



# Sharp Convergence of Nonlinear Functionals of a Class of Gaussian Random Fields

Weijun Xu<sup>1,2</sup>

Received: 7 July 2018 / Revised: 9 October 2018 / Accepted: 24 October 2018 /

Published online: 10 November 2018

© The Author(s) 2018

## Abstract

We present a self-contained proof of a uniform bound on multi-point correlations of trigonometric functions of a class of Gaussian random fields. It corresponds to a special case of the general situation considered in Hairer and Xu (large-scale limit of interface fluctuation models. ArXiv e-prints [arXiv:1802.08192](https://arxiv.org/abs/1802.08192), 2018), but with improved estimates. As a consequence, we establish convergence of a class of Gaussian fields composite with more general functions. These bounds and convergences are useful ingredients to establish weak universalities of several singular stochastic PDEs.

**Keywords** Multi-point correlation function · Trigonometric polynomial · Gaussian random fields

**Mathematics Subject Classification** 60H

## 1 Introduction

### 1.1 Motivation from Weak Universalities

The study of singular stochastic PDEs has received much attention recently, and powerful theories are being developed to enhance the general understanding of this area. We refer to the excellent surveys [8, 11] and references therein for recent breakthroughs in the field.

One of the motivations to study singular SPDEs is that many of them are expected to be universal objects in crossover regimes of their respective universality classes, a phenomenon known as weak universality. One well-known example is the KPZ equation [14], formally given by

---

✉ Weijun Xu  
weijunx@gmail.com

<sup>1</sup> University of Oxford, Oxford, UK

<sup>2</sup> New York University Shanghai, Shanghai, China

$$\partial_t h = \partial_x^2 h + \lambda(\partial_x h)^2 + \xi,$$

where  $\xi$  is the one-dimensional space-time white noise. The equation is only formal since it involves the square of a distribution. Nevertheless, the solution can be rigorously constructed in a few different ways, including the Cole–Hopf transform [1], pathwise solutions via rough paths/regularity structures [3, 9, 10] or paracontrolled distributions [6], or the notion of energy solution through a martingale problem [4, 7].

The KPZ equation is expected to be the universal model for weakly asymmetric interface growth at large scales. In [12], the authors considered continuous microscopic models of the type

$$\partial_t \tilde{h} = \partial_x^2 \tilde{h} + \sqrt{\varepsilon} F(\partial_x \tilde{h}) + \tilde{\xi} \quad (1.1)$$

for any even polynomial  $F$  and smooth stationary Gaussian random field  $\tilde{\xi}$ . The main result in [12] is that there exists  $C_\varepsilon \rightarrow +\infty$  such that the rescaled and re-centered height function

$$h_\varepsilon(t, x) := \varepsilon^{\frac{1}{2}} \tilde{h}(t/\varepsilon^2, x/\varepsilon) - C_\varepsilon t$$

converges to the solution of the KPZ equation with

$$\lambda = \frac{1}{2} \mathbf{E} F''(\tilde{\Psi}), \quad (1.2)$$

where  $\tilde{\Psi} = \partial_x P * \tilde{\xi}$  and  $P$  is the heat kernel. Hairer and Xu [13] extended the result to arbitrary even functions  $F$  with sufficient regularity and polynomial growth. Similar results have also been obtained in [5] for models at stationarity.

To see why the convergence holds with  $\lambda$  given by (1.2), we write down the equation for  $h_\varepsilon$ :

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \varepsilon^{-1} F(\sqrt{\varepsilon} \partial_x h_\varepsilon) + \xi_\varepsilon - C_\varepsilon,$$

where  $\xi_\varepsilon(t, x) = \varepsilon^{-\frac{3}{2}} \tilde{\xi}(t/\varepsilon^2, x/\varepsilon)$  approximates the space-time white noise  $\xi$  at scale  $\varepsilon$ . Let  $\Psi_\varepsilon = \partial_x P * \xi_\varepsilon$ , then  $\sqrt{\varepsilon} \Psi_\varepsilon$  is stationary Gaussian with finite variance. In addition, by analogy with the standard KPZ equation, it is reasonable to expect that the remainder  $\partial_x u_\varepsilon = \partial_x h_\varepsilon - \Psi_\varepsilon$  is almost bounded. Hence, one can Taylor expand the nonlinearity  $\varepsilon^{-1} F(\sqrt{\varepsilon} \partial_x h_\varepsilon)$  around  $\sqrt{\varepsilon} \Psi_\varepsilon$  and formally get

$$\begin{aligned} \varepsilon^{-1} F(\sqrt{\varepsilon} \partial_x h_\varepsilon) - C_\varepsilon &= \left( \varepsilon^{-1} F(\sqrt{\varepsilon} \Psi_\varepsilon) - C_\varepsilon \right) + \varepsilon^{-\frac{1}{2}} F'(\sqrt{\varepsilon} \Psi_\varepsilon) \cdot (\partial_x u_\varepsilon) \\ &\quad + F''(\sqrt{\varepsilon} \Psi_\varepsilon) \cdot (\partial_x u_\varepsilon)^2 + \mathcal{O}(\varepsilon^{\frac{1}{2}-}). \end{aligned}$$

One then needs to show the convergence of the objects  $\varepsilon^{-1} F(\sqrt{\varepsilon} \Psi_\varepsilon) - C_\varepsilon$ ,  $\varepsilon^{-\frac{1}{2}} F'(\sqrt{\varepsilon} \Psi_\varepsilon)$ ,  $F''(\sqrt{\varepsilon} \Psi_\varepsilon)$  as well as their products, which arise from the local expansion of  $\partial_x u_\varepsilon$ . Furthermore, such convergences need to be established in the optimal regularity space, which requires one to  $p$ th obtain moment bounds of these stochastic objects for arbitrarily large  $p$ th.

At least formally, by chaos expanding  $\varepsilon^{-1} F(\sqrt{\varepsilon} \Psi_\varepsilon)$  and taking  $C_\varepsilon = \varepsilon^{-1} \mathbf{E} F(\sqrt{\varepsilon} \Psi_\varepsilon)$ , one can see that

$$\varepsilon^{-1} F(\sqrt{\varepsilon} \Psi_\varepsilon) - C_\varepsilon \rightarrow \lambda \Psi^{\diamond 2}, \quad (1.3)$$

where  $\lambda$  is given in (1.2) and  $\Psi = \partial_x P * \xi$  is the limit of  $\Psi_\varepsilon$ . This is because all terms starting from the 4-th chaos vanish termwise as  $\varepsilon \rightarrow 0$ , and only the second chaos component survives in the limit.

When  $F$  is even polynomial, this heuristic indeed gives a direct proof of the convergence of the term in (1.3). However, when  $F$  is not polynomial, the actual proof of the convergence becomes much subtler. The main obstacle is that  $F(\sqrt{\varepsilon}\Psi_\varepsilon)$  expands into an infinite chaos series. If we brutally control their high moments termwise as in the polynomial case, then in order for these termwise moment bounds to be summable, we need to impose very strong conditions on  $F$  (namely, its Fourier transform being compactly supported), which is clearly too restrictive.

Instead, in [13], the authors expanded  $F(\sqrt{\varepsilon}\Psi_\varepsilon)$  in terms of Fourier transform, developed a procedure in obtaining pointwise correlation bounds on trigonometric functions of Gaussians, and deduced the desired convergence from those bounds.

Similar universality results are also present in the dynamical  $\Phi_3^4$  model. The weak universality of  $\Phi_3^4$  equation for a large class of symmetric phase coexistence models with polynomial potential was established in [13] for Gaussian noise and then extended in [17] to non-Gaussian noise. The extension beyond polynomial potential (even with Gaussian noise) has the same difficulties as in the KPZ case discussed above. In the recent work [2], the authors developed different methods based on Malliavin calculus to control similar objects. The methods developed in [2, 13] to treat general nonlinearities are both robust enough to cover both KPZ and  $\Phi_3^4$  equations as well as other similar situations.

In this article, we follow the ideas developed in [13] and prove a uniform bound in a special case considered in there. This special case is technically simpler to explain, but is also illustrative enough to reveal the main idea of the proof for the more general case. Furthermore, we obtain a better bound in this special case, thus yielding convergence results for functions  $F$  with lower regularity.

## 1.2 Main Statements

Fix a scaling  $\mathfrak{s} = (s_1, \dots, s_d)$  on  $\mathbf{R}^d$ , and let  $|\mathfrak{s}| = \sum_j s_j$ . The metric induced by  $\mathfrak{s}$  is

$$|x|_{\mathfrak{s}} := |x_1|^{\frac{1}{s_1}} + \dots + |x_d|^{\frac{1}{s_d}}.$$

Since the scaling is fixed throughout the article, we simply write  $|x|$  instead of  $|x|_{\mathfrak{s}}$ . For any Gaussian random field  $X$ , any function  $F : \mathbf{R} \rightarrow \mathbf{R}$  with at most exponential growth, and any integer  $m \geq 0$ , we write

$$\mathcal{H}_m(F(X)) = \sum_{n \geq m} C_n X^{\diamond n},$$

where  $X^{\diamond n}$  denotes the  $n$ -th Wick power of  $X$ , and  $C_n = \frac{1}{n!} \mathbf{E} F^{(n)}(X)$  is the coefficient of  $X^{\diamond n}$  in the chaos expansion of  $F(X)$ . In other words,  $\mathcal{H}_m(F(X))$  is  $F(X)$  with the first  $m - 1$  chaos removed. We refer to [16, Chapter 1] for more details on chaos expansion of random variables. We have the following bound.

**Theorem 1.1** Let  $\alpha \in (0, |\mathfrak{s}|)$ , and  $\{\Phi_\varepsilon\}_{\varepsilon \in (0,1)}$  be a class of centered Gaussian random fields satisfying

$$\frac{\varepsilon^\alpha}{\Lambda(|x-y|+\varepsilon)^\alpha} \leq \mathbf{E}(\Phi_\varepsilon(x)\Phi_\varepsilon(y)) \leq \frac{\Lambda\varepsilon^\alpha}{(|x-y|+\varepsilon)^\alpha} \quad (1.4)$$

for some  $\Lambda > 1$  and for all  $x, y \in \mathbf{R}^d$  and all  $\varepsilon \in (0, 1)$ . Then, for every  $K \geq 1$  and  $m, r \in \mathbf{N}$ , there exists  $C > 0$  depending on these parameters and  $\Lambda$  only such that

$$\left| \mathbf{E} \prod_{j=1}^K \partial_\theta^r \mathcal{H}_m \left( e^{i\theta \Phi_\varepsilon(x_j)} \right) \right| \leq C \mathbf{E} \prod_{j=1}^K \left( \Phi_\varepsilon^{\diamond m}(x_j) + \Phi_\varepsilon^{\diamond(m+1)}(x_j) \right) \quad (1.5)$$

for all  $\varepsilon \in (0, 1)$ ,  $\theta \in \mathbf{R}$  and  $\mathbf{x} = (x_k)_{k=1}^K$ .

Theorem 1.1 is the main technical ingredient to establish that if  $\{\Psi_\varepsilon\}$  approximates a certain Gaussian random field  $\Psi$ , then a large class of nonlinear functions of  $\Psi_\varepsilon$ , after proper rescaling and re-centering, converges to certain Wick powers of  $\Psi$ . We first give the assumption on the random field  $\Psi$ .

