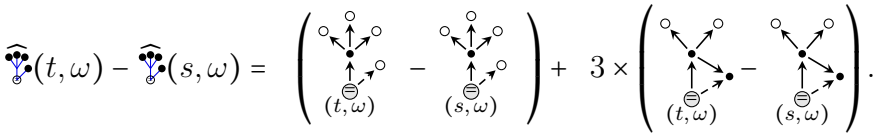
$$\mathbb{E}\Big[|\widehat{\Xi}^{(1)}(t,\omega)|^2\Big] \lesssim \langle\omega\rangle^{-2+4\lambda}\Big(\int_{-\infty}^{t} (t-u)^{\lambda}K_{t-u}(\omega)\mathrm{d}u \Big)^2 \lesssim \langle\omega\rangle^{-2},$$

where we have used the fact that for any (small) strictly positive λ, the integral can be bounded uniformly over the regularisation by $\langle\omega\rangle^{-2\lambda}$. This latter point can be checked as follows: for $\lambda=0$ and without the condition $\omega_1\sim\omega_2$, we had already calculated the integral in (4.17); the factor $(t-u)^{\lambda}$ makes an extra power $(\langle\omega_1\rangle^2+\langle\omega_2\rangle^2+\langle\omega+\omega_1+\omega_2\rangle^2)^{-\lambda}$ appear in the sum, which permits to invoke Lemma 4.1 and conclude. This completes the bound for $\tau = \Xi$.

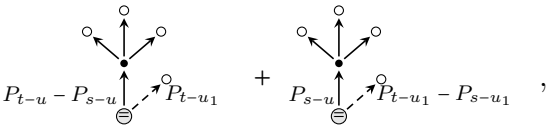
[5] This is slightly formal - here $\mathbf{c}'$ should denote the limit of $\mathbf{c}'_n$ as $n\to\infty$, which is infinite as discussed before. We are implicitly assuming that a regularisation is present, although we do not capture it in the notation.

5. Bounds for time differences

We finally discuss briefly how the reasoning in Section 4 should be modified to establish the bound (3.15) on the time differences $\mathbb{E}|\widehat{\tau}(t,\omega)-\widehat{\tau}(s,\omega)|^2$. We illustrate the (simple) modification necessary for the graph $\tau =$ . First, recall the Wiener chaos decomposition (4.30) for this graph, which yields the following expression for its time-differences:



The differences of graphs can be bounded separately. We only discuss the first difference here. We can rewrite this difference as



where we have made explicit which kernels are associated to the lower edges in the graph. (A more systematic treatment would suggest the use of new graphical notation for these!) The variance of each of these terms can then be bounded, by "glueing" two copies of each graph together, as in (4.32). We then use the elementary bounds

$$\int_{-\infty}^{t} |\widehat{P}_{t-u}(\omega)-\widehat{P}_{s-u}(\omega)|\,\mathrm{d}u \lesssim \frac{1}{\langle\omega\rangle^2}\big(1\wedge|t-s|\langle\omega\rangle^2\big),$$
$$\int_{-\infty}^{t} (\widehat{P}_{t-u_1}(\omega)-\widehat{P}_{s-u_1}(\omega))^2\,\mathrm{d}u_1 \lesssim \frac{1}{\langle\omega\rangle^2}\big(1\wedge|t-s|\langle\omega\rangle^2\big).$$

By interpolation, the right side can be replaced by $\lesssim \langle\omega\rangle^{-2+2\lambda}|t-s|^\lambda$, for any $\lambda\in[0,1]$. In other words, an extra factor $|t-s|^\lambda$ can be obtained by sacrificing a bit of the decay of the integral in ω. Then all of the arguments based on convolutions can be performed exactly as before, only with a slightly worse factor of ω at one place. We do not go through the details here, but leave it to the interested reader to check that this does not change the arguments in Section 4 in any significant way.

Appendix A. Alternative proof of Nelson's estimate

We now give a second proof of Proposition 3.3, based on the following logarithmic Sobolev inequality (see [4, Section 1.6] and [23, Sections 1.1 and 1.5]).

