# LARGE SCALE LIMIT OF INTERFACE FLUCTUATION MODELS

By Martin Hairer[*,1] and Weijun Xu[†,‡,2]

*Imperial College London*, *University of Oxford*[†] *and NYU Shanghai*[‡]

We extend the weak universality of KPZ in Hairer and Quastel [*Forum Math. Pi* **6** (2018) e3] to weakly asymmetric interface models with general growth mechanisms beyond polynomials. A key new ingredient is a pointwise bound on correlations of trigonometric functions of Gaussians in terms of their polynomial counterparts. This enables us to reduce the problem of a general nonlinearity with sufficient regularity to that of a polynomial.

## CONTENTS

## 1. Introduction.

1.1. *Weak universality of KPZ.* The weak universality conjecture for KPZ states that any "reasonable" weakly asymmetric interface fluctuation model in $1 + 1$ dimensions should rescale to the KPZ equation [28], formally given by

$$(1.1) \qquad \partial_t h = \partial_x^2 h + a(\partial_x h)^2 + \xi.$$

Here, $\xi$ denotes space–time white noise on the one-dimensional torus **T**, and $a \in \mathbf{R}$ is a coupling constant describing the strength of the asymmetry. The term "reasonable" refers to the following three features in the microscopic model:

- There is a smoothing mechanism which erodes high peaks and fills deep valleys.
- The fluctuation mechanism depends on the slope of the interface in a nontrivial way.
- The system is influenced by a random fluctuation with short range correlations.

These features are clearly visible in the macroscopic equation (1.1), represented by the Laplacian, the nonlinearity $(\partial_x h)^2$ and the noise $\xi$, respectively. The additional requirement, namely that of the microscopic model being "weakly asymmetric", is also essential for its large scale limit to be given by the KPZ equation. It refers to that the strength of the growth at microscopic level should be very weak, and is tuned according to the scale at which one looks at the system. The presence of this weak parameter features the so-called "crossover regime".

In fact, for completely symmetric models (i.e., the strength of the growth is 0), it is widely believed that their scaling limit is described by the stochastic heat equation, so that they belong to the Edwards–Wilkinson universality class [12], which exhibits Gaussian fluctuations at large scales. On the other hand, if the strength of asymmetry is of order 1, then it is believed that the large scale behaviour of such models is described by the "KPZ fixed point" and they are said to belong to the KPZ universality class. Very recently, the breakthrough by [30] established the

convergence of TASEP to the KPZ fixed point, which in particular yields a complete characterisation of the latter. There has also been substantial progress in the understanding of various statistics for other models in this class (see, e.g., [1, 2, 4] and references therein). So far, however, all these results strongly rely on the presence of a suitable "integrable structure" in the model, so that the underlying reason for the universal behaviour is still unclear from a mathematical perspective.

A natural related question is to investigate the crossover regime, that is, the object(s) that lie in between these two universality classes. This is where the KPZ equation comes in. The equation was first derived in [28] for height functions of droplets, and is expected to be a universal object for weakly asymmetric growth models at large scales. What is usually referred to as the "weak KPZ universality conjecture" can be interpreted as saying that the KPZ equation (1.1) is the only object that interpolates between the Edwards–Wilkinson and KPZ universality classes in the sense that solutions to the KPZ equations are expected to be the only stationary space–time Markov process that converges to the stochastic heat equation when "zooming in" and the KPZ fixed point when "zooming out". Unfortunately, such a characterisation of the KPZ equation appears to be far out of reach of current techniques. One possible step in this direction which has been investigated in recent years is to exhibit a class of models that is as large as possible and depends on a parameter $\varepsilon$ tuning their asymmetry in the sense that models with $\varepsilon = 0$ belong to the EW universality class and models with $\varepsilon \neq 0$ (are expected to) belong to the KPZ universality class. The weak universality conjecture then suggests that if one simultaneously sends $\varepsilon$ to 0 and considers such a system at a large scale $L(\varepsilon)$, then there exists a specific choice of $L$ such that one observes convergence to the KPZ equation as $\varepsilon \to 0$. We refer to the recent survey [23] for discussions in the more general context of crossover regimes between renormalisation group fixed points.

One mathematical obstacle towards the understanding of such a claim is that (1.1) is ill-posed. Nevertheless, much progress has been made towards this conjecture over the past years. The first rigorous statement was in the seminal work [3], where the authors showed that the height function of weakly asymmetric simple exclusion processes (WASEP) rescale to (1.1). The proof uses the Hopf–Cole transform of the KPZ equation, which amounts to showing that the exponential of the system converges to the multiplicative stochastic heat equation. There have been other recent works in this direction; see for example, [7–10, 29]. The systems considered in these works are all related to WASEP and rely on its various structures. Since most particle systems do not have the same structures, and they are usually not very well behaved under exponentiation, it is not clear how the methods employed in these works would generalise to other not so closely related situations.

In [16], the authors introduced the notion of an "energy solution" to the KPZ equation in equilibrium. A slightly stronger notion of solution was subsequently introduced in [20] and shown to be unique and coincide with the Hopf–Cole solution. This characterisation of the solutions to the KPZ equation was very fruitful in showing "weak universality" results as discussed above for a variety of models. In particular, this includes a large class of particle systems generalising WASEP but requiring much less structure [16], systems of interacting Brownian motion [11], as well as stochastic PDE models [18]. The only drawback of this technique is that it requires a priori knowledge of the invariant measure of the microscopic model. Furthermore, it only appears to cover Markovian models with the additional property that the decomposition $L = S + A$ of the generator into a symmetric and an antisymmetric part is such that both $A$ and $S$ are local operators (or at least "almost local").

In [21, 22], a pathwise notion of solution was developed based on the theory of rough paths/regularity structures. (See also [17, 19] for a pathwise approach using paracontrolled calculus that is quite different in its technical implementation but very similar in spirit.) This approach avoids using the Hopf–Cole transform as well as the use of the invariant measure, so that it covers non-Markovian models in particular. It also allows to show the stability of quite general discrete approximations [13], including some standard ones which had been shown to be stable earlier [24, 26]. In addition, a universality result was shown in [25] for quite a large class of continuous random PDE models with polynomial growth mechanism (see below for details). The present article pursues this line of investigation and shows that the polynomial growth mechanism that was essential in the proof of [25] can be replaced by an arbitrary sufficiently smooth function with subpolynomial growth at infinity.

1.2. *Main result.* In [25], the authors explored continuous microscopic growth models in weak asymmetry regime of the type

$$(1.2) \qquad \partial_t h = \partial_x^2 h + \sqrt{\varepsilon} F(\partial_x h) + \widehat{\xi},$$

where $F$ is an even polynomial and $\widehat{\xi}$ is a space–time Gaussian random field with smooth short range covariance that integrates to 1. They showed that the large scale behaviour of $h$, when properly rescaled and recentered, is described by the KPZ equation (1.1). One surprising fact revealed by this result is that the coupling constant $a$ in the limit depends on all coefficients of the polynomial $F$. The aim of this article is to remove the polynomial requirement on the nonlinearity. Our precise assumption on $F$ is the following.

ASSUMPTION 1.1. The function $F : \mathbf{R} \to \mathbf{R}$ is symmetric. Furthermore, there exist $\alpha \in (0, 1)$ and constants $C, M > 0$ such that $F \in \mathcal{C}^{7+\alpha}$ and satisfies the

bounds

$$\sup_{0\le\ell\le 7}\big|F^{(\ell)}(u)\big|\le C\big(1+|u|\big)^M,$$

$$\sup_{|h|<1}\frac{|F^{(7)}(u+h)-F^{(7)}(u)|}{|h|^\alpha}\le C\big(1+|u|\big)^M$$

for all $u\in\mathbf{R}$.

REMARK 1.2. The above assumption on $F$ implies certain decay of the local distributional norm of its Fourier transform, as stated in Proposition 4.9. In fact, the bounds (4.8) and (4.9) in that proposition are the only two properties of $F$ we will use.

One specific example of function $F$ that satisfies our assumptions is given by $F(u)=\sqrt{1+u^2}$. This function corresponds to the lateral growth, and is the most common one used in the literature for heuristic derivation of the KPZ equation.

ASSUMPTION 1.3. There exists a compactly supported smooth function $\rho\colon \mathbf{R}^2\to\mathbf{R}$ which is symmetric in its space variable $x$ and integrates to 1 (in space–time), and such that

$$(1.3)\qquad\qquad\qquad \widehat{\xi}\overset{\text{law}}{=}\xi*\rho,$$

where $\xi$ denotes space–time white noise.

REMARK 1.4. The assumption on the symmetry of $\rho$ in its space variable is mainly for technical convenience. Without the symmetry assumption, one also needs to shift the frame horizontally in order to establish the convergence. But that will not change the limiting equation, and we keep this symmetry assumption here for technical simplicity.

We now give our main result. Let $\widehat{\Psi}=P'*\widehat{\xi}$, where $P'$ denotes the spatial derivative of the heat kernel $P$, and $*$ is the convolution in space–time. Then $\widehat{\Psi}$ is stationary, and we use $\mu$ to denote its law at any space–time point. We define the constant $a$ to be

$$(1.4)\qquad\qquad a:=\frac{1}{2}\mathbf{E}F''(\widehat{\Psi})=\frac{1}{2}\int_{\mathbf{R}}F''(x)\mu(dx).$$

Let $h$ be the process (1.2), and define

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

Then $h_\varepsilon$ solves the equation

$$(1.5)\qquad\qquad \partial_t h_\varepsilon=\partial_x^2 h_\varepsilon+\varepsilon^{-1}F\big(\varepsilon^{\frac{1}{2}}\partial_x h_\varepsilon\big)+\xi_\varepsilon-C_\varepsilon,$$

where $\xi_\varepsilon(t, x) = \varepsilon^{-\frac{3}{2}} \widehat{\xi}(t/\varepsilon^2, x/\varepsilon)$ is an approximation (in law) to the space–time white noise $\xi$ at scale $\varepsilon$. There is a slight abuse of notation here since $h$ itself also depends on $\varepsilon$, but since we are mostly working with the rescaled process $h_\varepsilon$, the omission of $\varepsilon$ in the microscopic process should not create confusion.

From now, rather than considering (1.5) on all of $\mathbf{R}$, we will assume that the space variable takes values in the one-dimensional torus $\mathbf{T}$ of size 1. This means that the original equation (1.2), and hence the noise $\widehat{\xi}$ are actually defined on the torus of size $\varepsilon^{-1}$ and Assumption 1.3 should be interpreted accordingly. The constant $a$ defined in (1.4) however still uses the process defined on the whole space.

Returning to (1.5), if $F$ has sufficient regularity one can expand the nonlinearity as

$$\varepsilon^{-1} F\left(\varepsilon^{\frac{1}{2}} \partial_x h_\varepsilon\right) = \frac{a_0}{\varepsilon} + a_1 (\partial_x h_\varepsilon)^2 + a_2 \varepsilon (\partial_x h_\varepsilon)^4 + \cdots .$$

The term $\frac{a_0}{\varepsilon}$ can be killed by the choice of the large constant $C_\varepsilon$ in (1.5). Since each of the higher powers of $\partial_x h_\varepsilon$ ($\geq 4$) is multiplied by a positive power of $\varepsilon$, it seems that only the quadratic term would survive in the limit $\varepsilon \to 0$, and one might expect that $h_\varepsilon$ converges to the KPZ equation with coupling constant $a_1$. However, as already shown in [25] for polynomial $F$, this is simply *not* the case. The main result of this article is an extension of [25] to nonlinearities $F$ that satisfy Assumption 1.1. We state it as follows.

THEOREM 1.5. *Let $h_\varepsilon$ be the solution to (1.5) with initial data $h_0^{(\varepsilon)}$, where the nonlinearity $F$ satisfies Assumption 1.1. Suppose there exists $\eta \in (\frac{1}{2} - \frac{1}{M+4}, \frac{1}{2})$ and $h_0 \in \mathcal{C}^\eta(\mathbf{T})$ such that for some $\gamma \in (\frac{3}{2}, \frac{5}{3})$, $\|h_0^{(\varepsilon)}; h_0\|_{\gamma,\eta;\varepsilon} \to 0$ in the sense of (3.6). Then there exists $C_\varepsilon \to +\infty$ such that for every $T > 0$, $h_\varepsilon$ converges in probability in $\mathcal{C}^\eta([0, T], \mathbf{T})$ to the Hopf–Cole solution to the KPZ equation with coupling constant $a$ given by (1.4).*

REMARK 1.6. The large constant $C_\varepsilon$ is of the form $C_\varepsilon = \frac{\widehat{a}}{\varepsilon} + \mathcal{O}(1)$, where $\widehat{a} = \mathbf{E} F(\widehat{\Psi})$. This can be easily deduced by combining the definition of the renormalisation constants (5.1) and their behaviours in Section 5.3, the form of $C_\varepsilon$ in Theorem 3.8 and the convergence of renormalised models in Theorem 5.2.

In the case of polynomial $F$, one can see from (1.4) that $a$ is a linear combination of all its coefficients (except the zeroth term). This suggests that all the higher powers have contributions to the limit rather than simply vanishing. For the $F$ that we consider in this article, even though it is in general not infinitely differentiable to have a power series expansion, the combined effects of all these "higher powers" still exist, and is given explicitly by (1.4).