**Assumption 1.2**  $\Psi$  is a stationary Gaussian random field with correlation<sup>1</sup>

$$\mathbf{E}(\Psi(x)\Psi(y)) = G(x-y),$$

where  $G$  satisfies the bounds

$$\frac{c}{|x|^\alpha} \leq G(x) \leq \frac{C}{|x|^\alpha} \quad \text{and} \quad |(\partial_j G)(x)| \leq C|x|^{-\alpha-s_j}$$

for some  $\alpha \in (0, |\mathfrak{s}|)$  and all  $x \in \mathbf{R}^d$ . In addition, there exists a locally integrable function  $g$  such that

$$\varepsilon^\alpha G(\varepsilon^{s_1}x_1, \dots, \varepsilon^{s_d}x_d) \rightarrow g(x) \quad (1.6)$$

in  $L^1(\Omega)$  for every bounded subset  $\Omega$  of  $\mathbf{R}^d$ .

For  $M \in \mathbf{N}$  and open subset  $\mathcal{I} \subset \mathbf{R}$ , we define the norm  $\|\cdot\|_{M,\mathcal{I}}$  on distributions on  $\mathbf{R}$  by

$$\|\Upsilon\|_{M,\mathcal{I}} := \sup_{0 \leq r \leq M} \sup_{\substack{\varphi \in \mathcal{C}_c^M(\mathcal{I}): \\ \|\varphi\|_{\mathcal{C}^M(\mathcal{I})} \leq 1}} |\langle \Upsilon, \varphi^{(r)} \rangle|.$$

Our assumption on the function  $F : \mathbf{R} \rightarrow \mathbf{R}$  is the following.

<sup>1</sup> A rigorous way of saying the correlation is  $G$  is that

$$\mathbf{E}\langle \Psi, \varphi \rangle \langle \Psi, \phi \rangle = \int_{\mathbf{R}^d} G(x-y) \varphi(x) \phi(y) dx dy$$

for all  $\varphi, \phi \in \mathcal{C}_c^\infty(\mathbf{R}^d)$ .

**Assumption 1.3** There exists  $M \in \mathbf{N}$  such that the Fourier transform of  $F$  satisfies

$$\sum_{k \in \mathbf{Z}} \|\widehat{F}\|_{M, \mathcal{I}_k} < +\infty,$$

where  $\mathcal{I}_k = (k - 1, k + 1)$ .

For every  $\rho \in \mathcal{C}_c^\infty(\mathbf{R}^d)$  and  $\varepsilon > 0$ , let

$$\rho_\varepsilon(x) = \varepsilon^{-|\mathfrak{s}|} \rho(x_1/\varepsilon^{s_1}, \dots, x_d/\varepsilon^{s_d}).$$

The main convergence theorem is the following.

**Theorem 1.4** *Let  $\Psi$  and  $F$  satisfy the above assumptions, and  $g$  be the limiting  $L^1$  function of  $\varepsilon^\alpha G(\varepsilon \cdot)$  as in Assumption 1.2. Let  $\rho$  be a mollifier on  $\mathbf{R}^d$  and  $\Psi_\varepsilon = \Psi * \rho_\varepsilon$ . For every integer  $m$ , define*

$$a_m := \frac{1}{m!} \left( F^{(m)} * \mu \right) (0), \quad (1.7)$$

where  $\mu \sim \mathcal{N}(0, \sigma^2)$  is a Gaussian measure on  $\mathbf{R}$  with variance

$$\sigma^2 = \int g(x - y) \rho(x) \rho(y) dx dy. \quad (1.8)$$

Then, for every  $m < \frac{|\mathfrak{s}|}{\alpha}$  and every sufficiently small  $\kappa$ , we have

$$\varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m \left( F(\varepsilon^{\frac{\alpha}{2}} \Psi_\varepsilon) \right) \rightarrow a_m \Psi^{\diamond m}$$

as  $\varepsilon \rightarrow 0$  almost surely in  $\mathcal{C}^{-\frac{m\alpha}{2} - \kappa}$ . Here,  $\Psi^{\diamond m}$  is the  $m$ -th Wick power of  $\Psi$ .

We will first prove the main bound (1.5) in Theorem 1.1 and then establish the convergence in Theorem 1.4 by Fourier expanding  $F$  and applying (1.5) to  $\Phi_\varepsilon = \varepsilon^{\frac{\alpha}{2}} \Psi_\varepsilon$ . Note that although the bound in Theorem 1.1 holds for every integer  $m$ , the convergence in Theorem 1.4 requires  $m < \frac{|\mathfrak{s}|}{\alpha}$ . This can be easily seen from the fact that if  $m \geq \frac{|\mathfrak{s}|}{\alpha}$ , then  $\Psi^{\diamond m}$  would have divergent covariance and hence not well defined.

**Example 1.5** One typical example of the random field  $\Psi$  satisfying Assumption 1.2 is  $\Psi = \mathcal{L}^{-\frac{\beta}{2}} \xi$ , where  $\xi$  is the white noise on  $\mathbf{R}^d$ ,  $\beta = \frac{1}{2}(|\mathfrak{s}| - \alpha)$ , and  $\mathcal{L}$  is the differential operator given by

$$\mathcal{L} = \sum_{j=1}^d \left( -\partial_j^2 \right)^{\frac{1}{s_j}}.$$

This is the fractional Gaussian field. In this case, the convolution kernel  $K$  of  $\mathcal{L}^{-\frac{\beta}{2}}$  is homogeneous (in the scaling  $\mathfrak{s}$ ) of order  $-|\mathfrak{s}| + \beta$  in the sense that

$$K(\lambda^{s_1} x_1, \dots, \lambda^{s_d} x_d) = \lambda^{-|\mathfrak{s}| + \beta} K(x_1, \dots, x_d)$$

for all  $\lambda > 0$ . Hence, for  $g, G$  as in Assumption 1.2, we have the expression

$$g(x) = G(x) = \int_{\mathbf{R}^d} K(x+y)K(y)dy.$$

The same is true when  $\mathbf{R}^d$  is given the parabolic scaling  $(2, 1, \dots, 1)$  and  $\mathcal{L} = \partial_t - \Delta$  is the heat operator.

In the case of standard Euclidean scaling  $\mathfrak{s} = (1, \dots, 1)$ ,  $\Psi$  is simply the standard fractional Gaussian field  $(-\Delta)^{-\frac{\beta}{2}}\xi$ . We refer to the survey [15] for more details on fractional Gaussian fields.

**Example 1.6** As for the function  $F$ , all  $\mathcal{C}^{1+}$  functions with polynomial growth fall in Assumption 1.3. More precisely, if  $f \in \mathcal{C}^{1,\beta}(\mathbf{R})$  for some  $\beta > 0$ , and there exist  $C, M > 0$  such that

$$|f(x)| + |f'(x)| + \sup_{|h|<1} \frac{|f'(x+h) - f'(x)|}{|h|^\beta} \leq C(1 + |x|)^M$$

for all  $x \in \mathbf{R}$ , then  $F$  satisfies Assumption 1.3.

**Example 1.7** One very interesting example of the microscopic model (1.1) is with  $F(x) = |x|$ . It is almost linear, but one still expects its large-scale behavior to be nonlinear as described by the KPZ equation. This function  $F$  even not  $\mathcal{C}^1$ , but we still have

$$\|\widehat{F}\|_{1,\mathcal{I}_k} \leq C(1 + |k|)^{-2},$$

which clearly satisfies Assumption 1.3. Hence, as a consequence of Theorem 1.1, if  $\Psi = \partial_x P * \xi$  where  $\xi$  is the space-time white noise with one space dimension, and  $P$  is the heat kernel, then we have

$$\varepsilon^{-\frac{1}{2}} (|\Psi_\varepsilon| - \mathbf{E}|\Psi_\varepsilon|) \rightarrow a\Psi^{\diamond 2}$$

for some  $a > 0$  depending on the mollifier. This is the first step toward establishing convergence to the KPZ equation for the microscopic model of the form (1.1) with  $F(x) = |x|$ .

### 1.3 Remarks and Possible Generalizations

Theorem 1.1 is a special case of [13, Theorem 6.4] in that it allows only one frequency variable  $\theta$  rather than multiple ones. On the other hand, it is also more general since it allows subtraction of Wiener chaos up to any order. Furthermore, the bound (1.5) is completely independent of  $\theta$ , while the corresponding one in [13] is polynomial in  $\theta$ . As a consequence of this improvement, the condition on  $F$  for the convergence in Theorem 1.4 to hold is weaker.

The main technical difference that results in this improvement, as we shall see later in Sect. 2, is that in the clustering procedure, we are able to take the clustering distance  $L$  being independent of  $\theta$  rather than being quadratic in  $\theta$  as in [13].

We shall note that the convergence results in this article are not sufficient to establish weak universality in general situations. These would require convergence of the products of the objects considered in Theorem 1.4, with possible subtraction of extra chaos components after taking product. The convergence of these products requires a more general bound than Theorem 1.1 and [13, Theorem 6.2]. We leave them to future work.

## 2 Proof of Theorem 1.1

This section is devoted to the proof of Theorem 1.1. Assumptions 1.2 and 1.3 on  $\Psi$  and  $F$  are irrelevant here. We fix  $\alpha \in (0, |\mathfrak{s}|)$ , and let  $\{\Phi_\varepsilon\}_{\varepsilon \in (0,1)}$  be a family of Gaussian random fields with correlation functions satisfying (1.4). The following preliminary bounds on the correlation function will be used throughout the section.

**Proposition 2.1** *Let  $\gamma \geq 1$ . If  $|x' - y'| \leq \gamma|x - y|$ , then*

$$\mathbf{E}(\Phi_\varepsilon(x)\Phi_\varepsilon(y)) \leq \gamma^\alpha \Lambda^2 \mathbf{E}(\Phi_\varepsilon(x')\Phi_\varepsilon(y')). \quad (2.1)$$

*The bound is uniform over all  $\varepsilon \in (0, 1)$  and all pairs of points  $(x, y), (x', y') \in (\mathbf{R}^d)^2$  satisfying the above constraint. As a consequence, we have*

$$\mathbf{E}(\Phi_\varepsilon(x)\Phi_\varepsilon(y)) \mathbf{E}(\Phi_\varepsilon(x)\Phi_\varepsilon(z)) \leq \frac{2^\alpha \Lambda^3 \varepsilon^\alpha}{((|x - y| \wedge |x - z|) + \varepsilon)^\alpha} \cdot \mathbf{E}(\Phi_\varepsilon(y)\Phi_\varepsilon(z)) \quad (2.2)$$

*for all  $\varepsilon \in (0, 1)$  and all  $x, y, z \in \mathbf{R}^d$ .*

**Proof** The first bound follows from

$$\begin{aligned} \mathbf{E}(\Phi_\varepsilon(x')\Phi_\varepsilon(y')) &\geq \frac{\varepsilon^\alpha}{\Lambda(|x' - y'| + \varepsilon)^\alpha} \geq \frac{1}{\Lambda^2 \gamma^\alpha} \cdot \frac{\Lambda \varepsilon^\alpha}{(|x - y| + \varepsilon)^\alpha} \\ &\geq \frac{\mathbf{E}(\Phi_\varepsilon(x)\Phi_\varepsilon(y))}{\Lambda^2 \gamma^\alpha}, \end{aligned}$$

where we have used the assumption  $\gamma \geq 1$  in the second inequality. As for the second one, it suffices to notice

$$|y - z| \leq 2 \max\{|x - y|, |x - z|\}$$

and then apply (2.1).  $\square$

In what follows, we keep our notations same as in [13, Section 6]. For every finite set  $\mathcal{A}$ , let  $\mathbf{N}^{\mathcal{A}}$  be the set of multi-indices on  $\mathcal{A}$ . For  $\mathcal{A}$ -tuple of Gaussian random variables  $\mathbf{X} = (X_a)_{a \in \mathcal{A}}$  and  $\mathbf{n} \in \mathbf{N}^{\mathcal{A}}$ , we write  $\mathbf{X}^{\diamond \mathbf{n}} = \diamond_{a \in \mathcal{A}} X_a^{\diamond n_a}$ . Similarly, we write  $\mathbf{n}! = \prod_{a \in \mathcal{A}} n_a!$  and  $|\mathbf{n}| = \sum_{a \in \mathcal{A}} n_a$ . In general, we use standard letters for scalars and boldface ones to denote tuples.