Lemma A.1 (log-Sobolev inequality)**.** *Let μ be a Gaussian measure and $X\in W^{1,2}(\mu)$. We have*

$$\mathbb{E}(|X|^2\log|X|) \le \mathbb{E}|\mathbf{D}X|^2 + \frac{1}{2}\mathbb{E}|X|^2\log(\mathbb{E}|X|^2), \tag{A.1}$$

where $\mathbf{D}$ is the Malliavin derivative, and $\mathbb{E}$ is the expectation taken with respect to μ.

Now, let T_t be the Ornstein-Uhlenbeck semigroup defined by

$$T_t X = \sum_{n=0}^{+\infty} e^{-nt} X_n, \tag{A.2}$$

where X_n is the component of X in $\mathcal{H}_n$. The Ornstein-Uhlenbeck semigroup is closely related to the Malliavin derivative, because it determines the quadratic form associated with the infinitesimal generator L of T_t. More precisely, for sufficiently nice random variables X, Y, we have

$$\partial_t \mathbb{E}[(T_t X)Y] = \mathbb{E}[(LX)Y] = -\mathbb{E}[\langle \mathbf{D}X, \mathbf{D}Y\rangle]. \tag{A.3}$$

See [23, Section 1.4] for a more detailed discussion of these objects. The main use of the logarithmic Sobolev inequality will be to show the following hypercontractivity estimate.

Proposition A.2 (Hypercontractivity)**.** *Let T_t be the Ornstein-Uhlenbeck semigroup. We have*

$$\big(\mathbb{E}|T_t X|^q\big)^{\frac{1}{q}} \le \big(\mathbb{E}|X|^p\big)^{\frac{1}{p}}, \tag{A.4}$$

for all $p \ge 2$ and $q = 1 + (p-1)e^{2t}$.

Second proof of Proposition 3.3. If $X \in \mathcal{H}_n$, then $T_t X = e^{-nt} X$, and we can see that Proposition 3.3 is an immediate consequence of (A.2). It then remains to prove Proposition A.2. We can assume $X \ge 0$ without loss of generality.

Fix $p \ge 2$. Let $q(t) = 1 + (p-1)e^{2t}$, and let

$$F(t) = \mathbb{E}|T_t X|^{q(t)}, \qquad G(t) = F(t)^{\frac{1}{q(t)}}.$$

Our aim is to show that $G'(t) \le 0$ for all $t > 0$, and (A.4) will follow. In fact, we have

$$G'(t) = G(t)\bigg[-\frac{q'(t)}{q^2(t)} \log F(t) + \frac{F'(t)}{q(t)F(t)} \bigg].$$

Since $q'(t) \ge 0$, it suffices to show that

$$-\frac{1}{q(t)} F(t) \log F(t) + \frac{F'(t)}{q'(t)} \le 0. \tag{A.5}$$

Noting that

$$F'(t) = \mathbb{E}\bigg[(T_t X)^{q(t)} \bigg(q'(t) \log(T_t X) + q(t) \frac{L T_t X}{T_t X} \bigg) \bigg],$$

we see that (A.5) is equivalent to

$$-\frac{1}{q(t)} F(t) \log F(t) + \mathbb{E}\Big[(T_t X)^{q(t)} \log(T_t X) \Big] + \frac{q(t)}{q'(t)} \mathbb{E}\Big[(T_t X)^{q(t)-1} (L T_t X) \Big]. \tag{A.6}$$

Applying the log-Sobolev inequality to the random variable $(T_t X)^{\frac{q(t)}{2}}$ and using the integration by parts formula (A.3) which in the current context becomes

$$\mathbb{E}\langle \mathbf{D}Y, \mathbf{D}Z\rangle = -\mathbb{E}\big(Y(LZ)\big),$$

we see that (A.6) follows immediately. □

Acknowledgments. We are grateful to the very careful referees for their detailed and constructive criticisms, which led to many improvements over the whole paper. We were particularly impressed that one of the referees' reports turned out to be longer than our paper itself! JCM is partially supported by the ANR Grant LSD (ANR-15-CE40-0020-03). HW is supported by the Royal Society through the University Research Fellowship UF140187. WX is supported by the EPSRC through the fellowship EP/N021568/1.