Similar universality questions have been studied in the context of $\Phi_3^4$ equation. In [27], the authors considered the 3D microscopic phase coexistence models of

the type

$$\partial_t u = \Delta u - \varepsilon V'_\theta(u) + \widehat{\xi}$$

for an even polynomial potential $V$ near a critical point $\theta = 0$ and a smooth Gaussian field $\widehat{\xi}$. It was shown that the large scale behaviour of the field $u$ was described by the $\Phi^4_3$ equation. This result was extended to non-Gaussian noise [31] and general even potential with Gaussian noise [15].

The difficulties of extending from polynomial to general nonlinearities are essentially the same in both situations: one needs to control arbitrary high moments of a general function of a Gaussian field (or a more general random field). The methods developed in this article and the ones in [15] however are very different. We include a brief discussion on this towards the end of Section 2.

1.3. *Possible generalisations.* We discuss two possible generalisations of the result in this article.

1. *Regularity of $F$.* Our assumption of $7+$ differentiability of $F$ comes from the form of the bound we develop in this article and it is clear that this requirement is not optimal. In fact, the expression (1.4) suggests that it may be possible to take any $F$ that is Lipschitz continuous and not growing too fast at infinity.

One very interesting example would be $F(x) = |x|$, but it is not clear at this stage whether our technique could be sharpened to include that case. Since there is a big gap between Assumption 1.1 and Lipschitzness, we expect that new ideas are needed to treat low regularity functions.

2. *Non-Gaussian noise.* The essential part of the paper where we use Gaussianity of the noise is in Section 6, where we develop a pointwise bound for correlations of trigonometric functions of Gaussian fields. If one is able obtain a similar bound for a class of non-Gaussian random fields, then the arguments in the rest part of the paper can be used in exactly same way. However, it is not clear at this moment how such bounds can be obtained for non-Gaussian noises.

1.4. *Notation.* Throughout the article, we use $M$ to denote the growth of the derivatives of $F$ (in Assumption 1.1). For every random variable $X$, we write $\{X\} = X - \mathbf{E}X$ as its recentered version. We use $\diamond$ to denote Wick products between Gaussians, for example, $X^{\diamond k_1} \diamond Y^{\diamond k_2} \diamond Z^{\diamond k_3}$. We use $A \sqcup B$ to denote the union of two *disjoint* sets $A$ and $B$. Finally, we define the Fourier transform of $G$ such that

$$(1.6) \qquad\qquad G(x) = \int_{\mathbf{R}} \widehat{G}(\theta) e^{i\theta x} \, d\theta.$$

In this article, $G$ will always be the nonlinearity $F$, its derivatives, or their mollifications.

1.5. *Organisation of the article.* The rest of the article is organised as follows. In Section 2, we briefly explain the difficulties and give an outline of the strategy. In Section 3, we construct the regularity structures for (1.5), and solve the corresponding abstract equation in a suitable modeled distribution space associated to the regularity structures. Section 4 contains a few preliminary lemmas and bounds. In Section 5, we prove the convergence of the rescaled processes to the KPZ equation. The main new ingredient of the proof is a general pointwise bound for correlation functions, which we develop in Section 6. This bound enables us to reduce the problem of a general nonlinearity $F$ to that of a polynomial, which has already been treated in [25].

**2. Strategy.** In order to prove Theorem 1.5, we use the theory of regularity structures developed in [22] and its adaption to the case of polynomial $F$ in [25]. Let

$$(2.1) \qquad\qquad Z_\varepsilon = P * \xi_\varepsilon, \qquad \Psi_\varepsilon = P' * \xi_\varepsilon,$$

where $P$ is the heat kernel on the torus $\mathbf{T}$, and $*$ denotes space–time convolution. If $h_\varepsilon$ satisfies (1.5), then the remainder $u_\varepsilon = h_\varepsilon - Z_\varepsilon$ solves the equation

$$(2.2) \qquad \begin{aligned} \partial_t u_\varepsilon &= \partial_x^2 u_\varepsilon + \varepsilon^{-1} F\big(\varepsilon^{\frac{1}{2}}(\Psi_\varepsilon + \partial_x u_\varepsilon)\big) - C_\varepsilon, \\ u_0^{(\varepsilon)} &= h_0^{(\varepsilon)} - Z_0^{(\varepsilon)}, \end{aligned}$$

where $Z_0^{(\varepsilon)} = \int_{-\infty}^0 P_{t-s} * \xi_\varepsilon(s)\,ds$. Since $\varepsilon^{\frac{1}{2}}\Psi_\varepsilon \sim \mathcal{N}(0, \sigma^2)$ and, by analogy with [21], $\varepsilon^{\frac{1}{2}}\partial_x u_\varepsilon$ is expected to have size of almost $\varepsilon^{\frac{1}{2}}$, we can therefore expand $F$ near $\varepsilon^{\frac{1}{2}}\Psi_\varepsilon$, formally yielding

$$\begin{aligned} \varepsilon^{-1} F\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon &+ \varepsilon^{\frac{1}{2}}\partial_x u_\varepsilon\big) \\ &= \varepsilon^{-1} F\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon\big) + \varepsilon^{-\frac{1}{2}} F'\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon\big) \cdot (\partial_x u_\varepsilon) \\ (2.3) \qquad &+ \frac{1}{2} F''\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon\big) \cdot (\partial_x u_\varepsilon)^2 + \mathcal{O}(\varepsilon^{\frac{1}{2}-}). \end{aligned}$$

The natural step next is to characterise as $\varepsilon \to 0$ the limiting objects $F''(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon)$, $\varepsilon^{-1} F(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon)$, $\varepsilon^{-\frac{1}{2}} F'(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon)$, as well as the products between them and $\partial_x u_\varepsilon$ and $(\partial_x u_\varepsilon)^2$. If one expands $F''$, $\varepsilon^{-\frac{1}{2}} F'$ and $\varepsilon^{-1} F$ (all with the argument $\varepsilon^{\frac{1}{2}}\Psi_\varepsilon$) into Wiener chaos, then it is easy to see that arbitrary moments of all the higher order chaos vanish termwise as $\varepsilon \to 0$, and one should expect that the limiting objects are given by constant multiples of 1, the free field $\Psi$ (limit of $\Psi_\varepsilon$) and its Wick square $\Psi^{\diamond 2}$ (after recentering), respectively. The behaviour of product of these objects as $\varepsilon \to 0$ can be seen in a similar way.

In fact, this is the procedure taken in [25] for polynomial $F$. However, the main obstacle of implementing the same procedure for general $F$ is that it gives an

*infinite chaos series*. In order for the termwise $L^p$ moments to be summable for every $p$, even for fixed $\varepsilon$, one needs extremely fast decay of the coefficients in the chaos expansion. This translates into the condition that the Fourier transform of $F$ should decay faster than every Gaussian, which is clearly too restrictive for the main statement to be interesting and widely applicable.

This way of direct chaos expansion is of course far from being optimal when $F$ is not a polynomial. In fact, different homogeneous chaos are highly correlated in $L^p$ for large $p$, and the sum of termwise $L^p$ norms is simply a bad upper bound for the $L^p$ norm of the whole series—it does not capture large cancellations between different terms from the chaos expansion. On the other hand, in order to get a bound that is uniform in $\varepsilon$, one still needs to chaos expand the object at some point so that the negative power of $\varepsilon$ in front of $F$ or $F'$ can be balanced out by the positive powers carried by Wick powers in the expansion.

It is at this point that our approach starts to deviate from that in [25]. The main idea is to write the nonlinear function of the free fields in terms of its Fourier transform, and to use clustering arguments and trigonometric identities to encode cancellations *before* chaos expanding them. In this way, we obtain a pointwise bound on correlations of trigonometric functions of $\Psi_\varepsilon$ in terms of their polynomial counterparts. This bound is uniform in $\varepsilon$ and polynomial in the frequency of the trigonometric function.[3] Hence, as long as $F$ is sufficiently regular, it essentially enables us to reduce the problem of a general $F$ to that of a polynomial.

A systematic procedure to obtain such a pointwise bound in general situations, which applies to all the objects appearing in the expansion of (2.3), is developed in Section 6. The same technique applies straight away to all objects arising in the study of the dynamical $\Phi_3^4$ model and would in principle allow to recover the results of [15] where techniques from Malliavin calculus are employed to bound these objects. Conversely, it appears that these techniques may also in principle be able to treat the KPZ case with Gaussian noise.

One advantage of our present approach is that it isolates the reliance on the Gaussianity of the noise into the bound of Section 6, while the rest of the argument is essentially independent of it.

**3. Regularity structure and the abstract equation.**   The aim of this section is to construct the regularity structure that will enable us to solve the equation (2.2). As long as we can solve for $u_\varepsilon$, the process $h_\varepsilon$ is just $u_\varepsilon + Z_\varepsilon$. From now on, we focus on the remainder equation (2.2).

---

[3]After Fourier expanding $F$, if we chaos expand the trigonometric functions directly without doing clustering, then one would get a very crude bound (inverse Gaussian rather than polynomial) in the frequency variables. As a consequence, we would still end up requiring $\widehat{F}$ decaying faster than Gaussian in order for the stochastic objects to converge. See Remark 6.21 for more details.

3.1. *The regularity structure.* We start by introducing the collection of symbols in the regularity structure. Let $X^k$ denote the abstract polynomials, where $k = (k_0, k_1)$ is a two-dimensional index with $k_i \in \mathbf{N}$. In particular, we use the special symbol $\mathbf{1}$ for the case $k = \mathbf{0}$. We also use the symbols $\square$, $\circ$ and $\bullet$ to describe recentered and rescaled versions of $F''(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon)$, $\varepsilon^{-\frac{1}{2}} F'(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon)$ and $\varepsilon^{-1} F(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon)$, respectively. At the level of the regularity structure, we impose $\bullet = \circ^2$ even though the canonical lift used later on in this article does *not* satisfy the corresponding identity. The reason, as we shall see later, is that if we choose properly the constant multiples and renormalisations in the model, one converges to the "Wick square" of the other as $\varepsilon \to 0$.

Let $\mathcal{I}$ and $\mathcal{I}'$ denote the abstract integration maps with respect to the heat kernel and its spatial derivative. We then generate and add new symbols (basis vectors) to the regularity structure by applying $\mathcal{I}$ and $\mathcal{I}'$ and pairwise multiplication, as dictated by the structure of the equation (2.3) where we ignore for the moment the $\mathcal{O}(\varepsilon^{\frac{1}{2}-})$ error term. For convenience of notation, we use graphical notation analogous to those used in [21, 24] to denote the newly generated symbols by setting

$$ \raisebox{-0.5ex}{\scalebox{1}{$\mathord{\text{\raisebox{1ex}{$\circ$}}}$}} = \mathcal{I}'(\circ), \qquad \blacktriangledown = \mathcal{I}'(\bullet)^2, \qquad \underset{\circ}{\overset{\bullet}{\circ}} = \circ \cdot \mathcal{I}'(\circ \cdot \mathcal{I}'(\bullet)), \qquad \text{etc.} $$

Note that solid lines denote $\mathcal{I}'$ and not $\mathcal{I}$ since $\mathcal{I}$ does not play much of a role in our analysis. We now associate to every symbol $\tau$ a homogeneity $|\tau| \in \mathbf{R}$. For the Taylor polynomial $X^k$, we let $|X^k| = |k| = 2k_0 + k_1$ if $k = (k_0, k_1)$. Let $\kappa > 0$ be small. We set

$$ |\square| = -\kappa, \qquad |\circ| = -\frac{1}{2} - \kappa, $$

and define recursively

$$ |\tau \bar{\tau}| = |\tau| + |\bar{\tau}|, \qquad |\mathcal{I}(\tau)| = |\tau| + 2, \qquad |\mathcal{I}'(\tau)| = |\tau| + 1. $$

Note that since we decreed that $\bullet = \circ^2$, we have in particular $|\bullet| = -1 - 2\kappa$. The following is a list of all the symbols of negative homogeneity appearing in the regularity structure obtained in this way:

(3.1)
$$ \bullet, \; \underset{\circ}{\overset{\bullet}{\phantom{.}}}, \; \circ, \; \overset{\bullet\,\bullet}{\square}, \; \blacktriangledown, \; \underset{\circ}{\overset{\bullet}{\circ}}, \; \underset{\square}{\overset{\bullet}{\phantom{.}}}, \; \underset{\circ}{\overset{\circ}{\phantom{.}}}, \; \overset{\bullet}{\vert}, \; \square $$

The corresponding structure group can be defined in the same way as in [5, 22, 25], so we omit the details here.