Fix  $K \geq 1$  and  $m, r \in \mathbf{N}$ . Let  $[K] = \{1, \dots, K\}$ . We also fix  $\theta \in \mathbf{R}$  and  $\mathbf{x} = (x_k)_{k \in [K]} \in \mathbf{R}^K$  arbitrary. Write  $\langle \theta \rangle = 1 + |\theta|$ . All the constants  $C$  below depend on  $\Lambda, K, m$ , and  $r$  only unless otherwise mentioned. We seek bounds that are uniform in  $\varepsilon, \theta$ , and  $\mathbf{x}$ . We also write  $X_j = \Phi_\varepsilon(x_j)$  for simplicity since the bounds will be independent of  $\varepsilon$ .

## 2.1 Clustering and the First Bound

Let  $L > 0$  be a fixed large constant whose value, depending on  $\Lambda, K, m$ , and  $r$  only, will be specified later. Let  $\sim$  be an equivalence relation on  $[K]$  such that  $j \sim j'$  if there exists  $k \in \mathbf{N}$  and  $j_0, \dots, j_k \in [K]$  with  $j_0 = j$  and  $j_k = j'$  such that

$$|x_{j_{\ell+1}} - x_{j_\ell}| \leq L\varepsilon$$

for all  $\ell = 0, \dots, k-1$ . We let  $\mathcal{C}$  denote the partition of  $[K]$  into clusters obtained in this way. In other words,  $j$  and  $j'$  belong to the same cluster if and only if starting from  $x_j$ , one can reach  $x_{j'}$  by performing jumps with sizes at most  $L\varepsilon$  onto connecting points in  $\mathbf{x}$ .

We distinguish two cases depending on whether  $\mathcal{C}$  contains singletons or not. We first prove (1.5) when  $\mathcal{C}$  has no singleton; that is, every cluster in  $\mathcal{C}$  has at least two elements. In this case, we write down the explicit expression

$$\partial_\theta^r \mathcal{H}_m \left( e^{i\theta X_j} \right) = (iX_j)^r e^{i\theta X_j} - \sum_{n \leq m-1} \frac{i^n}{n!} \partial_\theta^r \left( e^{-\frac{\theta^2 \mathbf{E} X_j^2}{2}} \theta^n \right) X_j^{\otimes n}. \quad (2.3)$$

Since  $\mathbf{E} X_j^2 \in [\Lambda^{-1}, \Lambda]$ , the coefficients  $\partial_\theta^r (e^{-\frac{\theta^2 \mathbf{E} X_j^2}{2}} \theta^n)$  are uniformly bounded in  $\theta$ . We then plug the expansion (2.3) into the left-hand side of (1.5). The Gaussianity of the  $X_j$ 's and boundedness of the coefficients of the removed chaos imply

$$\left| \mathbf{E} \prod_{j=1}^K \partial_\theta^r \mathcal{H}_m \left( e^{i\theta X_j} \right) \right| \leq C \quad (2.4)$$

for some  $C$  independent of  $\varepsilon, \theta$ , and  $\mathbf{x}$ . It then remains to show that the right-hand side of (1.5) is bounded by a constant from below. For this, we use the assumption that  $\mathcal{C}$  contains no singletons.

Let  $u \in \mathcal{C}$  be arbitrary, and we label its elements by  $u = \{j_1, \dots, j_{|u|}\}$ . Since  $\mathcal{C}$  has no singleton set, we necessarily have  $|u| \geq 2$ . By clustering, any two points in  $u$  are at most  $KL\varepsilon$  away from each other. Hence, the assumption (1.4) implies (recalling that  $X_j = \Phi_\varepsilon(x_j)$ )

$$\mathbf{E} (X_j X_{j'}) \geq \frac{\varepsilon^\alpha}{\Lambda (|x_j - x_{j'}| + \varepsilon)^\alpha} \geq \frac{1}{\Lambda (KL + 1)^\alpha}$$



for every two points  $j, j' \in u$ . which implies

$$(\Lambda(KL + 1)^\alpha)^{-\lfloor \frac{m+1}{2} \rfloor |u|} \leq \left( \prod_{\ell=1}^{|u|} \mathbf{E}(X_{j_\ell} X_{j_{\ell+1}}) \right)^{\lfloor \frac{m+1}{2} \rfloor},$$

where we identified  $j_{|u|+1}$  with  $j_1$ . Multiplying the above bound over all  $u \in \mathcal{C}$  and using Wick's formula and positivity of the correlations, we obtain

$$\begin{aligned} (\Lambda(KL + 1)^\alpha)^{-\lfloor \frac{m+1}{2} \rfloor K} &\leq \prod_{u \in \mathcal{C}} \left( \prod_{\ell=1}^{|u|} \mathbf{E}(X_{j_\ell} X_{j_{\ell+1}}) \right)^{\lfloor \frac{m+1}{2} \rfloor} \\ &\leq \mathbf{E} \prod_{j=1}^K (X_j^{\diamond m} + X_j^{\diamond(m+1)}). \end{aligned} \quad (2.5)$$

This is the place where we use  $|u| \geq 2$  for every  $u \in \mathcal{C}$ , for otherwise the middle term above would contain the variance of a single random variable and the second inequality in (2.5) would be wrong. Combining (2.4) and (2.5), we obtain

$$\left| \mathbf{E} \partial_\theta^r \prod_{j=1}^K \mathcal{H}_m(e^{i\theta X_j}) \right| \leq C (\Lambda(KL + 1)^\alpha)^{K \lfloor \frac{m+1}{2} \rfloor} \cdot \mathbf{E} \prod_{j=1}^K (X_j^{\diamond m} + X_j^{\diamond(m+1)}), \quad (2.6)$$

which matches the right-hand side of (1.5). Since  $L$  (to be chosen later) is also independent of  $\varepsilon, \theta$ , and  $\mathbf{x}$ , this concludes the case when  $\mathcal{C}$  contains no singleton. The rest of the section is devoted to establishing (1.5) when at least one cluster in  $\mathcal{C}$  is singleton.

## 2.2 Expansion

Given the collection of points  $\mathbf{x}$  and the clustering above, let

$$\mathcal{S} = \{u \in \mathcal{C} : |u| = 1\}$$

be the set of singletons in  $\mathcal{C}$ , and let  $\mathcal{U} = \mathcal{C} \setminus \mathcal{S}$ . We write  $s \in \mathcal{S}$  for simplicity if  $\{s\}$  is a singleton set in  $\mathcal{C}$ .

For  $\mathbf{X} = (X_j)_{j \in [K]}$  and  $\mathbf{n} = (n_j)_{j \in [K]}$ , we write  $\mathbf{X}_u$  and  $\mathbf{n}_u$  for their restrictions to  $u$ , and  $\mathbf{X}_u^{\diamond \mathbf{n}_u} = \diamond_{j \in u} X_j^{\diamond n_j}$ . Splitting the left-hand side of (1.5) into sub-products within clusters  $u \in \mathcal{C}$  and chaos expanding each sub-product, we can rewrite it as

$$\mathbf{E} \prod_{j=1}^K \partial_\theta^r \mathcal{H}_m(e^{i\theta X_j}) = \sum_{N \geq 0} \sum_{\substack{\mathbf{n} \in \mathbf{N}^K: \\ |\mathbf{n}| = N}} \left( \prod_{u \in \mathcal{C}} C_{\mathbf{n}_u}(\theta, \mathbf{X}_u) \right) \left( \mathbf{E} \prod_{u \in \mathcal{C}} \mathbf{X}_u^{\diamond \mathbf{n}_u} \right). \quad (2.7)$$

Here,  $C_{\mathbf{n}_u}(\theta, \mathbf{X}_u)$  is the coefficient of  $\mathbf{X}_u^{\diamond \mathbf{n}_u}$  in the chaos expansion of  $\prod_{j \in u} \partial_\theta^r \mathcal{H}_m(e^{i\theta X_j})$  and has the expression

$$C_{\mathbf{n}_u}(\theta, \mathbf{X}_u) = \frac{1}{\mathbf{n}_u!} \mathbf{E} \left( \prod_{j \in u} \partial_\theta^r \partial_{X_j}^{n_j} \mathcal{H}_m(e^{i\theta X_j}) \right). \quad (2.8)$$

Note the product involving the expectation on the right-hand side of (2.7) is 0 if  $|\mathbf{n}|$  is odd. But we still sum over all integers  $N$  since this simplifies the notations later. The following lemma gives control on the coefficients  $C_{\mathbf{n}_u}$ .

**Lemma 2.2** *There exists  $C > 0$  depending on  $K, m, r$ , and  $\Lambda$  only such that*

$$|C_{\mathbf{n}_u}(\theta, \mathbf{X}_u)| \leq \frac{(C\langle\theta\rangle)^{|\mathbf{n}_u|}}{\mathbf{n}_u!} \quad (2.9)$$

for all  $\mathbf{n}_u \in \mathbf{N}^u$ , where we recall  $\langle\theta\rangle = 1 + |\theta|$ . As a consequence, we have

$$\sum_{|\mathbf{n}|=N} \prod_{u \in \mathcal{C}} |C_{\mathbf{n}_u}(\theta, \mathbf{X}_u)| \leq \frac{(C\langle\theta\rangle)^N}{N!}. \quad (2.10)$$

Furthermore, if  $\mathcal{S} \neq \emptyset$ , then we have

$$\sum_{|\mathbf{n}|=N} \prod_{u \in \mathcal{C}} |C_{\mathbf{n}_u}(\theta, X_u)| \leq e^{-\frac{\theta^2}{2\Lambda}} \cdot \frac{(C\langle\theta\rangle)^N}{N!}. \quad (2.11)$$

All the bounds are uniform in  $\varepsilon$  and  $\theta$ , and in the location of  $\mathbf{x}$  subject to whether  $\mathcal{C}$  contains a singleton or not.