References

[1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2011.

[2] L. Bertini, E. Presutti, B. Rüdiger, and E. Saada. Dynamical fluctuations at the critical point: convergence to a nonlinear stochastic PDE. *Teor. Veroyatnost. i Primenen.*, 38(4):689–741, 1993.

[3] M. Biskup. Reflection positivity and phase transitions in lattice spin models. In *Methods of contemporary mathematical statistical physics*, volume 1970 of *Lecture Notes in Math.*, pages 1–86. Springer, Berlin, 2009.

[4] V. Bogachev. *Gaussian Measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, first edition, 1998.

[5] R. Catellier and K. Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *Preprint*, arXiv:1310.6869, 2013.

[6] A. Chandra and M. Hairer. An analytic BPHZ theorem for regularity structures. *Preprint*, arXiv:1612.08138, 2016.

[7] A. Chandra and H. Weber. Stochastic PDEs, regularity structures, and interacting particle systems. *Preprint*, arXiv:1508.03616, 2015.

[8] G. Da Prato and A. Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.*, 31(4):1900–1916, 2003.

[9] W. E, A. Jentzen, and H. Shen. Renormalized powers of Ornstein-Uhlenbeck processes and well-posedness of stochastic Ginzburg-Landau equations. *Nonlinear Anal.*, 142:152–193, 2016.

[10] J. Fritz and B. Rüdiger. Time dependent critical fluctuations of a one-dimensional local mean field model. *Probab. Theory Related Fields*, 103(3):381–407, 1995.

[11] P. K. Friz and M. Hairer. *A course on rough paths, with an introduction to regularity structures*. Universitext. Springer, 2014.

[12] J. Glimm and A. Jaffe. *Quantum physics*. Springer-Verlag, New York, second edition, 1987.

[13] M. Gubinelli, P. Imkeller, and N. Perkowski. Paraproducts, rough paths and controlled distributions. *Preprint*, arXiv:1210.2684, 2012.

[14] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178:559–664, 2013.

[15] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[16] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ. *Preprint*, arXiv:1512.07845, 2015.

[17] M. Hairer, M. Ryser, and H. Weber. Triviality of the $2d$ stochastic Allen-Cahn equation. *Electr. Journ. Probab.*, 17(39):1–14, 2012.

[18] H.-H. Kuo. *Introduction to stochastic integration*. Universitext. Springer, New York, 2006.

[19] A. Kupiainen. Renormalization group and stochastic PDE's. *Preprint*, arXiv:1410.3094, 2014.

[20] J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4 model in the plane. *Preprint*, arXiv:1501.06191, 2015.

[21] J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4_3 model on the torus. *Preprint*, arXiv:1601.01234, 2016.

[22] J.-C. Mourrat and H. Weber. Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2. *Comm. Pure Appl. Math.*, in press, arXiv:1410.1179.

[23] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
[24] K. Osterwalder and R. Schrader. Axioms for Euclidean Green's functions. *Comm. Math. Phys.*, 31:83–112, 1973.
[25] K. Osterwalder and R. Schrader. Axioms for Euclidean Green's functions. II. *Comm. Math. Phys.*, 42:281–305, 1975. With an appendix by Stephen Summers.
[26] G. Parisi and Y. S. Wu. Perturbation theory without gauge fixing. *Sci. Sinica*, 24(4):483–496, 1981.
[27] M. Polyak. Feynman diagrams for pedestrians and mathematicians. In *Graphs and patterns in mathematics and theoretical physics*, volume 73 of *Proc. Sympos. Pure Math.*, pages 15–42. Amer. Math. Soc., Providence, RI, 2005.
[28] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 1994.

(Jean-Christophe Mourrat) Ecole normale supérieure de Lyon, CNRS, Lyon, France
E-mail address: jean-christophe.mourrat@ens-lyon.fr

(Hendrik Weber) University of Warwick, Coventry, United Kingdom
E-mail address: hendrik.weber@warwick.ac.uk

(Weijun Xu) University of Warwick, Coventry, United Kingdom
E-mail address: weijun.xu@warwick.ac.uk