3.2. *The models.* Recall the definition of $a$ in (1.4), and that $\Psi_\varepsilon = P' * \xi_\varepsilon$. For every $\varepsilon > 0$, we define a (random) representation $\mathbf{\Pi}^\varepsilon$ of the regularity structure by

$$ (\mathbf{\Pi}^\varepsilon \square)(z) = \frac{1}{2a} F''(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon(z)) - 1, $$

$$(3.2) \qquad (\mathbf{\Pi}^{\varepsilon}\circ)(z) = \frac{1}{2a\sqrt{\varepsilon}}F'\big(\varepsilon^{\frac{1}{2}}\Psi_{\varepsilon}(z)\big),$$

$$(\mathbf{\Pi}^{\varepsilon}\bullet)(z) = \frac{1}{a\varepsilon}F\big(\varepsilon^{\frac{1}{2}}\Psi_{\varepsilon}(z)\big) - C_{\bullet}^{(\varepsilon)},$$

where the constant $C_{\bullet}^{(\varepsilon)}$ is chosen in such a way that $\mathbf{E}(\mathbf{\Pi}^{\varepsilon}\bullet)(z) = 0$. This is extended canonically to the whole regularity structure by postulating that $\mathcal{I}$ and $\mathcal{I}'$ correspond to convolution by $K$ and $\partial_{x}K$, respectively, where $K$ is a suitable truncation of the heat kernel $P$ and equals $P$ in a domain containing the origin, and by setting $(\mathbf{\Pi}^{\varepsilon}\tau\bar{\tau})(z) = (\mathbf{\Pi}^{\varepsilon}\tau)(z) \cdot (\mathbf{\Pi}^{\varepsilon}\bar{\tau})(z)$.

The reason why we normalise and subtract constants in the way specified in (3.2) is that, with this choice of normalisation, we will be able to show that the action of $\mathbf{\Pi}^{\varepsilon}$ on these three symbols converges to 0, $\Psi$ and $\Psi^{\diamond 2}$, respectively, with these limits being independent of $F$ and of the covariance of the noise. This also justifies the relation $\bullet = \circ^{2}$ imposed in our regularity structure.

The canonical (random) model $\mathcal{L}_{\varepsilon} = (\Pi^{\varepsilon}, \Gamma^{\varepsilon}) = \mathcal{Z}(\mathbf{\Pi}^{\varepsilon})$ is then defined as in [5];[4] see [5], Definition 6.7, and the discussion preceding [5], Definition 6.23. The fact that the canonical model is indeed a model (i.e., it satisfies the corresponding analytical bounds) was shown in [5], Proposition 6.11 (see also [22], Proposition 8.27, for an essentially equivalent statement).

Given the canonical model $\mathcal{L}_{\varepsilon}$, we define a renormalised model $\widehat{\mathcal{L}}_{\varepsilon} = (\widehat{\Pi}^{\varepsilon}, \widehat{\Gamma}^{\varepsilon})$ by stipulating that $\widehat{\mathcal{L}}_{\varepsilon}$ is the model obtained from $\mathcal{L}_{\varepsilon}$ by BPHZ renormalisation, as defined in [5], Theorem 6.17. In our case, this renormalisation procedure is easy to describe explicitly: we set $\widehat{\mathbf{\Pi}}^{\varepsilon}\tau = \mathbf{\Pi}^{\varepsilon}\tau$ for $\tau \in \{\square, \circ, \bullet\}$ and keep the canonical actions of $\mathcal{I}, \mathcal{I}'$ and products intact, except that we set

$$(\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet\bullet}{\vee})(z) = \big(\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet}{\mathsf{l}}\big)^{2}(z) - C_{\overset{\bullet\bullet}{\vee}}^{(\varepsilon)},$$

$$(3.3) \qquad (\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet}{\underset{\circ}{\circ}\circ})(z) = \big(\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet}{\circlearrowleft}\big)(z) \cdot \big(\widehat{\mathbf{\Pi}}^{\varepsilon}\circ\big)(z) - C_{\overset{\bullet}{\circ}\circ}^{(\varepsilon)},$$

$$(\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet\bullet}{\underset{\mathsf{H}}{\vee}})(z) = \big(\widehat{\mathbf{\Pi}}^{\varepsilon}\square\big)(z) \cdot \big(\widehat{\mathbf{\Pi}}^{\varepsilon}\overset{\bullet\bullet}{\vee}\big)(z) - C_{\overset{\bullet\bullet}{\mathsf{H}}}^{(\varepsilon)},$$

where the values of the constants $C_{\tau}^{(\varepsilon)}$ are chosen in such a way that $\mathbf{E}(\widehat{\mathbf{\Pi}}^{\varepsilon}\tau)(0) = 0$ for all symbols $\tau$ appearing in (3.3). We then set $\widehat{\mathcal{L}}_{\varepsilon} = \mathcal{Z}(\widehat{\mathbf{\Pi}}^{\varepsilon})$, which is again a model by [5], Theorem 6.28. Using the Gaussianity of $\Psi_{\varepsilon}$, the fact that $F$ is even by assumption, and our assumption that the covariance of $\xi_{\varepsilon}$ is spatially symmetric, one can verify that this then implies that $\mathbf{E}(\widehat{\mathbf{\Pi}}^{\varepsilon}\tau)(0) = 0$ for all $\tau$ with $|\tau| < 0$, which implies that this is indeed the BPHZ model by the uniqueness statement in [5], Theorem 6.17.

---

[4]For the purpose of applying the results of [5], we consider $\circ$ and $\bullet$ as unrelated "noise types". The relation $\bullet = \circ^{2}$ is really only useful for comparing the limiting model to existing results on the KPZ equation and plays no role here. In particular, it does not restrict the space of admissible models for our regularity structure.

It was shown in [6] that, for a rather large class of noises and stochastic PDEs, the corresponding BPHZ renormalised model is well defined and stable under perturbations. Unfortunately, in order to apply this result to our situation, we would need sharp cumulant bounds of all orders on the three stochastic processes appearing in (3.3). In our particular example, this does not appear to be any easier than showing the full convergence of the models, so we will avoid using these results.

3.3. *Abstract equation.* The aim of this subsection is to formulate and solve an abstract fixed-point problem in a suitable modeled distribution space so that it can be reconstructed back to (2.2). The form of the equation (2.2), the Taylor expansion of $F$ in (2.3) and the definition of the model in (3.2) and (3.3) suggest that we may want to consider the fixed-point equation

$$U = \mathcal{P}\mathbf{1}_+\big(a(\circ + \mathcal{D}U)^2 + a\square(\mathcal{D}U)^2$$

$$(3.4) \qquad\qquad + \varepsilon^{-1}G\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon, \varepsilon^{\frac{1}{2}}\mathcal{R}_\varepsilon\mathcal{D}U\big) \cdot \mathbf{1}\big) + \widehat{\mathcal{P}}u_0$$

(here we do use the interpretation $\bullet = \circ^2$ when expanding the square) where

$$G(x, y) = F(x + y) - F(x) - F'(x)y - \frac{1}{2}F''(x)y^2,$$

and $\mathcal{D}$ is the abstract differentiation operator in the regularity structure. Using (3.2) and the fact that we are considering the canonical model, one verifies that solving (3.4) as a fixed-point problem in some $\mathcal{D}^\gamma$ space based on the canonical model $\mathcal{L}_\varepsilon$ yields a solution $U$ which is such that $u_\varepsilon = \mathcal{R}_\varepsilon U$ solves (2.2) with $C_\varepsilon = aC_\bullet^{(\varepsilon)}$. (Here, $\mathcal{R}_\varepsilon$ is the reconstruction operator associated to $\mathcal{L}_\varepsilon$.) On the other hand, it turns out that the solution associated to the model $\widehat{\mathcal{L}}_\varepsilon$ defined using (3.3) (in which case we replace of course the reconstruction operator $\mathcal{R}_\varepsilon$ appearing in (3.4) by the reconstruction operator $\widehat{\mathcal{R}}_\varepsilon$ associated to $\widehat{\mathcal{L}}_\varepsilon$) yields solutions $u_\varepsilon$ to (2.2) with

$$(3.5) \qquad\qquad C_\varepsilon = aC_\bullet^{(\varepsilon)} + a^3\big(C_{\bullet\!\!\vee}^{(\varepsilon)} + 4C_{\infty}^{(\varepsilon)} + C_{\Psi\!\!\bullet}^{(\varepsilon)}\big).$$

This will be justified in Theorem 3.8 below.

In order to solve the abstract fixed-point problem (3.4), we show that the right-hand side yields a contraction in a suitable space for small enough time $T$. This will give the existence of solution for short time, and then we can continue it to maximal time. Since $\Psi_\varepsilon$ is the stationary free field, this requires us in particular to be able to treat initial data for the equation for $u$ that have regularity just below $\mathcal{C}^{\frac{1}{2}}$.

There is a technical issue in carrying out this procedure. Since we cannot expect the solution at very short times to behave better than the solution to the heat equation, we can at best hope for a bound of order $t^{-\frac{1-\eta}{2}}$ on $\mathcal{R}\mathcal{D}U$ as $t \to 0$ if we start with a generic initial condition in $\mathcal{C}^\eta$. In particular, for any positive $\varepsilon$, we would not expect the term $G(\varepsilon^{\frac{1}{2}}\psi, \varepsilon^{\frac{1}{2}}\mathcal{R}\mathcal{D}U)$ to be integrable if $F(x)$ grows faster than

$|x|^{\frac{2}{1-\eta}}$ as $|x| \to \infty$. This would require $F$ to have less-than-quartic growth in order to start from initial data below $\mathcal{C}^{\frac{1}{2}}$. Fortunately, what saves us is that for any fixed $\varepsilon$, the solution is actually smooth, as long as we consider scales smaller than $\varepsilon$.

In order to quantify this, we proceed as in [25] and introduce $\varepsilon$-dependent spaces of functions and models. For $\eta \in (0, 1)$ and $\gamma \in (1, 2)$, we let $\mathcal{C}^{\gamma,\eta}_\varepsilon$ be the space of functions that are $\mathcal{C}^\eta$ at large scales (larger than $\varepsilon$) and $\mathcal{C}^\gamma$ at smaller scales. More precisely, we define the norm $\| \cdot \|_{\gamma,\eta;\varepsilon}$ by

$$\|u\|_{\gamma,\eta;\varepsilon} := \|u\|_{\mathcal{C}^\eta} + \frac{\|u'\|_\infty}{\varepsilon^{\alpha-1}} + \|u'\|_\infty + \sup_{\substack{x \neq y \\ |x-y| \leq \varepsilon}} \frac{|u'(x) - u'(y)|}{\varepsilon^{\eta-\gamma}|x-y|^{\gamma-1}}.$$

Note that $\mathcal{C}^{\gamma,\eta}_0$ is the same as $\mathcal{C}^\eta$. We can also compare two functions $u^{(\varepsilon)} \in \mathcal{C}^{\gamma,\eta}_\varepsilon$ and $u \in \mathcal{C}^\eta$ by

$$
\begin{aligned}
\|u^{(\varepsilon)}; u\|_{\gamma,\eta;\varepsilon} &= \|u^{(\varepsilon)} - u\|_{\mathcal{C}^\eta} \\
(3.6) \qquad\qquad &+ \frac{\|(u^{(\varepsilon)})'\|_\infty}{\varepsilon^{\alpha-1}} + \sup_{\substack{x \neq y \\ |x-y| \leq \varepsilon}} \frac{|(u^{(\varepsilon)})'(x) - (u^{(\varepsilon)})'(y)|}{\varepsilon^{\eta-\gamma}|x-y|^{\gamma-1}}.
\end{aligned}
$$

For a continuous function $\varphi : \mathbf{R}^+ \times \mathbf{T} \to \mathbf{R}$, $z \in \mathbf{R}^+ \times \mathbf{T}$ and $\lambda > 0$, we let $\varphi^\lambda_z$ be the function

$$\varphi^\lambda_z(z') = \lambda^{-3}\varphi((z' - z)/\lambda).$$

We let $\mathcal{B}$ denote the set of smooth functions which are compactly supported in the ball of radius one, and whose derivatives up to second order (including the function itself) are uniformly bounded by 1. We let $\mathcal{B}_0$ denote the class of functions $\phi \in \mathcal{B}$ such that $\int \phi(z)\, dz = 0$. Let $\mathcal{M}_\varepsilon$ be the space of admissible models[5] $(\Pi, \Gamma)$ that further satisfy the bound

$$\left|(\Pi_z\tau)(\phi^\lambda_z)\right| \lesssim \lambda^{\bar{\gamma}}\varepsilon^{|\tau|-\bar{\gamma}}, \qquad \tau \in \mathcal{U}', \bar{\gamma} = 1 - \kappa$$

for all test functions $\phi \in \mathcal{B}_0$. Let

$$\|\Pi\|_\varepsilon := \sup_z \sup_{\substack{\tau \in \mathcal{U}' \\ |\tau| < \bar{\gamma}}} \sup_{\lambda \leq \varepsilon} \sup_{\phi \in \mathcal{B}_0} \lambda^{-\bar{\gamma}}\varepsilon^{\bar{\gamma}-|\tau|}|(\Pi_z\tau)(\phi^\lambda_z)|.$$

We then define the family of "norms" on $\mathcal{M}_\varepsilon$ by

$$\|\!|\Pi|\!\|_\varepsilon := \|\!|\Pi|\!\| + \|\Pi\|_\varepsilon,$$

---

[5]Admissible models are the ones that act canonically on abstract Taylor polynomials and for which the abstract integration maps do represent convolution by $K$; see [22]. All models considered in this article are indeed admissible.

where $\|\!|\!| \cdot |\!|\!\|$ is the usual norm on space of models as in [22]. Note that $\mathscr{M}_\varepsilon$ consists of the same collection of models for all $\varepsilon > 0$, but their "norms" behave very differently as $\varepsilon \to 0$. We compare a model $\Pi^\varepsilon \in \mathscr{M}_\varepsilon$ and a model $\Pi \in \mathscr{M}$ by

$$\|\!|\!| \Pi^\varepsilon; \Pi |\!|\!\|_{\varepsilon,0} = \|\!|\!| \Pi^\varepsilon; \Pi |\!|\!\| + \|\!|\!| \Pi^\varepsilon |\!|\!\|_\varepsilon.$$

We have only $\Pi^\varepsilon$ under the norm $\| \cdot \|_\varepsilon$ above since $\|\Pi\|_\varepsilon$ may be infinity for positive $\varepsilon$.