**Proof** We express the right-hand side of (2.8) in a way that is convenient to estimate. For this, we introduce variables  $\boldsymbol{\beta} \in \mathbf{R}^K$ . Let  $\boldsymbol{\beta}_u$  denote the restriction of  $\boldsymbol{\beta}$  to  $u$ , and write  $\partial_{\boldsymbol{\beta}_u}^r = \prod_{j \in u} \partial_{\beta_j}^r$  and  $\boldsymbol{\beta}_u^{\mathbf{n}_u} = \prod_{j \in u} \beta_j^{n_j}$ . Using the identity

$$\partial_{X_j}^{n_j} \mathcal{H}_m(e^{i\beta_j X_j}) = (i\beta_j)^{n_j} \mathcal{H}_{m-n_j}(e^{i\beta_j X_j}),$$

where  $\mathcal{H}_{m-n} = \mathcal{H}_0$  if  $n \geq m$ , we can rewrite the right-hand side of (2.8) as

$$C_{\mathbf{n}_u}(\theta, \mathbf{X}_u) = \frac{i^{|\mathbf{n}_u|}}{\mathbf{n}_u!} \partial_{\boldsymbol{\beta}_u}^r \left( \boldsymbol{\beta}_u^{\mathbf{n}_u} \cdot \mathbf{E} \prod_{j \in u} \mathcal{H}_{m-n_j}(e^{i\beta_j X_j}) \right) \Big|_{\beta_j=\theta, \forall j \in u}.$$

When distributing  $r$  derivatives of each  $\beta_j$  into the two terms in the parenthesis, the differentiation of the first term ( $\boldsymbol{\beta}_u^{\mathbf{n}_u}$ ) yields an additional factor which is at most  $|\mathbf{n}_u|^{Kr} \leq C^{|\mathbf{n}_u|}$  for some fixed constant  $C$ , while the second term is uniformly bounded

both in  $\theta$  and  $\mathbf{n}_u$  since  $X_j$ 's all have bounded variance. This gives (2.9). The bound (2.10) follows from (2.9) and the multinomial theorem.

Finally, in order to obtain (2.11) when  $\mathcal{S} \neq \emptyset$ , it suffices to note that for  $s \in \mathcal{S}$ , we have

$$C_{n_s}(\theta, X_s) = \frac{i^{n_s}}{n_s!} \partial_\theta^r \left( \theta^{n_s} e^{-\frac{\theta^2 \mathbf{E} X_s^2}{2}} \right)$$

if  $n_s \geq m$ , and is 0 otherwise. Since  $\mathbf{E} X_s^2 \geq \frac{1}{\Lambda}$ , we gain an additional Gaussian factor  $e^{-\frac{\theta^2}{2\Lambda}}$  for every  $s \in \mathcal{S}$ , and hence  $e^{-\frac{\theta^2 |\mathcal{S}|}{2\Lambda}}$  in total. The bound (2.11) then follows from relaxing it to  $e^{-\frac{\theta^2}{2\Lambda}}$ .  $\square$

### 2.3 Representative Point

For  $|u| \geq 2$ , the corresponding term in the expectation on the right-hand side of (2.7) is a Wick product of multiple Gaussian random variables. We aim to reduce it to the Wick product of a single variable by choosing a representative point from each cluster.

For every  $u \in \mathcal{C}$ , choose  $u^*(u) \in u$  arbitrary. The choice for  $u^*$  is unique if  $u$  is singleton. We have the following proposition.

**Proposition 2.3** *There exists  $C > 0$  such that*

$$\mathbf{E} \left( \prod_{u \in \mathcal{C}} \mathbf{X}_u^{\diamond \mathbf{n}_u} \right) \leq C^{|\mathbf{n}|} \cdot \mathbf{E} \left( \prod_{u \in \mathcal{C}} X_{u^*(u)}^{\diamond |\mathbf{n}_u|} \right) \quad (2.12)$$

for every  $\mathbf{n} \in \mathbf{N}^K$  and every choice of  $u^*(u) \in u$ .

**Proof** If  $|\mathbf{n}| = \sum_{u \in \mathcal{C}} |\mathbf{n}_u|$  is odd, then both sides of (2.12) are 0, so we only need to consider the situation when  $|\mathbf{n}|$  is even.

In this case, the left-hand side is the sum of products of pairwise expectations  $\mathbf{E}(X_j X_{j'})$  for  $j$  and  $j'$  belonging to different clusters. The right-hand side (without the factor  $C^{|\mathbf{n}|}$ ) is the same except that each instance of  $X_j$  for  $j \in u$  is replaced by  $X_{u^*(u)}$ . It then suffices to control the effects of such replacements.

Let  $u, v$  be two different clusters in  $\mathcal{C}$ , and use  $u^*$  and  $v^*$  to denote  $u^*(u)$  and  $u^*(v)$ , respectively. For every  $j \in u$  and  $j' \in v$ , according to the clustering, we have

$$|x_{u^*} - x_j| \leq (K-1)|x_j - x_{j'}|, \quad |x_{v^*} - x_{j'}| \leq (K-1)|x_j - x_{j'}|,$$

which implies

$$|x_{u^*} - x_{v^*}| \leq |x_{u^*} - x_j| + |x_j - x_{j'}| + |x_{v^*} - x_{j'}| \leq 2K|x_j - x_{j'}|.$$

Hence, by (2.1), we deduce that

$$\mathbf{E}(X_j X_{j'}) \leq (2K)^\alpha \Lambda^2 \mathbf{E}(X_{u^*} X_{v^*}).$$

This is the effect of one such replacement. The claim then follows since there are  $\frac{|\mathbf{n}|}{2}$  pairwise expectations in each product in the sum. It also means we can take  $C = (2K)^{\frac{\alpha}{2}} \Lambda$  in (2.12).  $\square$

From now on, we restrict ourselves to the situation when  $|\mathcal{S}| \geq 1$ , and recall the notation  $\mathcal{U} = \mathcal{C} \setminus \mathcal{S}$ . We need to split the product on the right-hand side of (2.7) into sub-products in  $\mathcal{S}$  and in  $\mathcal{U}$ . For this, we introduce multi-indices  $\mathbf{k} = (k_s)_{s \in \mathcal{S}}$  and  $\ell = (\ell_u)_{u \in \mathcal{U}}$  and write  $|\mathbf{k}| = \sum_{s \in \mathcal{S}} k_s$  and  $|\ell| = \sum_{u \in \mathcal{U}} \ell_u$ . We then have the following proposition.

**Proposition 2.4** *Suppose  $|\mathcal{S}| \geq 1$ . Then, the left-hand side of (1.5) can be controlled by*

$$\left| \mathbf{E} \prod_{j=1}^K \partial_{\theta}^r \mathcal{H}_m \left( e^{i\theta X_j} \right) \right| \leq C e^{-\frac{\theta^2}{2\Lambda}} \sum_{N \geq 0} \left( \frac{(C \langle \theta \rangle)^{N+m|\mathcal{S}|}}{(N+m|\mathcal{S}|)!} \right. \\ \left. \times \sup_{|\mathbf{k}|+|\ell|=N} \mathbf{E} \left[ \left( \prod_{s \in \mathcal{S}} X_s^{\diamond(m+k_s)} \right) \left( \prod_{u \in \mathcal{U}} X_{u^*(u)}^{\diamond \ell_u} \right) \right] \right). \quad (2.13)$$

**Proof** We start with the expression (2.7). Note that for  $s \in \mathcal{S}$ ,  $C_{n_s}(\theta, X_s) = 0$  whenever  $n_s < m$ , so we can relax the expression to

$$\left| \mathbf{E} \prod_{j=1}^K \partial_{\theta}^r \mathcal{H}_m \left( e^{i\theta X_j} \right) \right| \leq \sum_{N \geq m|\mathcal{S}|} \left( \sum_{\mathbf{n}} \prod_{u \in \mathcal{C}} |C_{\mathbf{n}_u}(\theta, \mathbf{X}_{\mathbf{n}_u})| \right) \left( \sup_{\mathbf{n}} \mathbf{E} \prod_{u \in \mathcal{C}} \mathbf{X}_{\mathbf{n}_u}^{\diamond \mathbf{n}_u} \right),$$

where both the sum and supremum are taken over  $|\mathbf{n}| = N$  with the further restriction that  $n_s \geq m$  for all  $s \in \mathcal{S}$ . The claim follows immediately by applying Lemma 2.2 and Proposition 2.3 to the right-hand side above and noting the range of the sum and supremum.  $\square$

## 2.4 Graphic Representation

It remains to control the term involving the expectation on the right-hand side of (2.13). Since all  $X_j$ 's are Gaussian, it can be written as a sum over products of pairwise expectations. The number of terms in each product (and hence the total power) can be arbitrarily large since  $N$  will be summed over all integers. Following [13], we introduce graphic notations to describe these objects.

Given a set  $\mathbb{V}$ , we write  $\mathbb{V}_2$  for the set of all subsets of  $\mathbb{V}$  with exactly two elements. A (generalized) graph is a triple  $\Gamma = (\mathbb{V}, \mathbb{E}, R)$ . Here,  $\mathbb{V}$  is the set of vertices, and  $\mathbb{E} : \mathbb{V}_2 \rightarrow \mathbf{N}$  is the set of edges with multiplicities. More precisely, each edge  $\{x, y\} \in \mathbb{V}_2$  has multiplicity  $\mathbb{E}(x, y) = \mathbb{E}(y, x)$ . We do not allow self-loops, so  $\mathbb{E}(x, x) = 0$  for all  $x \in \mathbb{V}$ . Finally,  $R : \mathbb{V}_2 \rightarrow \mathbf{R}$  is a function that assigns a value to each pair of vertices.

Given a graph  $\Gamma = (\mathbb{V}, \mathbb{E}, R)$ , we define the degree of a point  $x \in \mathbb{V}$  and of  $\Gamma$ , respectively, by

$$\deg(x) := \sum_{y \in \mathbb{V}} \mathbb{E}(x, y), \quad \deg(\Gamma) := \sum_{x \in \mathbb{V}} \deg(x).$$

The value of  $\Gamma$  is defined by

$$|\Gamma| := \prod_{e \in \mathbb{V}_2} (R(e))^{\mathbb{E}(e)}.$$

In what follows, we always take  $\mathbb{V} = \{x_j\}_{j=1}^K$  fixed (so is the clusters in  $\mathcal{C}$ ), and  $R(x_j, x_{j'}) = \mathbb{E}(X_j X_{j'})$ . We also fix the representative points  $u^*(u)$  chosen for each  $u \in \mathcal{C}$ . Hence, the only variable in our graph is the multiplicity  $\mathbb{E}$  of the edges. Recall the decomposition  $\mathcal{C} = \mathcal{S} \cup \mathcal{U}$  into singletons and clusters with at least two points. We introduce the following definition to characterize the pairings that appear in the expectation term on the right-hand side of (2.13).

**Definition 2.5** For each  $\mathbf{k} \in \mathbf{N}^{\mathcal{S}}$  and  $\ell \in \mathbf{N}^{\mathcal{U}}$ , the set  $\Omega_{\mathbf{k}, \ell}$  consists of graphs with  $\mathbb{V}$  and  $R$  specified above, and such that  $\deg(x_s) = m + k_s$  for all  $s \in \mathcal{S}$ ,  $\deg(x_{u^*(u)}) = \ell_u$  for all  $u \in \mathcal{U}$ , and  $\deg(x) = 0$  for all other  $x \in \mathbb{V}$ .

Let  $\Omega^*$  be the set of graphs  $\Gamma$  such that both of the following hold:

1.  $\Gamma \in \Omega_{\mathbf{k}, \ell}$  for some  $\mathbf{k} \in \mathbf{N}^{\mathcal{S}}$  and  $\ell \in \mathbf{N}^{\mathcal{U}}$  with the restriction that  $k_s \in \{0, 1\}$  for every  $s \in \mathcal{S}$  and  $\ell_u \leq m + 1$  for each  $u \in \mathcal{U}$ .
2. If  $\ell_u \geq 1$  for some  $u \in \mathcal{U}$ , then there exists  $s \in \mathcal{S}$  such that  $\mathbb{E}(x_s, x_{u^*(u)}) = \ell_u$ .

**Remark 2.6** The first requirement for  $\Omega^*$  above is equivalent to that  $\deg(x_s) \in \{m, m + 1\}$  for all  $s \in \mathcal{S}$ ,  $\deg(x_{u^*(u)}) \leq m + 1$  for all  $u \in \mathcal{U}$ , and is 0 for all other points. The second requirement says that if  $x_{u^*(u)}$  has a nonzero degree, then all its edges must be connected to a single point  $x_s$  for some  $s \in \mathcal{S}$ . We will see later that the definition of  $\Omega^*$  corresponds to “minimal graphs” after the reduction procedure in the next subsection.

**Remark 2.7** The clustering depends on the choice of  $L$ , and hence so do the definitions of  $\Omega_{\mathbf{k}, \ell}$  and  $\Omega^*$ . On the other hand, these are just intermediate steps and our final bound (1.5) does not involve clustering at all. Furthermore, the choice of  $L$  later [in (2.18)] is also independent of the location of  $\mathbf{x}$ . Hence, we omit the dependence of the clustering on  $L$  here for notational simplicity.

## 2.5 Reduction

We now start to control the right-hand side of (2.13). If  $m|\mathcal{S}| + |\mathbf{k}| + |\ell|$  is odd, then the term with the expectation is 0. So we only need to deal with the case when  $m|\mathcal{S}| + |\mathbf{k}| + |\ell|$  is even.

In that case, the number of different pairings contributing to the expectation in (2.13) is at most  $(m|\mathcal{S}| + |\mathbf{k}| + |\ell| - 1)!!$ , so with Definition 2.5, we have

$$\mathbb{E} \left[ \left( \prod_{s \in \mathcal{S}} X_s^{\diamond(m+k_s)} \right) \left( \prod_{u \in \mathcal{U}} X_{u^*(u)}^{\diamond \ell_u} \right) \right] \leq (|\mathbf{k}| + |\ell| + m|\mathcal{S}| - 1)!! \cdot \sup_{\Gamma \in \Omega_{\mathbf{k}, \ell}} |\Gamma|. \quad (2.14)$$

Comparing the above bound and the right-hand side of (2.13), we see that we need to control  $|\Gamma|$  for  $\Gamma \in \Omega_{\mathbf{k}, \ell}$  with arbitrarily large  $\mathbf{k}$  and  $\ell$ . We first bound it by values of the graphs in  $\Omega^*$ , which is done via a reduction procedure. After that, we enhance the graphs in  $\Omega^*$  to match the right-hand side of (1.5) to conclude the proof.

We start with the reduction step. This is where we need to choose the clustering distance  $L$  sufficiently large, which will ensure the uniform in  $\theta$  bound after summing over  $\mathbf{k}$  and  $\ell$ . We first give the following proposition, which reduces graphs in  $\Omega_{\mathbf{k}, \ell}$  to those in  $\Omega^*$ .

**Proposition 2.8** *There exists  $C > 0$  depending on  $\Lambda$  only such that*

$$\max_{\Gamma \in \Omega_{\mathbf{k}, \ell}} |\Gamma| \leq \max_{\Gamma^* \in \Omega^*} \left( (CL^{-\alpha})^{\frac{1}{2}(\deg(\Gamma) - \deg(\Gamma^*))} \cdot |\Gamma^*| \right)$$

for every pair  $(\mathbf{k}, \ell)$  and every  $L > 0$ . The constant  $C$  does not depend on the choice of  $L$ , though the clusters and the definitions of  $\Omega_{\mathbf{k}, \ell}$  and  $\Omega^*$  do.

**Proof** Fix  $\mathbf{k} \in \mathbb{N}^{\mathcal{S}}$ ,  $\ell \in \mathbb{N}^{\mathcal{U}}$ , and  $\Gamma \in \Omega_{\mathbf{k}, \ell}$  arbitrary. It suffices to show that if  $\Gamma \notin \Omega^*$ , then we can find a  $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}}, \bar{\ell}}$  with  $\bar{\mathbf{k}} \leq \mathbf{k}$ ,  $\bar{\ell} \leq \ell$  and  $|\bar{\mathbf{k}}| + |\bar{\ell}| < |\mathbf{k}| + |\ell|$  strictly such that

$$|\Gamma| \leq (CL^{-\alpha})^{\frac{1}{2}(|\mathbf{k} - \bar{\mathbf{k}}| + |\ell - \bar{\ell}|)} |\bar{\Gamma}|. \quad (2.15)$$

One can then iterate this bound until the graph is reduced to some  $\Gamma^* \in \Omega^*$  to conclude the proposition. This necessarily happens since each time the total degree of the graph decreases strictly. Here, the inequality on  $\mathbf{k}$  and  $\ell$  means the inequality in each component.

To see the existence of such a  $\bar{\Gamma}$  when  $\Gamma \in \Omega_{\mathbf{k}, \ell} \setminus \Omega^*$ , we consider the two situations where either one of the two conditions for  $\Omega^*$  in Definition 2.5 is violated. We first consider the violation of Condition 1. Since  $\Gamma \in \Omega_{\mathbf{k}, \ell}$ , failure of Condition 1 means there exists  $j \in \mathcal{S} \cup \{u^*(u) : u \in \mathcal{U}\}$  such that  $\deg(x_j) \geq m + 2$ . We fix this  $j$ , and there are two possibilities in this situation.

*Case 1.* There exist  $i \neq i'$  such that  $\mathbb{E}(x_j, x_i) \geq 1$  and  $\mathbb{E}(x_j, x_{i'}) \geq 1$ . In this case, we let  $\bar{\Gamma}$  be the graph obtained from  $\Gamma$  by performing the following operations:

$$\mathbb{E}(x_j, x_i) \mapsto \mathbb{E}(x_j, x_i) - 1, \quad \mathbb{E}(x_j, x_{i'}) \mapsto \mathbb{E}(x_j, x_{i'}) - 1, \quad \mathbb{E}(x_i, x_{i'}) \mapsto \mathbb{E}(x_i, x_{i'}) + 1.$$

The only point whose degree has been changed in this operation is  $x_j$  (reduced by 2). Hence, we have  $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}}, \bar{\ell}}$  with  $(\bar{\mathbf{k}}, \bar{\ell}) \leq (\mathbf{k}, \ell)$  and  $|\mathbf{k} - \bar{\mathbf{k}}| + |\ell - \bar{\ell}| = 2$ . To see the bound (2.15), we note that by definition of  $\Omega_{\mathbf{k}, \ell}$ ,  $x_j$  is at least  $L\varepsilon$  away from both  $x_i$  and  $x_{i'}$ . Hence, by (2.2), we have

$$\mathbb{E}(X_j X_i) \cdot \mathbb{E}(X_j X_{i'}) \leq \frac{2^\alpha \Lambda^3}{(L+1)^\alpha} \cdot \mathbb{E}(X_i X_{i'}).$$

In graphic notation, this means

$$\left| \begin{array}{c} i \quad \quad c \quad \quad i' \\ \quad \swarrow \quad \searrow \\ \quad a+1 \quad b+1 \\ \quad \quad \downarrow \\ \quad \quad j \end{array} \right| \leq \frac{2^\alpha \Lambda^3}{(L+1)^\alpha} \cdot \left| \begin{array}{c} i \quad \quad c+1 \quad \quad i' \\ \quad \swarrow \quad \searrow \\ \quad a \quad \quad b \\ \quad \quad \downarrow \\ \quad \quad j \end{array} \right|,$$

where we have omitted  $x$  and simply write the indices to denote vertices. Since all the other parts of the graph remain unchanged, this operation gives a desired  $\bar{\Gamma}$  with (2.15).

*Case 2.*

If for the  $x_j$  that violates Condition 1, all its edges are connected to another point  $x_i$ , then we necessarily have  $\deg(x_i) \geq m+2$ . Thus, we let  $\bar{\Gamma}$  be the graph obtained from  $\Gamma$  by reducing  $\mathbb{E}(x_j, x_i)$  by two. Then,  $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}}, \bar{\ell}}$  with  $(\bar{\mathbf{k}}, \bar{\ell}) \leq (\mathbf{k}, \ell)$  but this time  $|\mathbf{k} - \bar{\mathbf{k}}| + |\ell - \bar{\ell}| = 4$ . Since  $|x_j - x_i| \geq L\varepsilon$ , we also have the bound

$$|\Gamma| \leq \frac{\Lambda^2}{(L+1)^{2\alpha}} |\bar{\Gamma}|,$$

which is also of the form (2.15). This completes the treatment of the violation of Condition 1.

We now turn to the situation when Condition 2 is violated. This means there exists  $u \in \mathcal{U}$  such that

- (a) either  $x_{u^*(u)}$  is connected to two other different points  $x_i$  and  $x_{i'}$ ;
- (b) or  $x_{u^*(u)}$  is connected to  $x_{u^*(u')}$  for some  $u' \in \mathcal{U}$ .

For (a), we perform exactly the same operation as Case 1 in the above situation. This will give rise to a graph in  $\Omega_{\bar{\mathbf{k}}, \bar{\ell}}$  with  $\bar{\mathbf{k}} = \mathbf{k}$ ,  $\bar{\ell}_u = \ell_u - 2$  and  $\bar{\ell}_{u'} = \ell_{u'}$  for all other  $u' \in \mathcal{U}$ , and satisfying (2.15). For (b), we simply reduce  $\mathbb{E}(x_{u^*(u)} x_{u^*(u')})$  by 1, which results a graph in  $\Omega_{\bar{\mathbf{k}}, \bar{\ell}}$  with  $|\mathbf{k} - \bar{\mathbf{k}}| + |\ell - \bar{\ell}| = 2$  and the desired bound (2.15).

Since the above cases have covered all the possibilities for  $\Gamma \in \Omega_{\mathbf{k}, \ell} \setminus \Omega^*$ , we have completed the proof of the proposition.  $\square$

The following proposition is then a simple consequence.

**Proposition 2.9** *There exist  $C > 0$  and  $L > 0$  depending on  $\Lambda$ ,  $K$ ,  $m$ , and  $r$  only such that for every  $\theta \in \mathbf{R}$ , every location  $\mathbf{x} \in (\mathbf{R}^d)^K$ , and every  $\varepsilon \in (0, 1)$ , we have the bound*

$$\left| \mathbf{E} \prod_{j=1}^K \partial_{\theta}^r \mathcal{H}_m \left( e^{i\theta X_j} \right) \right| \leq C \max_{\Gamma \in \Omega^*} |\Gamma|. \quad (2.16)$$

**Remark 2.10** The bound is completely independent of  $\theta$  and  $\varepsilon$ , and its dependence on the location of  $\mathbf{x}$  is via  $\Omega^*$  only. Also note that the clustering, and hence  $\Omega^*$ , depends on the choice of  $L$ .