We also introduce $\varepsilon$-dependent spaces of modelled distributions as given in [25], Definition 2.17. Given a model $(\Pi, \Gamma) \in \mathscr{M}_\varepsilon$, define the $\mathcal{D}_\varepsilon^{\gamma,\eta}$ space to be the modelled distributions $U$ with the norm

$$\|U\|_{\gamma,\eta;\varepsilon} := \|U\|_{\gamma,\eta} + \sup_z \sup_{\alpha > \gamma} \frac{|U(z)|_\alpha}{\varepsilon^{\eta-\alpha}}$$

$$+ \sup_{\substack{|z-z'| \leq \sqrt{|t| \wedge |t'|} \\ |z-z'| \leq \varepsilon}} \frac{|U(z) - \Gamma_{z,z'} U(z')|}{|z - z'|^{\gamma-\alpha} \varepsilon^{\eta-\gamma}}.$$

Similarly, we compare two functions $U^{(\varepsilon)} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ and $U \in \mathcal{D}^{\gamma,\eta}$ by

$$\|U^{(\varepsilon)}; U\|_{\gamma,\eta;\varepsilon} = \|U^{(\varepsilon)}; U\|_{\gamma,\eta} + \sup_z \sup_{\alpha > \gamma} \frac{|U^{(\varepsilon)}(z)|_\alpha}{\varepsilon^{\eta-\alpha}}$$

$$+ \sup_{\substack{|z-z'| \leq \sqrt{|t| \wedge |t'|} \\ |z-z'| \leq \varepsilon}} \frac{|U^{(\varepsilon)}(z) - \Gamma_{z,z'} U^{(\varepsilon)}(z')|}{|z - z'|^{\gamma-\alpha} \varepsilon^{\eta-\gamma}}.$$

The reason that $U$ does not appear in the latter two terms on the right-hand side above is the same as before—these two supremum may be infinity for general $U \in \mathcal{D}_0^{\gamma,\eta}$. Also, here $\eta$ is allowed to be any real number less than $\gamma$ (not necessarily positive).

REMARK 3.1.   The readers may have noticed that we have the abuse of notation $\| \cdot \|_{\gamma,\eta;\varepsilon}$ to denote both $\mathcal{C}_\varepsilon^{\gamma,\eta}$ and $\mathcal{D}_\varepsilon^{\gamma,\eta}$ norms. But since the precise function space we are referring to should be clear in relevant contexts, we keep this same notation for both for simplicity.

PROPOSITION 3.2.   *Let $\gamma \in (1, 2)$ and $\eta \in (0, 1)$. Let $u \in \mathcal{C}_\varepsilon^{\gamma,\eta}$, and $\widehat{P}u$ be the harmonic extension of $u$. Then $\widehat{P}u \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ and $\|\widehat{P}u\|_{\gamma,\eta;\varepsilon} \lesssim \|u\|_{\gamma,\eta;\varepsilon}$. Furthermore, if $\bar{u} \in \mathcal{C}^\eta$, then one has $\|\widehat{P}u; \widehat{P}\bar{u}\|_{\gamma,\eta;\varepsilon} \lesssim \|u; \bar{u}\|_{\gamma,\eta;\varepsilon}$.*

PROOF.   Same as [25], Proposition 4.7.  □

PROPOSITION 3.3. *Let $\eta \leq 1 - \kappa$, $\gamma = \frac{3}{2} + 2M\kappa$, and $T > 0$. Let $U_\varepsilon \in \mathcal{D}^{\gamma,\eta}$ be based on some model $\Pi^\varepsilon \in \mathcal{M}_\varepsilon$. Then, for every $t > 0$ such that $[t - \varepsilon^2, t + \varepsilon^2] \subset [0, T]$, the function $u_t^{(\varepsilon)} = (\mathcal{R}^\varepsilon U_\varepsilon)(t, \cdot)$ belongs to $\mathcal{C}_\varepsilon^{\gamma,\eta}$ with the bound*

$$\|u_t^{(\varepsilon)}\|_{\gamma,\eta;\varepsilon} \lesssim \|U_\varepsilon\|_{\gamma,\eta} \||\Pi^\varepsilon\||_\varepsilon,$$

*where the proportionality constant is independent of $\varepsilon$. Furthermore, if $U \in \mathcal{D}^{\gamma,\eta}$ based on some model $\Pi \in \mathcal{M}$, then $u_t = (\mathcal{R}U)(t, \cdot) \in \mathcal{C}^\eta$ and one has the bound*

$$\|u_t^{(\varepsilon)}; u_t\|_{\gamma,\eta;\varepsilon} \lesssim \|U_\varepsilon; U\|_{\gamma,\eta}\big(\||\Pi^\varepsilon\||_\varepsilon + \||\Pi\||\big)$$
$$+ \||\Pi^\varepsilon; \Pi\||_{\varepsilon,0}\big(\|U_\varepsilon\|_{\gamma,\eta} + \|U\|_{\gamma,\eta}\big).$$

PROOF. Same as [25], Proposition 4.8. □

PROPOSITION 3.4. *Let $U \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ for some $\gamma > 1$ and $\eta \in \mathbf{R}$. Then $\mathscr{D}U \in \mathcal{D}_\varepsilon^{\gamma-1,\eta-1}$ with the bound*

$$\|\mathscr{D}U\|_{\gamma-1,\eta-1;\varepsilon} \leq C\|U\|_{\gamma,\eta;\varepsilon}.$$

*Furthermore, if $\bar{U} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$, then one has*

$$\|\mathscr{D}U; \mathscr{D}\bar{U}\|_{\gamma-1,\eta-1;\varepsilon} \leq \|U; \bar{U}\|_{\gamma,\eta;\varepsilon}.$$

PROOF. Same as [25], Proposition 4.9. □

PROPOSITION 3.5. *For $i = 1, 2$, let $U_i \in \mathcal{D}_\varepsilon^{\gamma_i,\eta_i}(V^{(i)})$, where $V^{(i)}$ is a sector with regularity $\alpha_i$. Then, $U = U_1 U_2 \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ with*

$$\gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1), \qquad \eta = (\eta_1 + \alpha_2) \wedge (\eta_2 + \alpha_1) \wedge (\eta_1 + \eta_2),$$

*and one has the bound*

$$\|U\|_{\gamma,\eta;\varepsilon} < C\|U_1\|_{\gamma_1,\eta_1;\varepsilon}\|U_2\|_{\gamma_2,\eta_2;\varepsilon}\big(1 + \||\Pi\||_\varepsilon\big)^2.$$

*Furthermore, if $\bar{U}_i \in \mathcal{D}_\varepsilon^{\gamma_i,\eta_i}$ with the same parameters as above, and $\bar{U} = \bar{U}_1 \bar{U}_2$, then one has the bound*

$$\|U; \bar{U}\|_{\gamma,\eta;\varepsilon} < C\big(\|U_1; \bar{U}_1\|_{\gamma_1,\eta_1;\varepsilon} + \|U_2; \bar{U}_2\|_{\gamma_2,\eta_2;\varepsilon}\big)\big(1 + \||\Pi; \bar{\Pi}\||_\varepsilon\big).$$

PROOF. Same as [25], Proposition 4.10. □

PROPOSITION 3.6. *Let $V$ be a sector of regularity $\alpha$, and let $U \in \mathcal{D}_\varepsilon^{\gamma,\eta}(V)$ with $-2 < \eta < \gamma \wedge \alpha$. Let $\mathbf{1}_+$ be the restriction of time variables to be positive, and $T$ denote the length of the interval where the abstract integration takes place. Then, provided that $\gamma$ and $\eta$ are not integers, there exists $\theta > 0$ such that*

$$\|\mathcal{P}\mathbf{1}_+ U\|_{\gamma+2,\eta+2;\varepsilon} \leq C(T + \varepsilon)^\theta\big(\|U\|_{\gamma,\eta;\varepsilon} + \||\Pi\||_\varepsilon\big).$$

*Furthermore, if $\bar{U} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$, then one has*

$$\|\mathcal{P}U; \mathcal{P}\bar{U}\|_{\gamma+2,\eta+2;\varepsilon} \leq C(T + \varepsilon)^\theta\big(\|U; \bar{U}\|_{\gamma,\eta;\varepsilon} + \||\Pi; \bar{\Pi}\||_\varepsilon\big).$$

PROOF. Same as [22], Theorem 7.1, Lemma 7.3, and [25], Proposition 4.13.
□

We are now ready to prove the main theorem on the existence of the solution to the abstract equation. We restrict our regularity structure to the space spanned by polynomials and the basis elements listed in (3.1).

THEOREM 3.7. *Let $\gamma \in (\frac{3}{2}, \frac{5}{3})$ and $\eta \in (\frac{1}{2} - \frac{1}{M+4}, \frac{1}{2})$, where $M$ is the same as in Assumption* 1.1. *Let $u_0^{(\varepsilon)} \in \mathcal{C}_\varepsilon^{\gamma,\eta}$, and let $\psi_\varepsilon : \mathbf{R}^+ \times \mathbf{T} \to \mathbf{R}$ be a family of smooth functions such that*

$$\sup_{\varepsilon \in (0,1)} \sup_{z \in [0,\bar{T}] \times \mathbf{T}} \varepsilon^{\frac{1}{2}+\kappa} |\psi_\varepsilon(z)| < +\infty$$

*for every $\bar{T} > 0$, and for some sufficiently small $\kappa > 0$. Consider the fixed-point problem*

$$(3.7) \quad U = \mathcal{P}\mathbf{1}_+ \big( a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2 + \varepsilon^{-1} G(\varepsilon^{\frac{1}{2}}\psi_\varepsilon, \varepsilon^{\frac{1}{2}}\mathcal{R}\mathscr{D}U) \cdot \mathbf{1} \big) + \widehat{P}u_0^{(\varepsilon)},$$

*where $\widehat{P}$ is the harmonic extension operator, and*

$$G(x, y) = F(x + y) - F(x) - F'(x)y - \frac{1}{2}F''(x)y^2.$$

*Then there exists $\varepsilon_0 > 0$ such that for every $\varepsilon \in [0, \varepsilon_0]$ and every $\Pi^\varepsilon \in \mathscr{M}_\varepsilon$, (3.7) has a unique solution $U^{(\varepsilon)} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ to up to time $T$. Furthermore, the terminal time $T$ can be chosen uniformly jointly over bounded sets of initial conditions in $\mathcal{C}_\varepsilon^{\gamma,\eta}$ and bounded sets in $\mathscr{M}_\varepsilon$.*

*Now, suppose $\{u_0^{(\varepsilon)}\}$ is a sequence in $\mathcal{C}_\varepsilon^{\gamma,\eta}$ such that $\|u_0^{(\varepsilon)}; u_0\|_{\gamma,\eta;\varepsilon} \to 0$ for some function $u_0 \in \mathcal{C}^\eta$, and $\{\Pi^\varepsilon\}$ is a sequence of models in $\mathscr{M}_\varepsilon$ such that $\|\Pi^\varepsilon; \Pi\|_{\varepsilon,0} \to 0$ for some model $\Pi \in \mathscr{M}$. Let $T > 0$ be fixed and suppose $U \in \mathcal{D}^{\gamma,\eta}$ solves the fixed-point problem*

$$(3.8) \quad U = \mathcal{P}\mathbf{1}_+ \big( a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2 \big) + \widehat{P}u_0,$$

*with the model $\Pi$ and initial data $u_0$ up to time $T$. Then, for every $\varepsilon > 0$ sufficiently small, there exists a unique solution $U^{(\varepsilon)} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ to (3.7) with the model $\Pi^\varepsilon$ and initial condition $u_0^{(\varepsilon)}$ up to the same terminal time $T$. Furthermore, we have*

$$\lim_{\varepsilon \to 0} \|U^{(\varepsilon)}; U\|_{\gamma,\eta;\varepsilon} = 0.$$

PROOF. We first show that the fixed-point problem (3.7) can be solved in $\mathcal{D}_\varepsilon^{\gamma,\eta}$ with local existence time uniform in $\varepsilon$. Consider the map

$$\mathcal{M}_T^{(\varepsilon)}(U) = \mathcal{P}\mathbf{1}_+ \big( a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2 + \varepsilon^{-1} G(\varepsilon^{\frac{1}{2}}\psi_\varepsilon, \varepsilon^{\frac{1}{2}}\mathcal{R}^\varepsilon \mathscr{D}U) \cdot \mathbf{1} \big)$$

$$(3.9) \qquad + \widehat{P}u_0^{(\varepsilon)},$$

where $\mathcal{R}^\varepsilon$ is the reconstruction operator associated to $\Pi^\varepsilon \in \mathcal{M}_\varepsilon$, and $T$ is the length of the time interval on which the modelled distribution $U$ is defined. We want to show that for sufficiently small $T$ and $\varepsilon$, $\mathcal{M}_T^{(\varepsilon)}$ is a contraction map from $\mathcal{D}_\varepsilon^{\gamma,\eta}$ to itself, and this $T$ can be chosen uniformly over all small $\varepsilon$. For convenience of notation, we omit the subscript $T$ below and simply denote the map by $\mathcal{M}^{(\varepsilon)}$.