**Proof of Proposition 2.9** Note that graphs in  $\Omega_{\mathbf{k}, \ell}$  have degree  $|\mathbf{k}| + |\ell| + m|\mathcal{S}|$ , so by Proposition 2.8, there exists  $\Gamma^* \in \Omega^*$  such that

$$\max_{\Gamma \in \Omega_{\mathbf{k}, \ell}} |\Gamma| \leq (CL^{\frac{\alpha}{2}})^{\deg(\Gamma^*)} \cdot (CL^{-\alpha})^{\frac{1}{2}(|\mathbf{k}| + |\ell| + m|\mathcal{S}|)} \cdot |\Gamma^*|$$

for all  $\mathbf{k}$  and  $\ell$ . Combining it with Proposition 2.4 and (2.14), we get

$$\left| \mathbf{E} \prod_{j=1}^K \partial_{\theta}^r \mathcal{H}_m \left( e^{i\theta X_j} \right) \right| \leq CL^{\frac{\alpha}{2} \deg(\Gamma^*)} \cdot \exp \left( -\frac{\theta^2}{2\Lambda} + \frac{C_0 \langle \theta \rangle^2}{L^{\alpha}} \right) \cdot \max_{\Gamma^* \in \Omega^*} |\Gamma^*| \quad (2.17)$$

for some constant  $C_0$ . One can choose  $L$  sufficiently large depending on  $C_0$  and  $\Lambda$  only such that

$$\frac{1}{L^{\alpha}} < \frac{1}{4C_0\Lambda}. \quad (2.18)$$

This guarantees that the exponential term is uniformly bounded in  $\theta$ . Since  $C_0$  depends on  $\Lambda$ ,  $K$ ,  $m$ , and  $r$  only, so does  $L$ . Finally,  $L^{\frac{\alpha}{2} \deg(\Gamma^*)}$  is also uniformly bounded since graphs in  $\Omega^*$  have degrees at most  $(m+1)K$ . This completes the proof.  $\square$

**Remark 2.11** The reason why we need to choose  $L$  large is to ensure the exponential and hence the whole right-hand side of (2.17) being uniformly bounded in  $\theta$ . As we see now, the Gaussian factor  $e^{-\frac{\theta^2}{2\Lambda}}$  in (2.11) allows us to choose such  $L$  being independent of  $\theta$ . Together with the enhancement procedure in Sect. 2.6, this ensures the bounds in Proposition 2.9 and hence in Theorem 1.1 are completely independent of  $\theta$ .

Without the Gaussian factor, one would need to take  $L$  quadratic in  $\theta$  to make the exponential in (2.17) bounded, and the enhancement procedure in below would produce a bound that is polynomial in  $\theta$  with its degree depending on  $m$  and  $K$ .

## 2.6 Enhancement and Conclusion of the Proof

From now on, we fix the choice of  $L$  in (2.18). We need to control the right-hand side of (2.16) by that of (1.5). To achieve this, we enhance every  $\Gamma \in \Omega^*$  to a graph  $\text{Enh}(\Gamma)$



where  $\deg(x_j) \in \{m, m+1\}$  for every  $j \in [K]$ , which matches the pairing occurring in the desired upper bound. The enhancement procedure will also be performed in such a way that  $|\text{Enh}(\Gamma)|$  is an upper bound for  $|\Gamma|$  up to some proportionality constant, which is uniform in  $\varepsilon$ ,  $\theta$ , and  $\mathbf{x}$  subject to  $|\mathcal{S}| \geq 1$ . This will lead to bound (1.5). The procedure is similar to the one used to obtain (2.6) when  $|\mathcal{S}| = 0$ .

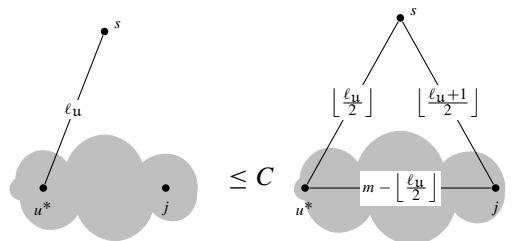
Fix  $\Gamma \in \Omega^*$  arbitrary, so in particular,  $\Gamma \in \Omega_{\mathbf{k}, \ell}$  for some  $\mathbf{k} \in \mathbf{N}^{\mathcal{S}}$  and  $\ell \in \mathbf{N}^{\mathcal{U}}$ . By the definition of  $\Omega^*$ ,  $\deg(x_s) \in \{m, m+1\}$  for all  $s \in \mathcal{S}$ . For every  $u \in \mathcal{U}$ , we have  $\deg(x_{u^*}) = \ell_u \leq m+1$ , and all of them are connected to one single  $x_s$  for some  $s \in \mathcal{S}$  if  $\ell_u \geq 1$ . All other points in  $u$  have degree 0. To construct  $\text{Enh}(\Gamma)$ , we add new edges to vertices in  $u \in \mathcal{U}$  and also move around existing edges, but keep  $\deg(x_s)$  unchanged for all  $s \in \mathcal{S}$  throughout the procedure. We do this cluster by cluster and write  $u^* = u^*(u)$  for simplicity.

Fix  $u \in \mathcal{U}$  arbitrary. To perform the enhancement operation for  $u$ , we let  $s \in \mathcal{S}$  be such that  $x_s$  is the unique singleton point connected to  $x_{u^*(u)}$  if  $\ell_u \geq 1$ . This also includes  $\ell_u = 0$ , in which case  $s$  could be arbitrary. We distinguish several situations depending on the number of points in  $u$ .

*Case 1.*  $|u| = 2$ . Let  $j \neq u^*(u)$  denote the other point in  $u$ . By definition of  $\Omega^*$ , we have  $\deg(x_j) = 0$ . We then perform the following operations. We move  $\lfloor (\ell_u + 1)/2 \rfloor$  of the  $\ell_u$  edges between  $x_s$  and  $x_{u^*}$  to connecting  $x_s$  and  $x_j$  and add  $m - \ell_u$  edges between  $x_{u^*}$  and  $x_j$ . By clustering, we have  $|x_{u^*} - x_j| \leq L\varepsilon$  and  $|x_s - x_j| \leq 2|x_s - x_{u^*}|$ . Hence, Proposition 2.1 gives the bounds

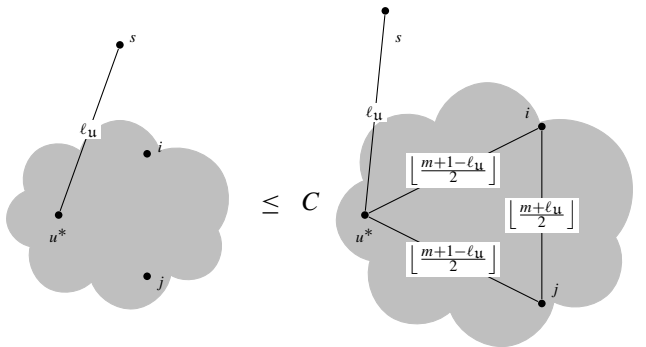
$$\begin{aligned} (\mathbf{E}X_s X_{u^*})^{\ell_u} &\leq C(\mathbf{E}X_s X_{u^*})^{\lfloor \frac{\ell_u}{2} \rfloor} (\mathbf{E}X_s X_j)^{\lfloor \frac{\ell_u+1}{2} \rfloor}, \\ 1 &\leq C((L+1)^\alpha \mathbf{E}(X_{u^*} X_j))^{m - \lfloor \frac{\ell_u}{2} \rfloor}, \end{aligned}$$

where  $L$  as chosen in (2.18) is independent of  $\theta$ ,  $\varepsilon$ , and  $\mathbf{x}$ . So in graphic notation, the above operation gives



where the gray area indicates the cluster  $u$ , and we have omitted drawing the remaining  $(m - \ell_u)$  or  $(m + 1 - \ell_u)$  edges from  $x_s$ . We also drop  $\lfloor \cdot \rfloor$  and simply use the graph itself to denote its value. Then,  $\deg(x_s) = m$  or  $m + 1$  is unchanged in the procedure. Furthermore, we have  $\deg(x_{u^*}) = m$  and  $\deg(x_j) \in \{m, m+1\}$  after the operation. This also includes the situation  $\ell_u = 0$ .

Case 2.  $|u| = 3$ . Let  $i, j$  denote the two other points in  $u$ . We then perform the operation



We see  $\deg(x_s)$  is unchanged. One can also check that  $\deg(x_{u^*(u)}) = m$  or  $m + 1$ , and  $\deg(x_i) = \deg(x_j) = m$ . So we have the correct degrees of the vertices as well as the desired bound.

Case 3.  $|u| \geq 4$ . We denote the other  $|u| - 1$  points in the cluster by  $j_1, \dots, j_{|u|-1}$ . For  $|u| - 2$  of them, say  $x_{j_1}, \dots, x_{j_{|u|-2}}$ , we perform the same operation as in Sect. 2.1 by cyclically connecting them with edges of multiplicities  $\lfloor \frac{m+1}{2} \rfloor$ . This yields the bound

$$1 \leq C (\mathbf{E}(X_{j_1} X_{j_2}) \cdots \mathbf{E}(X_{j_{|u|-3}} X_{j_{|u|-2}}) \mathbf{E}(X_{j_{|u|-2}} X_{j_1}))^{\lfloor \frac{m+1}{2} \rfloor}.$$

For the remaining points  $u^*$  and  $j_{|u|-1}$ , we perform the same operation as in Case 1 above. This again raises the degrees of all points in  $u$  to  $m$  or  $m + 1$  with a desired bound.

Every cluster  $u \in \mathcal{U}$  falls into one of the above three cases. The graph  $\text{Enh}(\Gamma)$  is obtained by performing the above operations to all  $u \in \mathcal{U}$ . It is clear from the bounds in the above three situations that there exists  $C > 0$  such that

$$|\Gamma| \leq C |\text{Enh}(\Gamma)|.$$

It is also straightforward to check that in  $\text{Enh}(\Gamma)$ , we have  $\deg(x_j) \in \{m, m + 1\}$  for all  $j \in [K]$ , and hence it represents one of the pairings from the expectation

$$\mathbf{E} \prod_{j=1}^K \left( X_j^{\diamond m} + X_{j+1}^{\diamond(m+1)} \right).$$

Hence, we deduce there exists  $C > 0$  such that

$$|\Gamma| \leq C |\text{Enh}(\Gamma)| \leq C \mathbf{E} \prod_{j=1}^K \left( X_j^{\diamond m} + X_{j+1}^{\diamond(m+1)} \right) \quad (2.19)$$

for all  $\theta, \varepsilon$ , and  $\mathbf{x}$ , and this is true for all  $\Gamma \in \Omega^*$ . Combining (2.19) and Proposition 2.9, we obtain the bound (1.5) in the case  $|\mathcal{S}| \geq 1$ .

Since the bound when  $\mathcal{S} = \emptyset$  has already been established in (2.6), we have thus completed the proof of Theorem 1.1.

### 3 Convergence of the Fields—Proof of Theorem 1.4

We are now ready to prove Theorem 1.4. For notational simplicity, we write  $A \lesssim_\alpha B$  to denote that  $A \leq CB$ , where the constant  $C$  depends only on the parameter(s) in the subscripts of the symbol  $\lesssim$  (and in this case  $\alpha$ ).

In order to apply the bound in Theorem 1.1, we use the convention for  $\widehat{F}$  such that

$$F(x) = \int_{\mathbf{R}} \widehat{F}(\theta) e^{i\theta x} d\theta.$$

But this only appears in intermediate steps, and the final statement does not depend on the definition of  $\widehat{F}$ . For every  $\varphi : \mathbf{R}^d \rightarrow \mathbf{R}$ , every  $x \in \mathbf{R}^d$ , and every  $\lambda > 0$ , we let

$$\varphi_x^\lambda(y) = \lambda^{-|\mathbf{s}|} \varphi\left(\frac{y_1 - x_1}{\lambda^{s_1}}, \dots, \frac{y_d - x_d}{\lambda^{s_d}}\right).$$

Recall that  $\Psi_\varepsilon = \rho_\varepsilon * \Psi$ , where  $\rho_\varepsilon$  is the rescaled mollifier. Also recall the form of  $a_m$  in (1.7). We first give the convergence criterion.

**Proposition 3.1** *Let  $\kappa > 0$  and  $m < \frac{|\mathbf{s}|}{\alpha}$ . If for every compact  $\mathcal{K} \subset \mathbf{R}^d$  and every  $n \in \mathbf{N}$ , we have*

$$\sup_{\lambda \in (\varepsilon, 1)} \sup_{x \in \mathcal{K}} \sup_{\substack{\varphi \in C_c^\infty(\mathcal{K}): \\ \|\varphi\|_{C^{\frac{m\alpha}{2}}} \leq 1}} \lambda^{\frac{m\alpha}{2} + \kappa} \left( \mathbf{E} \left| \langle \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m \left( F(\varepsilon^{\frac{\alpha}{2}} \Psi_\varepsilon) \right) - a_m \Psi^{\diamond m}, \varphi_x^\lambda \rangle |^{2n} \right)^{\frac{1}{2n}} \rightarrow 0 \quad (3.1)$$

as  $\varepsilon \rightarrow 0$ , then  $\varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\varepsilon^{\frac{\alpha}{2}} \Psi_\varepsilon)) \rightarrow a_m \Psi^{\diamond m}$  in  $\mathcal{C}^{-\frac{m\alpha}{2} - \kappa'}$  for every  $\kappa' > \kappa$ .

The proof of the proposition is standard Kolmogorov's criterion, and so it remains to prove (3.1). By stationarity, we can simply restrict to the case  $x = 0$  in (3.1). Writing  $\|\cdot\|_{2n} := (\mathbf{E}|\cdot|^{2n})^{\frac{1}{2n}}$  as well as  $\Phi_\varepsilon = \varepsilon^{\frac{\alpha}{2}} \Psi_\varepsilon$ , we need to show for all small  $\kappa$  that

$$\lambda^{\frac{m\alpha}{2} + \kappa} \|\langle \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m \Psi^{\diamond m}, \varphi^\lambda \rangle\|_{2n} \rightarrow 0 \quad (3.2)$$

as  $\varepsilon \rightarrow 0$ , uniformly over  $\lambda \in (\varepsilon, 1)$  and smooth  $\varphi$  supported in a ball of radius 1 such that  $\|\varphi\|_{C^{\frac{m\alpha}{2}}} \leq 1$ . The rest of the section is devoted to the proof of (3.2).

Since  $\Psi$  is stationary, so is  $\Psi_\varepsilon = \rho_\varepsilon * \Psi$ . Hence,  $\Phi_\varepsilon$  has stationary Gaussian distribution  $\mu_\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$  with

$$\sigma_\varepsilon^2 = \mathbf{E}\Phi_\varepsilon^2 = \varepsilon^\alpha \int G(x - y) \rho_\varepsilon(x) \rho_\varepsilon(y) dx dy, \quad (3.3)$$

where  $G$  is the correlation function of  $\Psi$  as in Assumption 1.2. The coefficient of the  $m$ -th term in the chaos expansion of  $F(\Phi_\varepsilon)$  is given by

$$a_m^{(\varepsilon)} = \frac{1}{m!} \left( F^{(m)} * \mu_\varepsilon \right) (0).$$

We split the difference  $\varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m \Psi^{\diamond m}$  into three parts by

$$\begin{aligned} \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m \Psi^{\diamond m} &= \left( \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m^{(\varepsilon)} \Psi_\varepsilon^{\diamond m} \right) \\ &\quad + a_m^{(\varepsilon)} (\Psi_\varepsilon^{\diamond m} - \Psi^{\diamond m}) + (a_m^{(\varepsilon)} - a_m) \Psi^{\diamond m}, \end{aligned} \quad (3.4)$$

and we show that each of them satisfies a bound of the form of (3.2).

The latter two terms are simpler. For the second one, we notice that  $m < \frac{|\mathbf{s}|}{\alpha}$  ensures  $\Psi^{\diamond m}$  is well defined, and for all sufficiently small  $\kappa$ , we have

$$\| \langle \Psi_\varepsilon^{\diamond m} - \Psi^{\diamond m}, \varphi^\lambda \rangle \|_{2n} \lesssim_n \varepsilon^\kappa \lambda^{-\frac{m\alpha}{2} - \kappa}$$

uniformly over  $\lambda \in (\varepsilon, 1)$ . Also, since  $a_m^{(\varepsilon)}$  is uniformly bounded in  $\varepsilon$ , it then follows immediately that

$$\lambda^{\frac{m\alpha}{2} + \kappa} \| \langle \Psi_\varepsilon^{\diamond m} - \Psi^{\diamond m}, \varphi^\lambda \rangle \|_{2n} \rightarrow 0$$

as  $\varepsilon \rightarrow 0$ , uniformly over  $\lambda \in (\varepsilon, 1)$  and  $\varphi$  in the range required in Proposition 3.1.

For the third term, it suffices to notice that assumption (1.6) on  $G$  guarantees that  $\sigma_\varepsilon^2 \rightarrow \sigma^2$ , where  $\sigma_\varepsilon^2$  and  $\sigma^2$  are given by (3.3) and (1.8). Hence, we immediately have  $a_m^{(\varepsilon)} \rightarrow a_m$ . The desired bound of the form (3.2) then follows immediately from the boundedness of  $\Psi^{\diamond m}$  in  $\mathcal{C}^{-\frac{m\alpha}{2} - \kappa}$ .

We now turn to the first term on the right-hand side of (3.4), which requires the use of the bound in Theorem 1.1. We first note that the covariance of  $\Phi_\varepsilon = \varepsilon^{\frac{\alpha}{2}} \rho_\varepsilon * \Psi$  has the form

$$\mathbf{E}(\Phi_\varepsilon(x) \Phi_\varepsilon(y)) = (\rho_\varepsilon^{\star 2} \star G)(x - y),$$

where  $\star$  denotes the forward convolution in the sense that  $(f \star g)(x) = \int f(x + y)g(y)dy$ . Assumption 1.2 on  $G$  guarantees that  $\Phi_\varepsilon$  satisfies the assumption (1.4) in Theorem 1.1 with some  $\Lambda > 1$ . Since  $\Psi_\varepsilon^{\diamond m} = \varepsilon^{-\frac{m\alpha}{2}} \Phi_\varepsilon^{\diamond m}$ , and  $a_m^{(\varepsilon)}$  is precisely the  $m$ -th coefficient in the chaos expansion of  $F(\Phi_\varepsilon)$ , we have

$$\varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m^{(\varepsilon)} \Psi_\varepsilon^{\diamond m} = \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_{m+1}(F(\Phi_\varepsilon)). \quad (3.5)$$

We leave aside the factor  $\varepsilon^{-\frac{m\alpha}{2}}$  and focus on  $\mathcal{H}_{m+1}(F(\Phi_\varepsilon))$  for a moment. Fourier expanding  $F$  and changing the order of integration, we get the identity

$$\langle \mathcal{H}_{m+1}(F(\Phi_\varepsilon)), \varphi^\lambda \rangle = \langle \widehat{F}, \mathcal{A}\Phi_\varepsilon \rangle_\theta,$$

where

$$(\mathcal{A}\Phi_\varepsilon)(\theta) = (\mathcal{A}_{\varphi, \lambda, m} \Phi_\varepsilon)(\theta) = \int_{\mathbf{R}^d} \mathcal{H}_{m+1} \left( e^{i\theta \Phi_\varepsilon(x)} \right) \varphi^\lambda(x) dx, \quad (3.6)$$

and the subscript  $\theta$  on the inner product indicates that the testing is taken with respect to the Fourier variable  $\theta \in \mathbf{R}$ . We now omit the subscripts in  $\mathcal{A}$  for simplicity. Recall that  $\mathcal{I}_k = (k - 1, k + 1)$ . Multiplying  $\mathcal{A}\Phi_\varepsilon$  by a partition of unity subordinate to the intervals  $\{\mathcal{I}_k\}_{k \in \mathbf{Z}}$  and separating these terms, we get the bound

$$|\langle \mathcal{H}_{m+1}(F(\Phi_\varepsilon)), \varphi^\lambda \rangle| \leq C_M \sum_{k \in \mathbf{Z}} \|\widehat{F}\|_{M, \mathcal{I}_k} \sup_{0 \leq r \leq M} \sup_{\theta \in \mathcal{I}_k} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|,$$

where  $M \in \mathbf{N}$  is as in Assumption 1.3. Now, taking  $2n$ -th moments on both sides and using triangle inequality, we get

$$\|\langle \mathcal{H}_{m+1}(F(\Phi_\varepsilon)), \varphi^\lambda \rangle\|_{2n} \leq C_M \sum_{k \in \mathbf{Z}} \|\widehat{F}\|_{M, \mathcal{I}_k} \left( \mathbf{E} \sup_{0 \leq r \leq M} \sup_{\theta \in \mathcal{I}_k} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n} \right)^{\frac{1}{2n}}. \quad (3.7)$$

There are two suprema inside the expectation. The first supremum is taken over  $M + 1$  elements, so we can move it out of  $(\mathbf{E}|\cdot|^{2n})^{\frac{1}{2n}}$  at the cost of a constant multiple depending on  $M$  and  $n$  only. The second one is taken over an interval, so we need the following lemma to interchange it with the expectation.