For the term with the initial condition, by Proposition 3.2, we have

$$\|\widehat{P}u_0^{(\varepsilon)}\|_{\gamma,\eta;\varepsilon} < C\|u_0^{(\varepsilon)}\|_{\gamma,\eta;\varepsilon}. \tag{3.10}$$

Note that the left-hand side above is the norm for $\mathcal{D}_\varepsilon^{\gamma,\eta}$, while the right-hand side is for $\mathcal{C}_\varepsilon^{\gamma,\eta}$. We now treat the terms $(\circ + \mathscr{D}U)^2$ and $\square \cdot (\mathscr{D}U)^2$.

Since $\mathscr{D}U \in \mathcal{D}_\varepsilon^{\gamma-1,\eta-1}$ with a sector of regularity $-2\kappa$, and the two "constant terms" $\circ$ and $\square$ are both in $\mathcal{D}_\varepsilon^{\infty,\infty}$ with sector regularities $-\frac{1}{2} - \kappa$ and $-\kappa$ respectively, it follows from Proposition 3.5 that

$$\|(\circ + \mathscr{D}U)^2\|_{\gamma_1,\eta_1;\varepsilon} \lesssim (1 + \|\!\|\Pi^\varepsilon\|\!\|_\varepsilon)^2(1 + \|U\|_{\gamma,\eta;\varepsilon})^2,$$

$$\gamma_1 = \gamma - \frac{3}{2} - \kappa, \eta_1 = 2\eta - 2,$$

$$\|\square \cdot (\mathscr{D}U)^2\|_{\gamma_2,\eta_2;\varepsilon} \lesssim (1 + \|\!\|\Pi^\varepsilon\|\!\|_\varepsilon)^3(1 + \|U\|_{\gamma,\eta;\varepsilon})^2,$$

$$\gamma_2 = \gamma - 1 - 3\kappa, \eta_2 = 2\eta - 2 - \kappa.$$

Thus, the modelled distribution $a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2$ belongs to $\mathcal{D}_\varepsilon^{\bar\gamma,\bar\eta}$ with a sector of regularity $\bar\alpha$ with

$$\bar\gamma = \gamma - \frac{3}{2} - \kappa, \qquad \bar\eta = 2\eta - 2 - \kappa, \qquad \bar\alpha = -1 - 2\kappa.$$

One can check that these parameters satisfy

$$-2 < \bar\eta < \bar\gamma \wedge \bar\alpha, \qquad \bar\gamma + 2 > \gamma, \qquad \bar\eta + 2 > \eta,$$

so one can apply Proposition 3.6 and use the natural inclusion of the $\mathcal{D}_\varepsilon^{\gamma,\eta}$ spaces to conclude that there exists $\theta_1 > 0$ such that

$$\|\mathcal{P}\mathbf{1}_+\big(a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2\big)\|_{\gamma,\eta;\varepsilon}$$

$$\lesssim (T + \varepsilon)^{\theta_1}(1 + \|\!\|\Pi^\varepsilon\|\!\|_\varepsilon)^3(1 + \|U\|_{\gamma,\eta;\varepsilon})^2. \tag{3.11}$$

As for the "Taylor remainder" term $G$, we have

$$G(x, y) = \frac{1}{6}F^{(3)}(x + h)y^3$$

for some $h$ between $0$ and $y$. Thus, the growth condition $|F^{(3)}(x + h)| \lesssim (1 + |x + h|)^M$ and the fact $|h| \leq |y|$ imply that

$$\|\varepsilon^{-1}G\big(\varepsilon^{\frac{1}{2}}\psi_\varepsilon(t, \cdot), \varepsilon^{\frac{1}{2}}\mathcal{R}^\varepsilon \mathscr{D}U(t, \cdot)\big)\|_{L^\infty(\mathbf{T})}$$

$$\lesssim \varepsilon^{\frac{1}{2}}\|(\mathcal{R}^\varepsilon \mathscr{D}U)(t, \cdot)\|_{L^\infty(\mathbf{T})}^3\big(1 + \|\varepsilon^{\frac{1}{2}}\psi_\varepsilon\|_\infty + \|\varepsilon^{\frac{1}{2}}(\mathcal{R}^\varepsilon \mathscr{D}U)(t, \cdot)\|_{L^\infty(\mathbf{T})}\big)^M.$$

Since $\mathscr{D}U \in \mathcal{D}_\varepsilon^{\gamma-1,\eta-1}$, it follows from the assumption $\Pi^\varepsilon \in \mathcal{M}_\varepsilon$ and the reconstruction theorem ([22], Theorem 3.10, Proposition 6.9) that

$$\left\|(\mathcal{R}^\varepsilon \mathscr{D}U)(t,\cdot)\right\|_{L^\infty(\mathbf{T})} \lesssim \left\|\!\left\|\Pi^\varepsilon\right\|\!\right\|_\varepsilon \|U\|_{\gamma,\eta;\varepsilon} \cdot (\sqrt{t}+\varepsilon)^{\eta-1}.$$

Together with the assumption on $\psi_\varepsilon$, we see that for every $\delta > 0$, we have

$$\left\|\varepsilon^{-1}G\big(\varepsilon^{\frac{1}{2}}\psi_\varepsilon(t,\cdot),\varepsilon^{\frac{1}{2}}(\mathcal{R}^\varepsilon\mathscr{D}U)(t,\cdot)\big)\right\|_{L^\infty(\mathbf{T})}$$

$$\lesssim \varepsilon^{\frac{1}{2}}(\sqrt{t}+\varepsilon)^{3(\eta-1)}\big(\varepsilon^{-M\kappa}+(\sqrt{t}+\varepsilon)^{M(\eta-\frac{1}{2})}\big)$$

$$\times \big(1+\left\|\!\left\|\Pi^\varepsilon\right\|\!\right\|_\varepsilon\|U\|_{\gamma,\eta;\varepsilon}\big)^{M+3}.$$

One can easily check that for $\eta \in (\frac{1}{2}-\frac{1}{M+4},\frac{1}{2})$ and sufficiently small $\kappa$, we have

$$\varepsilon^{\frac{1}{2}}(\sqrt{t}+\varepsilon)^{3(\eta-1)}\big(\varepsilon^{-M\kappa}+(\sqrt{t}+\varepsilon)^{M(\eta-\frac{1}{2})}\big) \lesssim \varepsilon^\delta t^{-1+\delta}$$

for some $\delta > 0$. In particular, the singularity in $t$ near the origin is integrable. Hence, there exist $C, \theta_2 > 0$ such that

$$\varepsilon^{-1}\big\|\mathcal{P}\mathbf{1}_+ G\big(\varepsilon^{\frac{1}{2}}\psi_\varepsilon,\varepsilon^{\frac{1}{2}}\mathcal{R}^\varepsilon\mathscr{D}U\big)\big\|_{\gamma,\eta;\varepsilon}$$

$$(3.12) \qquad \leq C(\varepsilon T)^{\theta_2}\big(1+\left\|\!\left\|\Pi^\varepsilon\right\|\!\right\|_\varepsilon\|U\|_{\gamma,\eta;\varepsilon}\big)^{M+3}.$$

Combining (3.10), (3.11) and (3.12), we see that for every $R > 0$ sufficiently large, there exists a final time $T > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in [0,\varepsilon_0]$, the map $\mathcal{M}_T^{(\varepsilon)}$ defined in (3.7) maps a ball of radius $R$ in $\mathcal{D}_\varepsilon^{\gamma,\eta}$ into itself provided $\|u_0^{(\varepsilon)}\|_{\gamma,\eta;\varepsilon} < \frac{\Lambda}{3C}$ for the $C$ in (3.10) and

$$(T+\varepsilon)^\theta < \frac{R}{3C(1+\left\|\!\left\|\Pi^\varepsilon\right\|\!\right\|_\varepsilon)^{M+3}(1+R)^{M+3}}.$$

We now fix $\Pi^\varepsilon \in \mathcal{M}_\varepsilon$, and we will show that the map $\mathcal{M}_T^{(\varepsilon)}$ is locally Lipschitz from $R$-balls of $\mathcal{D}_\varepsilon^{\gamma,\eta}$ into itself for with Lipschitz constant less than 1 for small $T$. This will imply that it is a contraction on the $R$-ball in $\mathcal{D}_\varepsilon^{\gamma,\eta}$. In fact, using the relation

$$(\mathscr{D}U)^2 - (\mathscr{D}\bar{U})^2 = (\mathscr{D}U - \mathscr{D}\bar{U})(\mathscr{D}U + \mathscr{D}\bar{U}),$$

we can use the same argument as above to show that the map

$$(3.13) \qquad U \mapsto \mathcal{P}\mathbf{1}_+\big(a(\circ + \mathscr{D}U)^2 + \square \cdot (\mathscr{D}U)^2\big)$$

is locally Lipschitz with the Lipschitz constant bounded by

$$C(T+\varepsilon)^\theta(1+R)^2\big(1+\left\|\!\left\|\Pi^\varepsilon\right\|\!\right\|_\varepsilon\big)^2.$$

As for the term $\varepsilon^{-1}G(\varepsilon^{\frac{1}{2}}\psi_\varepsilon,\varepsilon^{\frac{1}{2}}\mathcal{R}^\varepsilon\mathscr{D}U)$, by Taylor's theorem, there exist space–time functions $\phi_\varepsilon,\bar{\phi}_\varepsilon$ with

$$\big|\phi_\varepsilon - \varepsilon^{\frac{1}{2}}\psi_\varepsilon\big| \leq \varepsilon^{\frac{1}{2}}\big|\mathcal{R}^\varepsilon\mathscr{D}U\big|, \qquad \big|\bar{\phi}_\varepsilon - \varepsilon^{\frac{1}{2}}\psi_\varepsilon\big| \leq \varepsilon^{\frac{1}{2}}\big|\mathcal{R}^\varepsilon\mathscr{D}\bar{U}\big|$$

M. HAIRER AND W. XU

such that

$$\varepsilon^{-1} G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \varepsilon^{\frac{1}{2}} \mathcal{R}^\varepsilon \mathscr{D} U\big) = \frac{\sqrt{\varepsilon}}{6} F^{(3)}(\phi_\varepsilon)(\mathcal{R}^\varepsilon \mathscr{D} U)^3,$$

$$\varepsilon^{-1} G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \varepsilon^{\frac{1}{2}} \mathcal{R}^\varepsilon \mathscr{D} \bar{U}\big) = \frac{\sqrt{\varepsilon}}{6} F^{(3)}(\bar{\phi}_\varepsilon)(\mathcal{R}^\varepsilon \mathscr{D} \bar{U})^3.$$

Using the triangle inequality, we see that

$$\varepsilon^{-1} \big| G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \mathcal{R}^\varepsilon \mathscr{D} U\big) - G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \mathcal{R}^\varepsilon \mathscr{D} \bar{U}\big) \big|$$

$$\lesssim \varepsilon^{\frac{1}{2}} \big| F^{(3)}(\phi_\varepsilon) \big| \cdot \big| \mathcal{R}^\varepsilon \mathscr{D} U - \mathcal{R}^\varepsilon \mathscr{D} \bar{U} \big| \cdot \big( \big| \mathcal{R}^\varepsilon \mathscr{D} U \big|^2 + \big| \mathcal{R}^\varepsilon \mathscr{D} \bar{U} \big|^2 \big)$$

$$+ \varepsilon^{\frac{1}{2}} \big| F^{(3)}(\phi_\varepsilon) - F^{(3)}(\bar{\phi}_\varepsilon) \big| \cdot \big| \mathcal{R}^\varepsilon \mathscr{D} \bar{U} \big|^3.$$

For the second term on the right-hand side above, we have

$$\big| F^{(3)}(\phi_\varepsilon) - F^{(3)}(\bar{\phi}_\varepsilon) \big| \le \big| F^{(4)}(\widetilde{\phi}_\varepsilon) \big| \cdot \big| \phi_\varepsilon - \bar{\phi}_\varepsilon \big|$$

for some $\widetilde{\phi}_\varepsilon$ with $\big| \widetilde{\phi}_\varepsilon - \varepsilon^{\frac{1}{2}} \psi_\varepsilon \big| \le \big| \varepsilon^{\frac{1}{2}} \mathcal{R}^\varepsilon \mathscr{D} U \big| + \big| \varepsilon^{\frac{1}{2}} \mathcal{R}^\varepsilon \mathscr{D} \bar{U} \big|$. Thus, using the assumptions on the uniform boundedness of $\varepsilon^{\frac{1}{2}+\kappa} \psi_\varepsilon$, the growth of the derivatives of $F$, and that $|\phi_\varepsilon - \bar{\phi}_\varepsilon| \le \varepsilon^{\frac{1}{2}} |\mathcal{R}^\varepsilon (\mathscr{D} U - \mathscr{D} \bar{U})|$, we can perform a similar argument as before to obtain

$$\varepsilon^{-1} \big\| G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \mathcal{R}^\varepsilon \mathscr{D} U\big) - G\big(\varepsilon^{\frac{1}{2}} \psi_\varepsilon, \mathcal{R}^\varepsilon \mathscr{D} \bar{U}\big) \big\|_{L^\infty(\mathbf{T})}$$

$$\lesssim \varepsilon^\delta t^{-1+\delta} \big(1 + \big\|\!\|\Pi^\varepsilon\|\!\big\|_\varepsilon \big)^{M+4} (1+R)^{M+4} \|U - \bar{U}\|_{\gamma,\eta;\varepsilon}$$

for some $\delta > 0$. This implies that the map

$$U \mapsto \mathcal{P} \mathbf{1}_+ \big(\varepsilon^{-1} G\big(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon, \mathcal{R}^\varepsilon \mathscr{D} U\big) \cdot \mathbf{1}\big)$$

is locally Lipschitz with constant bounded by

$$C(\varepsilon T)^\theta \big(1 + \big\|\!\|\Pi^\varepsilon\|\!\big\|_\varepsilon \big)^{M+4} (1+R)^{M+4}$$

for some $C, \theta > 0$. Combining it with the local Lipschitzness of the map in (3.13) (and the Lipschitz constant there is bounded by $R^4 (T + \varepsilon)^\theta$), we deduce that for sufficiently small $T$, there is a unique solution in $\mathcal{D}_\varepsilon^{\gamma,\eta}$ to (3.7) up to time $T$. Furthermore, by the expressions for the local Lipschitz constants above, we see that there exists $\varepsilon_0 > 0$ such that this local existence time can be chosen uniformly over $\varepsilon \in (0, \varepsilon_0)$, over bounded sets of $\Pi^\varepsilon \in \mathscr{M}_\varepsilon$ and bounded sets of initial data $u_0^{(\varepsilon)} \in \mathcal{C}_\varepsilon^{\gamma,\eta}$.

We now turn to the second part of the theorem, namely the convergence of the solutions $U^{(\varepsilon)}$ up to the time $T$ for which the limiting abstract solution $U$ is defined. Here, $U^{(\varepsilon)} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ and $U \in \mathcal{D}_0^{\gamma,\eta}$ are local solutions to the fixed-point problems (3.7) and (3.8) based on $\Pi^\varepsilon \in \mathscr{M}_\varepsilon$ and $\Pi \in \mathscr{M}$, respectively. By the same arguments as above, we know there exists $S < T$ such that (3.7) has a fixed-point

solution up to time $S$ for all small $\varepsilon$. Let $\mathcal{M}_S^\varepsilon$ and $\mathcal{M}_S$ denote the maps associated to the fixed-point problem (3.7) for $\varepsilon > 0$ and $\varepsilon = 0$, respectively. Previous arguments have already shown that

$$\left\| \mathcal{M}_S^{(\varepsilon)}(U^{(\varepsilon)}); \mathcal{M}_S^{(\varepsilon)}(U) \right\|_{\gamma,\eta;\varepsilon}$$
$$\lesssim (S+\varepsilon)^\theta \| U^{(\varepsilon)}; U \|_{\gamma,\eta;\varepsilon} + \| \Pi^\varepsilon; \Pi \|_{\varepsilon;0} + \| u_0^{(\varepsilon)}; u_0 \|_{\gamma,\eta;\varepsilon},$$

where the proportionality constant is independent of $\varepsilon$. If $U^{(\varepsilon)}$ and $U$ are fixed-point solutions, then for small enough $S$, we have

$$(3.14) \qquad \| U^{(\varepsilon)}; U \|_{\gamma,\eta;\varepsilon} \lesssim \| \Pi^\varepsilon; \Pi \|_{\varepsilon;0} + \| u_0^{(\varepsilon)}; u_0 \|_{\gamma,\eta;\varepsilon},$$

which surely converges to 0 (up to time $S$) as $\varepsilon \to 0$. Furthermore, Proposition 3.3 guarantees that we can iterate (3.14) up to time $T$ in finitely many steps, thus completing the proof. $\square$

3.4. *Renormalised equation.* Let $\widehat{\mathcal{L}}_\varepsilon$ be the model defined in (3.3) and let $\mathcal{R}$ be the reconstruction map associated with $\widehat{\mathcal{L}}_\varepsilon$. We have the following theorem.

THEOREM 3.8. *Let $U \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ be the abstract solution to the fixed-point problem (3.7) with initial data $u_0 \in \mathcal{C}_\varepsilon^{\gamma,\eta}$. Then, $u = \mathcal{R}U$ solves the classical PDE*

$$\partial_t u = \partial_x^2 u + \varepsilon^{-1} F\left(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon + \varepsilon^{\frac{1}{2}} \partial_x u\right) - C_\varepsilon$$

*with initial condition $u_0$, where $C_\varepsilon$ is given by (3.5).*

PROOF. Applying the reconstruction operator to both sides of (3.7), we get

$$u = P * \mathbf{1}_+ \left( \mathcal{R}(a(\circ + \mathscr{D}U)^2 + a\square(\mathscr{D}U)^2) + \varepsilon^{-1} G(\varepsilon^{\frac{1}{2}} \Psi_\varepsilon, \varepsilon^{\frac{1}{2}} \mathcal{R}\mathscr{D}U) \right)$$
$$(3.15) \qquad + P_t u_0.$$

A straightforward calculation shows that, modulo terms of homogeneity strictly greater than $\frac{1}{2}$, any fixed point $U$ to (3.7) has the form

$$\mathscr{D}U = a\,\mathop{\bullet}\limits_{\vert} + v'\,\mathbf{1} + 2a^2\,\mathop{\propto}\limits^{\bullet} + 2av'\,\mathop{\circ}\limits_{},$$

for some space–time function $v'$. As a consequence, the factor multiplying $\mathbf{1}_+$ in the right-hand side of (3.7) is of the form

$$a\,\bullet + 2a^2\,\mathop{\bullet}\limits_{\circ} + 2av'\circ + 4a^3\,\mathop{\bullet}\limits_{\circ\circ} + a^3\,\mathop{\blacktriangledown}\limits + 2a^2v'\,\mathop{\bullet}\limits_{\vert} + a^3\,\mathop{\bullet}\limits_{\sqcup} + 2a^2v'\,\mathop{\bullet}\limits_{\square} + w\,\mathbf{1},$$

for some continuous function $w$, modulo terms of strictly positive homogeneity. The claim then follows from the definition of the model in (3.2) and (3.3) in essentially the same way as [22], Proposition 9.10, for example. $\square$

**4. Preliminary bounds.** In this section, we give some preliminary lemmas and bounds that will be useful later when we establish a general pointwise bound and prove the main convergence.

4.1. *Bounds on the free field.* Recall from (1.3) that

$$\xi_\varepsilon \stackrel{\text{law}}{=} \xi * \rho_\varepsilon$$

for some mollifier $\rho$ symmetric in its space variable and rescaled at scale $\varepsilon$. Here, "$*$" denotes the space–time convolution. Also recall that $\Psi_\varepsilon = P' * \xi_\varepsilon$, where $P'$ is the spatial derivative of the heat kernel $P$ on the torus. We have the following bound on the correlation function of $\Psi_\varepsilon$. It is the basis for most of the bounds throughout the article.

LEMMA 4.1. *The correlation $\mathbf{E}(\Psi_\varepsilon(x)\Psi_\varepsilon(y))$ is a function of $x - y$, and this function is symmetric in its space variable. Furthermore, there exists $\Lambda > 0$ such that*

$$(4.1) \qquad \frac{1}{\Lambda(|x - y| + \varepsilon)} \le \mathbf{E}\big(\Psi_\varepsilon(x)\Psi_\varepsilon(y)\big) \le \frac{\Lambda}{|x - y| + \varepsilon}$$

*for all $x, y \in \mathbf{R}^+ \times \mathbf{T}$ and $\varepsilon > 0$. Here, $|x - y|$ denotes the parabolic distance between the space–time points $x$ and $y$.*

PROOF. Since $\Psi_\varepsilon = P' * \xi_\varepsilon = (P' * \rho_\varepsilon) * \xi$, the delta correlation of the space–time white noise gives

$$\mathbf{E}\big(\Psi_\varepsilon(x)\Psi_\varepsilon(y)\big) = \big((P')^{\star 2} * (\rho_\varepsilon^{\star 2})\big)(x - y),$$

where "$*$" is the usual space–time convolution, and "$\star$" is the forward convolution in space–time in the sense that

$$(f \star g)(x) = \int f(x + y)g(y)\,dy.$$

Assumption 1.3 on the symmetry of covariance implies that $\rho_\delta^{\star 2}$ is also symmetric in its space variable. Since $(P')^{\star 2}$ also has the same symmetry, so does $\mathbf{E}(\Psi_\varepsilon(x)\Psi_\varepsilon(y))$.

For the second claim, it is easy to check that $(P')^{\star 2} = \frac{1}{2}P$ and $\rho_\varepsilon^{\star 2} = (\rho^{\star 2})_\varepsilon$. Thus, both bounds in (4.1) follow from the properties of the heat kernel. □

By stationarity, we let

$$\varrho_\varepsilon(x - y) = \mathbf{E}\big(\Psi_\varepsilon(x)\Psi_\varepsilon(y)\big).$$

The following lemma, based on Lemma 4.1 and the triangle inequality, will also be used throughout Section 6 to get the general pointwise bound.

LEMMA 4.2. *Let $K \geq 1$. For every $z, z' \in \mathbf{R}^+ \times \mathbf{T}$ with $|z'| \leq K|z|$ in the parabolic metric, we have*

(4.2) $$\varrho_\varepsilon(z) \leq K\Lambda^2 \varrho_\varepsilon(z').$$

PROOF. *By Lemma* 4.1, *we have*

$$\varrho_\varepsilon(z') \geq \frac{1}{\Lambda(|z'| + \varepsilon)} \geq \frac{1}{K\Lambda^2} \cdot \frac{\Lambda}{|z| + \varepsilon} \geq \frac{1}{K\Lambda^2} \cdot \varrho_\varepsilon(z),$$

*where in the second inequality we have used the assumption $|z'| \leq K|z|$ as well as $K \geq 1$.* □

4.2. *Interchanging the supremum with the expectation.* Fix an integer $d \geq 1$. For each $k = 1, \ldots, d$, let $J_k = [c_k, C_k]$ be a fixed interval. For every function $\Phi$ on $\mathcal{J} = J_1 \times \cdots \times J_d$ and every subset $I = \{i_1, \ldots, i_m\} \subset \{1, \ldots, d\}$, we write

$$\partial_I \Phi = \partial_{i_1} \cdots \partial_{i_m} \Phi.$$

We also let $\Phi_I$ be the function of $|I|$ variables such that

$$\Phi_I : J_{i_1} \times \cdots \times J_{i_m} \to \mathbf{R}$$

is the restriction of $\Phi$ to variables from $I$ while the remaining variables take the value $c_k$ for $k \notin I$. For convenience, we write

$$\mathcal{J}_I = J_{i_1} \times \cdots \times J_{i_m},$$

and use $|\mathcal{J}_I|$ to denote the volume of $\mathcal{J}_I$. Finally, we let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ and $\boldsymbol{\theta}_I = (\theta_{i_1}, \ldots, \theta_{i_m})$. With this notation, we have the following lemma.

LEMMA 4.3. *Let $\Phi$ be a random smooth function on $\mathcal{J} = J_1 \times \cdots \times J_d$. Then we have*

$$\mathbf{E} \sup_{\boldsymbol{\theta} \in \mathcal{J}} |\Phi(\boldsymbol{\theta})|^{2n} \leq 2^{2nd} \sum_I |\mathcal{J}_I|^{2n} \sup_{\boldsymbol{\theta} \in \mathcal{J}} \mathbf{E} |(\partial_I \Phi)(\boldsymbol{\theta})|^{2n},$$

*where the sum is taken over all subsets $I$ of $\{1, \ldots, d\}$, and $|\mathcal{J}_I| = 1$ if $I = \varnothing$.*

PROOF. By repeatedly applying the fundamental theorem of calculus, we see that for every $\boldsymbol{\theta} \in \mathcal{J}$ one has the identity

$$\Phi(\boldsymbol{\theta}) = \sum_{m=0}^d \sum_{I=(i_1,\ldots,i_m)} \int_{c_{i_1}}^{\theta_{i_1}} \cdots \int_{c_{i_m}}^{\theta_{i_m}} (\partial_I \Phi)_I(\mathbf{u}_I) \, d\mathbf{u}_I,$$

where the sum on the right-hand side is taken over all size-$m$ subsets $I$ of $\{1, \ldots, d\}$, and then over all $m$. We have also used the convention that the term

corresponding to $m = 0$ (or $I = \varnothing$) is the constant term $\Phi(c_1, \ldots, c_d)$. We then have the bound

$$\sup_{\boldsymbol{\theta} \in \mathcal{J}} |\Phi(\boldsymbol{\theta})| \leq \sum_I \int_{\mathcal{J}_I} |(\partial_I \Phi)_I(\mathbf{u}_I)| \, d\mathbf{u}_I.$$

Since there are $2^d$ terms in the sum, raising both sides to the power $2n$, and using Hölder's inequality for each of the integrals on the right-hand side, we obtain

$$\sup_{\boldsymbol{\theta} \in \mathcal{J}} |\Phi(\boldsymbol{\theta})|^{2n} \leq 2^{2nd} \sum_I |\mathcal{J}_I|^{2n-1} \int_{\mathcal{J}_I} |(\partial_I \Phi)_I(\mathbf{u}_I)|^{2n} \, d\mathbf{u}_I.$$

Taking expectation on both sides, and replacing the integrals by the supremum norm, we get

$$\mathbf{E} \sup_{\boldsymbol{\theta} \in \mathcal{J}} |\Phi(\boldsymbol{\theta})|^{2n} \leq 2^{2nd} \sum_I |\mathcal{J}_I|^{2n} \sup_{\boldsymbol{\theta}_I \in \mathcal{J}_I} \mathbf{E} |(\partial_I \Phi)_I(\boldsymbol{\theta}_I)|^{2n}.$$

Finally, note that the $(\partial_I \Phi)_I$ is a function of $|I|$ variables, and the other $d - |I|$ ones are fixed at values $c_k$'s. Thus, by allowing taking the supremum of all the variables $\theta_k$'s, we enlarge the upper bound to match the statement of the lemma. This completes the proof. $\square$

REMARK 4.4. In the proof of the lemma, one controls the $L^\infty$ norm of $\Phi$ by the $L^1$ norms of its derivatives up to order $d$. From Sobolev embedding, one can reduce the number of derivatives by increasing the integrability of them on the right-hand side. But it turns out later that in our context, both ways yield the same control for $(\mathbf{E} \sup_{\boldsymbol{\theta}} |\Phi(\boldsymbol{\theta})|^{2n})^{\frac{1}{2n}}$.