**Lemma 3.2** *Suppose  $f$  is a random  $\mathcal{C}^1$  function on an interval  $\mathcal{I}$ . For every  $p \geq 1$ , there exists  $C$  depending on  $p$  and  $|\mathcal{I}|$  only such that*

$$\mathbf{E} \sup_{\theta \in \mathcal{I}} |f(\theta)|^p \leq C_{p, |\mathcal{I}|} \sup_{\theta \in \mathcal{I}} \mathbf{E} (|f(\theta)|^p + |f'(\theta)|^p).$$

**Proof** Fix  $\theta_0 \in \mathcal{I}$  arbitrary. By fundamental theorem of calculus and Hölder's inequality, we have

$$|f(\theta)| \leq |f(\theta_0)| + |\mathcal{I}|^{1-\frac{1}{p}} \left( \int_{\mathcal{I}} |f'(x)|^p dx \right)^{\frac{1}{p}}.$$

Raising both sides to  $p$ -th power, we get

$$\sup_{\theta} |f(\theta)|^p \leq C \left( |f(\theta_0)|^p + \int_{\mathcal{I}} |f'(\theta)|^p d\theta \right),$$

where  $C$  depends on  $p$  and  $\mathcal{I}$ . The assertion then follows by taking expectation on both sides and noting that

$$\mathbf{E} \int_{\mathcal{I}} |f'(\theta)|^p d\theta = \int_{\mathcal{I}} \mathbf{E} |f'(\theta)|^p d\theta \leq |\mathcal{I}| \sup_{\theta \in \mathcal{I}} \mathbf{E} |f'(\theta)|^p.$$

This completes the proof of the lemma.  $\square$

Using Lemma 3.2 to interchange the expectation and supremum, we have

$$\mathbf{E} \sup_{0 \leq r \leq M} \sup_{\theta \in \mathcal{I}_k} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n} \lesssim_{n, M} \sup_{0 \leq r \leq M+1} \sup_{\theta \in \mathcal{I}_k} \mathbf{E} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n},$$

where the supremum over  $r$  is taken over  $r \leq M + 1$  to include the one additional derivative required in the interchange. Plugging it back into the right-hand side of (3.7), we obtain

$$\|\langle \mathcal{H}_{m+1}(F(\Phi_\varepsilon)), \varphi^\lambda \rangle\|_{2n} \lesssim_{n,M} \sum_{k \in \mathbb{Z}} \|\widehat{F}\|_{M, \mathcal{I}_k} \sup_{0 \leq r \leq M+1} \sup_{\theta \in \mathcal{I}_k} \left( \mathbf{E} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n} \right)^{\frac{1}{2n}}. \quad (3.8)$$

It then remains to control the quantity  $\mathbf{E} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n}$ . Recalling the expression of  $\mathcal{A}\Phi_\varepsilon$  in (3.6), we have

$$\mathbf{E} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n} = \int_{(\mathbb{R}^d)^{2n}} \left( \mathbf{E} \prod_{j=1}^{2n} \partial_\theta^r \mathcal{H}_{m+1} \left( e^{i\theta \Phi_\varepsilon(x_j)} \right) \right) \cdot \left( \prod_{j=1}^{2n} \varphi^\lambda(x_j) \right) d\mathbf{x},$$

where we used the shorthand notation  $\mathbf{x} = (x_1, \dots, x_{2n})$ . We now apply Theorem 1.1 to the expectation part above, so that we get

$$\mathbf{E} \prod_{j=1}^{2n} \partial_\theta^r \mathcal{H}_{m+1} \left( e^{i\theta \Phi_\varepsilon(x_j)} \right) \lesssim_r \mathbf{E} \prod_{j=1}^{2n} \left( \Phi_\varepsilon^{\diamond(m+1)}(x_j) + \Phi_\varepsilon^{\diamond(m+2)}(x_j) \right).$$

Plugging it into the integral on the right-hand side above and using the identity

$$\begin{aligned} & \int \left[ \mathbf{E} \prod_{j=1}^{2n} \left( \Phi_\varepsilon^{\diamond(m+1)}(x_j) + \Phi_\varepsilon^{\diamond(m+2)}(x_j) \right) \right] \cdot \left( \prod_{j=1}^{2n} \varphi^\lambda(x_j) \right) d\mathbf{x} \\ &= \mathbf{E} |\langle \Phi_\varepsilon^{\diamond(m+1)} + \Phi_\varepsilon^{\diamond(m+2)}, \varphi^\lambda \rangle|^{2n}, \end{aligned}$$

we get

$$\left( \mathbf{E} |(\mathcal{A}\Phi_\varepsilon)^{(r)}(\theta)|^{2n} \right)^{\frac{1}{2n}} \lesssim_{r,n} \sum_{\ell=1}^2 \left( \mathbf{E} |\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^{2n} \right)^{\frac{1}{2n}}. \quad (3.9)$$

In particular, the bound is uniform in both  $\varepsilon$  and  $\theta$ . We have the following lemma controlling the right-hand side above.

**Lemma 3.3** *For every integer  $\ell \geq 1$  and every sufficiently small  $\kappa$ , we have*

$$\left( \mathbf{E} |\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^{2n} \right)^{\frac{1}{2n}} \lesssim_{n,\ell} \varepsilon^{\frac{m\alpha}{2} + \kappa} \lambda^{-\frac{m\alpha}{2} - \kappa}$$

uniformly over  $\varepsilon, \lambda \in (0, 1)$ .

**Proof** Since  $\Phi_\varepsilon$  is Gaussian, by equivalence of moments, the left-hand side above can be controlled by

$$\left( \mathbf{E} |\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^{2n} \right)^{\frac{1}{2n}} \lesssim_n \left( \mathbf{E} |\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^2 \right)^{\frac{1}{2}},$$

so we only need to bound the second moment. By Wick's formula, we have

$$\mathbf{E}|\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^2 = (m+\ell)! \int (\mathbf{E}\Phi_\varepsilon(x)\Phi_\varepsilon(y))^{m+\ell} \varphi^\lambda(x)\varphi^\lambda(y) dx dy.$$

Since  $\ell \geq 1$ , we have

$$(\mathbf{E}\Phi_\varepsilon(x)\Phi_\varepsilon(y))^{m+\ell} \leq \frac{\Lambda^{m+\ell} \varepsilon^{\alpha(m+\ell)}}{(|x-y|+\varepsilon)^{\alpha(m+\ell)}} \leq \frac{\Lambda^{m+\ell} \varepsilon^{m\alpha+\kappa}}{|x-y|^{m\alpha+\kappa}}$$

for all  $\kappa \in (0, \ell\alpha)$ . For all  $\kappa$  sufficiently small such that  $m\alpha + \kappa < |\mathfrak{s}|$ , the singularity on the right-hand side above is integrable, so we have

$$\mathbf{E}|\langle \Phi_\varepsilon^{\diamond(m+\ell)}, \varphi^\lambda \rangle|^2 \lesssim \varepsilon^{m\alpha+\kappa} \int \frac{\varphi^\lambda(x)\varphi^\lambda(y)}{|x-y|^{m\alpha+\kappa}} dx dy \lesssim \varepsilon^{m\alpha+\kappa} \lambda^{-m\alpha-\kappa}.$$

The proof is complete by taking square roots on both sides and replacing  $\kappa$  by  $2\kappa$ .  $\square$

Now, combining (3.8), (3.9) and Lemma 3.3, and using Assumption 1.3 on  $F$ , we get

$$\begin{aligned} \|\langle \mathcal{H}_{m+1}(F(\Phi_\varepsilon)), \varphi^\lambda \rangle\|_{2n} &\lesssim_{n,M} \varepsilon^{\frac{m\alpha}{2}+\kappa} \lambda^{-\frac{m\alpha}{2}-\kappa} \sum_{k \in \mathbb{Z}} \|\widehat{F}\|_{M, \mathcal{I}_k} \\ &\lesssim_{n,M} \varepsilon^{\frac{m\alpha}{2}+\kappa} \lambda^{-\frac{m\alpha}{2}-\kappa}. \end{aligned}$$

Substituting the above bound back to (3.5), we deduce that

$$\lambda^{\frac{m\alpha}{2}+\kappa} \|\langle \varepsilon^{-\frac{m\alpha}{2}} \mathcal{H}_m(F(\Phi_\varepsilon)) - a_m^{(\varepsilon)} \Psi_\varepsilon^{\diamond m}, \varphi^\lambda \rangle\|_{2n} \lesssim_{n,M} \varepsilon^\kappa,$$

which is the desired bound. The proof of Theorem 1.4 is thus complete.

**Acknowledgements** The author acknowledges the support from the Engineering and Physical Sciences Research Council through the fellowship EP/N021568/1. I also thank the anonymous referee for carefully reading the draft version of the article and providing helpful suggestions on improving the presentation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bertini, L., Giacomin, G.: Stochastic burgers and KPZ equations from particle systems. Commun. Math. Phys. **183**(3), 571–607 (1997). <https://doi.org/10.1007/s002200050044>
- Furlan, M., Gubinelli, M.: Weak universality for a class of 3D stochastic reaction-diffusion models. ArXiv e-prints (2017). [arXiv:1708.03118](https://arxiv.org/abs/1708.03118)
- Friz, P.K., Hairer, M.: A Course on Rough Paths. Universitext. With an Introduction to Regularity Structures. Springer, Cham (2014). <https://doi.org/10.1007/978-3-319-08332-2>
- Goncalves, P., Jara, M.: Nonlinear fluctuations of weakly asymmetric interacting particle systems. Arch. Rational Mech. Anal. **212**(2), 597–644 (2014). <https://doi.org/10.1007/s00205-013-0693-x>. [arXiv:1309.5120](https://arxiv.org/abs/1309.5120)

5. Gubinelli, M., Perkowski, N.: The Hairer–Quastel universality result at stationarity. In: Stochastic analysis on large scale interacting systems, RIMS Kôkyûroku Bessatsu, vol. B59, pp. 101–115. Research Institute for Mathematical Sciences (RIMS), Kyoto (2016)
6. Gubinelli, M., Perkowski, N.: KPZ reloaded. *Commun. Math. Phys.* **349**(1), 165–269 (2017). <https://doi.org/10.1007/s00220-016-2788-3>. [arXiv:1508.03877](https://arxiv.org/abs/1508.03877)
7. Gubinelli, M., Perkowski, N.: Energy solutions of KPZ are unique. *J. Am. Math. Soc.* **312**, 427–471 (2018). <https://doi.org/10.1090/jams/889>. [arXiv:1508.07764](https://arxiv.org/abs/1508.07764)
8. Gubinelli, M.: A panorama of singular SPDEs. *Proc. Int. Cong. Math.* **2**, 2277–2304 (2018)
9. Hairer, M.: Solving the KPZ equation. *Ann. Math.* **1782**, 559–664 (2013). <https://doi.org/10.4007/annals.2013.178.2.4>. [arXiv:1109.6811](https://arxiv.org/abs/1109.6811)
10. Hairer, M.: A theory of regularity structures. *Invent. Math.* **2014**, 269–504 (1982). <https://doi.org/10.1007/s00222-014-0505-4>. [arXiv:1303.5113](https://arxiv.org/abs/1303.5113)
11. Hairer, M.: Renormalisation of parabolic stochastic PDEs. *Jpn. J. Math.* **13**(2), 187–233 (2018)
12. Hairer, M., Quastel, J.: A class of growth models rescaling to KPZ. *ArXiv e-prints* (2015). [arxiv:1512.07845](https://arxiv.org/abs/1512.07845)
13. Hairer, M., Xu, W.: Large-scale limit of interface fluctuation models. *Ann Probab.* *ArXiv e-prints* (2018). [arXiv:1802.08192](https://arxiv.org/abs/1802.08192)
14. Kardar, M., Parisi, G., Zhang, Y.-C.: Dynamical scaling of growing interfaces. *Phys. Rev. Lett.* **56**, 889–892 (1986). <https://doi.org/10.1103/PhysRevLett.56.889>
15. Lodhia, A., Sheffield, S., Sun, X., Watson, S.: Fractional Gaussian fields: a survey. *Probab. Surv.* **13**, 1–56 (2016)
16. Nualart, D.: The Malliavin Calculus and Related Topics. Probability and its Applications (New York), 2nd edn. Springer, Berlin (2006)
17. Shen, H., Xu, W.: Weak universality of dynamical  $\phi_3^4$ : non-Gaussian noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **62**, 211–254 (2018). [arXiv:1601.05724](https://arxiv.org/abs/1601.05724)