4.3. *Localisation and decomposition of $\widehat{F}$.* For every $d \geq 1$ and $\boldsymbol{\theta} \in \mathbf{R}^d$, let $\mathfrak{R}_{\boldsymbol{\theta}}$ be the rectangle of side length 2 in $\mathbf{R}^d$ centred at $\boldsymbol{\theta}$. The value of $d$ varies in different contexts in this article, and hence so is the dimension of $\mathfrak{R}_{\boldsymbol{\theta}}$. We let $|\boldsymbol{\theta}| := \sum_{i=1}^d |\theta_i|$. For every open bounded set $\Omega \subset \mathbf{R}^d$ and $M \in \mathbf{N}$, we define a norm $\| \cdot \|_{\mathcal{B}_M(\Omega)}$ on $\mathcal{C}_c^\infty(\Omega)$ functions by

$$\|\boldsymbol{\phi}\|_{\mathcal{B}_M(\Omega)} := \sup_{\mathbf{r}: 0 \leq r_i \leq M} \sup_{\mathbf{x} \in \Omega} |(\partial^{\mathbf{r}} \boldsymbol{\phi})(\mathbf{x})|,$$

where $\mathbf{x} = (x_1, \ldots, x_d)$, $\mathbf{r} = (r_1, \ldots, r_d)$, and $\partial^{\mathbf{r}} = \partial_{x_1}^{r_1} \cdots \partial_{x_d}^{r_d}$. Let $\mathcal{B}_M(\Omega)$ be the closure of $\mathcal{C}_c^\infty(\Omega)$ under that norm. The space $\mathcal{B}_M$ is different from the usual Hölder space $\mathcal{C}^M$ in that it requires every directional derivative of $\boldsymbol{\phi}$ up to order $M$, not just the total number of its derivatives. But they do coincide when $d = 1$. For every integer $M$ and open set $\Omega \subset \mathbf{R}^d$, we define a norm $\| \cdot \|_{M, \Omega}$ on distributions on $\mathbf{R}^d$ by

$$\|\Upsilon\|_{M, \Omega} := \sup_{\boldsymbol{\phi}: \|\boldsymbol{\phi}\|_{\mathcal{B}_M(\Omega)} \leq 1} |\langle \Upsilon, \boldsymbol{\phi} \rangle|,$$

where the supremum is taken over all $\phi \in \mathcal{C}_c^\infty(\Omega)$ with the $\mathcal{B}_M(\Omega)$ norm bounded by 1. We have the following lemma.

LEMMA 4.5. *Suppose $\Upsilon$ is a distribution on $\mathbf{R}^d$ with the form $\Upsilon = \bigotimes_{i=1}^{d} \Upsilon^i$, where each $\Upsilon^i$ is a distribution on $\mathbf{R}$. Then we have*

$$\|\Upsilon\|_{M, \mathfrak{R}_{\mathbf{K}}} \leq \prod_{i=1}^{d} \|\Upsilon^i\|_{M, \mathfrak{R}_{K_i}}$$

*for all $\mathbf{K} = (K_1, \ldots, K_d) \in \mathbf{Z}^d$.*

PROOF. The case $d = 1$ is immediate. For $d \geq 2$, we have

$$|\langle \Upsilon, \boldsymbol{\phi} \rangle| = \left| \left\langle \Upsilon^d, \left\langle \bigotimes_{i=1}^{d-1} \Upsilon^i, \boldsymbol{\phi}(\cdot, \ldots, \cdot, x_d) \right\rangle \right\rangle \right|$$

$$\leq \|\Upsilon^d\|_{M, \mathfrak{R}_{K_d}} \cdot \sup_{r_d \leq M} \sup_{x_d \in \mathfrak{R}_{K_d}} \left| \left\langle \bigotimes_{i=1}^{d-1} \Upsilon^i, (\partial_{x_d}^{r_d} \phi)(\cdot, \ldots, \cdot, x_d) \right\rangle \right|,$$

and the claim follows by induction. $\quad\square$

PROPOSITION 4.6. *Let $\rho$ be a mollifier on $\mathbf{R}$. For every $\delta > 0$, let $\rho_\delta(\cdot) = \delta^{-1} \rho(\cdot/\delta)$. For $i = 1, \ldots, d$, let $G^i$ be a Schwartz distribution on $\mathbf{R}$ and write $G_\delta^i = G^i * \rho_\delta$. Let*

$$\Upsilon = \bigotimes_{i=1}^{d} \widehat{G^i}, \qquad \Upsilon_\delta = \bigotimes_{i=1}^{d} \widehat{G_\delta^i},$$

*where $\widehat{G^i}$ and $\widehat{G_\delta^i}$ are Fourier transforms of $G^i$ and $G_\delta^i$. Then for every $M, N \in \mathbf{Z}$, there exists $C = C(d, M, N) > 0$ such that*

$$(4.3) \qquad \|\Upsilon_\delta\|_{M, \mathfrak{R}_{\mathbf{K}}} \leq C \delta^{-N} (1 + |\mathbf{K}|)^{-N} \prod_{i=1}^{d} \|\widehat{G^i}\|_{M, \mathfrak{R}_{K_i}}$$

*for all $\mathbf{K} = (K_1, \ldots, K_d) \in \mathbf{Z}^d$ and all $\delta \in (0, 1)$. Furthermore, there exists $C = C(d, M) > 0$ such that*

$$(4.4) \qquad \|\Upsilon - \Upsilon_\delta\|_{M, \mathfrak{R}_{\mathbf{K}}} \leq C \delta^\beta \prod_{i=1}^{d} [(1 + |K_i|)^\beta \|\widehat{G^i}\|_{M, \mathfrak{R}_{K_i}}]$$

*for all $\mathbf{K} \in \mathbf{Z}^d$, all $\beta \in (0, 1)$ and all $\delta \in (0, 1)$.*

PROOF. We first prove the bound for $\Upsilon_\delta$. By Lemma 4.5, we have

$$(4.5) \qquad \|\Upsilon_\delta\|_{M, \mathfrak{R}_{\mathbf{K}}} \leq \prod_{i=1}^{d} \|\widehat{G_\delta^i}\|_{M, \mathfrak{R}_{K_i}}.$$

For every $i$, we have $\langle \widehat{G_\delta^i}, \phi \rangle = \langle \widehat{G^i}, \widehat{\rho}(\delta \cdot) \phi \rangle$. Since for every $\phi \in \mathcal{B}_M(\mathfrak{R}_{K_i})$, the function $\widehat{\rho}(\delta \cdot) \phi$ also belongs to $\mathcal{B}_M(\mathfrak{R}_{K_i})$ with the bound

$$\| \widehat{\rho}(\delta \cdot) \phi \|_{\mathcal{B}_M(\mathfrak{R}_{K_i})} \leq C_{M,N} (1 + \delta |K_i|)^{-N} \| \phi \|_{\mathcal{B}_M(\mathfrak{R}_{K_i})},$$

we can deduce that

$$(4.6) \qquad \| \widehat{G_\delta^i} \|_{M,\mathfrak{R}_{K_i}} \leq C_{M,N} (1 + \delta |K_i|)^{-N} \| \widehat{G^i} \|_{M,\mathfrak{R}_{K_i}}.$$

Plugging (4.6) into (4.5) and noticing that

$$\prod_i (1 + \delta |K_i|)^{-1} \leq (1 + \delta |\mathbf{K}|)^{-1} \leq \delta^{-1} (1 + |\mathbf{K}|)^{-1}$$

for all $\delta < 1$, we obtain the bound (4.3) for $T_\delta$. As for the difference $\Upsilon - \Upsilon_\delta$, it can be expressed by

$$\Upsilon - \Upsilon_\delta = \sum_{j=1}^{d} \left[ \bigotimes_{i=1}^{j-1} \widehat{G_\delta^i} \otimes (\widehat{G^j} - \widehat{G_\delta^j}) \otimes \bigotimes_{i=j+1}^{d} \widehat{G^i} \right].$$

By Lemma 4.5 and the bound (4.6) with $N = 0$, we have

$$\| \Upsilon - \Upsilon_\delta \|_{M,\mathfrak{R}_{\mathbf{K}}} \leq C_M \sum_{j=1}^{d} \left[ \| \widehat{G^j} - \widehat{G_\delta^j} \|_{M,\mathfrak{R}_{K_j}} \prod_{i \neq j} \| \widehat{G^i} \|_{M,\mathfrak{R}_{K_i}} \right].$$

It then suffices to bound $\| \widehat{G^j} - \widehat{G_\delta^j} \|_{M,\mathfrak{R}_{K_j}}$ for each $j$. In fact, since $\langle \widehat{G^j} - \widehat{G_\delta^j}, \phi \rangle = \langle \widehat{G^j}, (1 - \widehat{\rho}(\delta \cdot)) \phi \rangle$, and

$$\| (1 - \widehat{\rho}(\delta \cdot)) \phi \|_{\mathcal{B}_M(\mathfrak{R}_{K_j})} \leq C_M \delta^\beta (1 + |K_j|)^\beta \| \phi \|_{\mathcal{B}_M(\mathfrak{R}_{K_j})}$$

for every $\beta \in (0,1)$, we deduce immediately from the definition of $\| \cdot \|_{M,\mathfrak{R}_{K_j}}$ that

$$\| \widehat{G^j} - \widehat{G_\delta^j} \|_{M,\mathfrak{R}_{K_j}} \leq C_M \delta^\beta (1 + |K_j|)^\beta \| \widehat{G^j} \|_{M,\mathfrak{R}_{K_j}}.$$

The bound (4.4) then follows. $\quad \square$

LEMMA 4.7. *Let $G$ be a Schwartz distribution on $\mathbf{R}$. Then, for every $\ell, M \in \mathbf{N}$, there exists a constant $C > 0$ such that*

$$\| \widehat{G^{(\ell)}} \|_{M,\mathfrak{R}_K} \leq C (1 + |K|)^\ell \| \widehat{G} \|_{M,\mathfrak{R}_K}$$

*for all $K \in \mathbf{Z}$.*

PROOF. We have $\langle \widehat{G^{(\ell)}}, \phi \rangle = \langle \widehat{G}, \tilde{\phi} \rangle$, where $\tilde{\phi}(\theta) = (i\theta)^\ell \phi(\theta)$. For $\phi \in \mathcal{C}_c^\infty(\mathfrak{R}_K)$, $\tilde{\phi}$ also belongs to $\mathcal{C}_c^\infty(\mathfrak{R}_K)$ with the bound

$$\| \tilde{\phi} \|_{\mathcal{B}_M(\mathfrak{R}_K)} \lesssim_{\ell,M} (1 + |K|)^\ell \| \phi \|_{\mathcal{B}_M(\mathfrak{R}_K)}.$$

The claim then follows immediately. $\quad \square$

LEMMA 4.8.   *Suppose $F \in \mathcal{C}^{k+\alpha}$ for some integer $k$ and $\alpha \in [0, 1)$, and its derivatives satisfy the bounds*

$$\sup_{0 \leq \ell \leq k} \left| F^{(\ell)}(x) \right| \leq C(1 + |x|)^M,$$

$$\sup_{|h| < 1} \frac{|F^{(k)}(x + h) - F^{(k)}(x)|}{|h|^\alpha} \leq C(1 + |x|)^M$$

*for all $x \in \mathbf{R}$. Then there exists $C = C(k, M) > 0$ such that*

$$\|\widehat{F}\|_{M+2, \mathfrak{R}_K} \leq C(1 + |K|)^{-k-\alpha}$$

*for every $K \in \mathbf{Z}$.*

PROOF.   For $\phi : \mathbf{R} \to \mathbf{R}$, let $\phi_K(\cdot) = \phi(\cdot - K)$. By definition, we have

(4.7)       $$\|\widehat{F}\|_{M+2, \mathfrak{R}_K} = \sup_{\phi} |\langle \widehat{F}, \phi_K \rangle| = \sup_{\phi} \left| \int_{\mathbf{R}} F(x) \widehat{\phi}(x) e^{-iKx} \, dx \right|,$$

where the supremum is taken over all $\phi \in \mathcal{C}_c^{\infty}(\mathfrak{R}_0)$ such that $\|\phi\|_{\mathcal{C}^{M+2}(\mathfrak{R}_0)} \leq 1$. Here, the $B_M$ and $C^M$ norms are equivalent since we are in dimension one. We need to look at $\| \cdot \|_{M+2; \mathfrak{R}_K}$ norm of $\widehat{F}$ since $F$ has growth of order $M$, and the $(M + 2)$-differentiability of $\phi$ implies the derivatives of its Fourier transform has the decay

$$|\widehat{\phi}^{(\ell)}(x)| \lesssim_\ell (1 + |x|)^{-M-2}$$

for all $\ell \geq 1$, so the right-hand side of (4.7) is integrable. If $k = 0$, we can write the right-hand side of (4.7) as

$$\int_{\mathbf{R}} F(x) \widehat{\phi}(x) e^{-iKx} \, dx$$

$$= \frac{1}{2} \int_{\mathbf{R}} \left( F(x) \widehat{\phi}(x) - F\left(x - \frac{\pi}{K}\right) \widehat{\phi}\left(x - \frac{\pi}{K}\right) \right) e^{-iKx} \, dx.$$

Then, by the Hölder continuity of $F$ and the decay of $\widehat{\phi}$ and its derivatives, we have

$$\|\widehat{F}\|_{M+2, \mathfrak{R}_K} \leq C(1 + |K|)^{-\alpha}.$$

If $k \geq 1$, integrating by parts $k$ times, we have (for $|K| \neq 0$)

$$\left| \int_{\mathbf{R}} F(x) \widehat{\phi}(x) e^{-iKx} \, dx \right|$$

$$= |K|^{-k} \left| \sum_{\ell=0}^{k} \binom{k}{\ell} \int_{\mathbf{R}} F^{(\ell)}(x) \widehat{\phi}^{(k-\ell)}(x) e^{-iKx} \, dx \right|.$$

For the terms with $\ell \leq k - 1$ in the sum above, one can further integrate by parts to get the decay of order $\mathcal{O}(|K|^{-k-1})$. For $\ell = k$, since $F^{(k)} \in \mathcal{C}^\alpha$, the decay $\mathcal{O}(|K|^{-k-\alpha})$ follows from the prefactor $|K|^{-k}$ and the previous case for $\mathcal{C}^\alpha$ functions. This completes the proof. $\quad\square$

The next proposition is the main property of $F$ we will use in this article.

PROPOSITION 4.9. *Let $F \colon \mathbf{R} \to \mathbf{R}$ satisfy Assumption* 1.1. *Let $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d) \in \mathbf{N}^d$, and*

$$\Upsilon = \bigotimes_{i=1}^d \widehat{F^{(\ell_i)}}, \qquad \Upsilon_\delta = \bigotimes_{i=1}^d \widehat{F_\delta^{(\ell_i)}}.$$

*Then, for every $N > 0$, there exists a constant $C$ depending on $N$ and $\boldsymbol{\ell}$ such that*

$$(4.8) \qquad \|\Upsilon_\delta\|_{M+2,\mathfrak{R}_{\mathbf{K}}} \leq C\delta^{-N}(1 + |\mathbf{K}|)^{-N}$$

*for all $\mathbf{K} \in \mathbf{Z}^d$ and $\delta \in (0, 1)$. For the difference $T - T_\delta$, we have the bound*

$$(4.9) \qquad \|\Upsilon - \Upsilon_\delta\|_{M+2,\mathfrak{R}_{\mathbf{K}}} \leq C\delta^\beta \prod_{i=1}^d (1 + |K_i|)^{-7-\alpha+\ell_i+\beta}$$

*for all $\mathbf{K} \in \mathbf{Z}^d$, $\delta \in (0, 1)$ and $\beta \in (0, 1)$. Here, $\alpha$ and $M$ are the same as in Assumption* 1.1.

PROOF. This follows directly from Assumption 1.1, Lemmas 4.7, 4.8 and Proposition 4.6 by setting $G^i = F^{(\ell_i)}$. $\quad\square$

We also need the following proposition later.

PROPOSITION 4.10. *Let $\Upsilon$ be a distribution on $\mathbf{R}^d$ and $\Phi \in \mathcal{C}^\infty(\mathbf{R}^d)$, which is allowed to grow at infinity. Then we have*

$$(4.10) \qquad |\langle \Upsilon, \Phi \rangle| \lesssim_M \sum_{\mathbf{K} \in \mathbf{Z}^d} \|\Upsilon\|_{M+2,\mathfrak{R}_{\mathbf{K}}} \sup_{\mathbf{r}:r_j \leq M+2} \sup_{\boldsymbol{\theta} \in \mathfrak{R}_{\mathbf{K}}} |\partial^{\mathbf{r}} \Phi(\boldsymbol{\theta})|.$$

PROOF. Multiply $\Phi$ by a partition of unity and estimate the terms separately. $\quad\square$

4.4. *Behaviour of the coefficients in the chaos expansion.* In [25], after Wick renormalisation, each homogeneous chaos component of a polynomial model has logarithmically divergent, finite and vanishing parts as $\varepsilon \to 0$. The same is true in our context with a general $F$. Thus, in order to identify the limits of our models with that of the KPZ equation, we need to distinguish these three parts for each of our objects. The situation here is more complicated as the models are constructed

from a general function $F$ rather than a polynomial, so we need to get fine control on the coefficients of the terms in the chaos expansion. The aim of this section is to give a few lemmas that provide the necessary controls for the object $\tau = \overset{\bullet}{\underset{\circ}{\circ}}$. This is the most complicated object among those in Table (3.1), and will be treated in detail in Section 5.

Throughout, we fix the mollifier $\rho$ on $\mathbf{R}$. For every $\delta > 0$, write $\rho_\delta(\cdot) = \delta^{-1}\rho(\cdot/\delta)$ and $F_\delta = F * \rho_\delta$. Let $a_\delta$ be defined as the same as the coupling constant $a$ in (1.4) except that one replaces $F''$ by $F_\delta''$ in (1.4). It is easy to check that $|a - a_\delta| \lesssim \delta$. For space–time points $x$, $y$, $z$, we will use the notation

$$(4.11) \qquad X = \varepsilon^{\frac{1}{2}}\Psi_\varepsilon(x), \qquad Y = \varepsilon^{\frac{1}{2}}\Psi_\varepsilon(y), \qquad Z = \varepsilon^{\frac{1}{2}}\Psi_\varepsilon(z).$$

We omit the dependence of these random fields on $\varepsilon$ for notational simplicity. Our aim is to have control on the coefficients that is uniform over all space–time points and all $\varepsilon > 0$. Also, recall the recentering notation $\{W\} = W - \mathbf{E}W$. We have the following lemmas.

LEMMA 4.11. *With the notation* (4.11), *we have the decomposition*

$$\mathbf{E}(F''(X)F'(Y)F'(Z)) - 8a^3\mathbf{E}YZ$$

$$(4.12) \qquad = \mathbf{E}(\{F''(X)\}F'(Y)F'(Z)) + 2a\mathbf{E}(F'(Y)F'(Z) - 4a^2YZ).$$

*For the first term on the right-hand side, there exist* $N, \beta > 0$ *such that*

$$(4.13) \qquad \left|\mathbf{E}(\{F''(X)\}F'(Y)F'(Z))\right| \lesssim \delta^{-N}(\mathbf{E}XY)(\mathbf{E}XZ) + \delta^\beta\mathbf{E}YZ,$$

*where the proportionality constant is uniform over* $\delta \in (0, 1)$. *For the second term, there exists* $\beta' > 0$ *such that*

$$(4.14) \qquad \left|\mathbf{E}(F'(Y)F'(Z) - 4a^2YZ)\right| \lesssim (\mathbf{E}YZ)^{1+\beta'}.$$

*Both bounds above are uniform over all* $\varepsilon \in (0, 1)$ *and all locations of the points* $x, y, z$.

PROOF. The identity (4.12) is immediate since $\mathbf{E}F''(X) = 2a$. To see the bound (4.13), we let $\Upsilon = \widehat{F''} \otimes \widehat{F'} \otimes \widehat{F'}$, and $T_\delta$ be the distribution that replaces every appearance of $F^{(\ell)}$ in $\Upsilon$ by $F_\delta^{(\ell)}$. We then write

$$\mathbf{E}(\{F''(X)\}F'(Y)F'(Z)) = \langle \Upsilon_\delta, \mathbf{E}\Phi_\varepsilon \rangle + \langle \Upsilon - \Upsilon_\delta, \mathbf{E}\Phi_\varepsilon \rangle,$$

where

$$\Phi_\varepsilon(\boldsymbol{\theta}; x, y, z) = \{\cos(\theta_\mathfrak{x}X)\}\sin(\theta_\mathfrak{y}Y)\sin(\theta_\mathfrak{z}Z),$$

is viewed as a test function in the variable $\boldsymbol{\theta} = (\theta_\mathfrak{x}, \theta_\mathfrak{y}, \theta_\mathfrak{z})$. Following the notation in Section 6.1, the type space here is $\tilde{\mathcal{T}} = \{\mathfrak{x}, \mathfrak{y}, \mathfrak{z}\}$ with $\tilde{\mathcal{O}} = \{\mathfrak{y}, \mathfrak{z}\}$ and $\tilde{\mathcal{E}} = \{\mathfrak{x}\}$. For the term $\langle \Upsilon_\delta, \mathbf{E}\Phi_\varepsilon \rangle$, we have by Proposition 4.10

$$(4.15) \qquad \left|\langle \Upsilon_\delta, \mathbf{E}\Phi_\varepsilon \rangle\right| \leq \sum_{\mathbf{K} \in \mathbf{Z}^3} \|\Upsilon_\delta\|_{M+2; \mathfrak{R}_\mathbf{K}} \sup_{|\mathbf{r}|_\infty \leq M+2} \sup_{\boldsymbol{\theta} \in \mathfrak{R}_\mathbf{K}} \left|\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\mathbf{E}\Phi_\varepsilon\right|,$$

where $|\mathbf{r}|_\infty$ denotes the largest component of the multi-index $\mathbf{r}$. Interchanging differentiation and expectation and applying the bound (6.6), we get

$$\sup_{\boldsymbol{\theta}\in\mathfrak{R}_\mathbf{K}}\left|\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\mathbf{E}\Phi_\varepsilon\right| = \sup_{\boldsymbol{\theta}\in\mathfrak{R}_\mathbf{K}}\left|\mathbf{E}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon)\right| \lesssim (1+|\mathbf{K}|)^4(\mathbf{E}XY)(\mathbf{E}XZ).$$

Here, the power of $|\mathbf{K}|$ is 4 since $\tilde{\mathcal{O}}$ has cardinality 2 and $\tilde{\mathcal{E}}$ is a singleton. Plugging the above bound back into (4.15) and applying (4.8) to the term $\|\Upsilon_\delta\|_{M+2;\mathfrak{R}_\mathbf{K}}$, we get

$$\left|\langle\Upsilon_\delta, \mathbf{E}\Phi_\varepsilon\rangle\right| \lesssim \delta^{-N}(\mathbf{E}XY)(\mathbf{E}XZ)$$

for some $N$. As for the term with $\Upsilon - \Upsilon_\delta$, we let

$$\boldsymbol{\phi}_\varepsilon = \cos(\theta_{\mathfrak{r}}X)\sin(\theta_{\mathfrak{y}}Y)\sin(\theta_{\mathfrak{z}}Z),$$

and write $\Phi_\varepsilon = \boldsymbol{\phi}_\varepsilon + (\Phi_\varepsilon - \boldsymbol{\phi}_\varepsilon)$. We estimate the action of $\Upsilon - \Upsilon_\delta$ on these two terms separately. For the one with $\boldsymbol{\phi}_\varepsilon$, (6.7) implies

$$\sup_{\boldsymbol{\theta}\in\mathfrak{R}_\mathbf{K}}\left|\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\mathbf{E}\boldsymbol{\phi}_\varepsilon\right| \lesssim_{\mathbf{r}} (1+|\mathbf{K}|)^2(\mathbf{E}YZ).$$

Here, there are only two powers on $|\mathbf{K}|$ since we counted the cardinality of the set $\tilde{\mathcal{O}}$ only. We then use Proposition 4.10 and the bound (4.9) to conclude that

$$\left|\langle\Upsilon - \Upsilon_\delta, \mathbf{E}\boldsymbol{\phi}_\varepsilon\rangle\right| \lesssim \delta^\beta \mathbf{E}YZ.$$

The same bound holds for $\langle\Upsilon - \Upsilon_\delta, \Phi_\varepsilon - \boldsymbol{\phi}_\varepsilon\rangle$. The derivation is only simpler since the field involving $X$ is averaged out. This proves the bound (4.13).

To see (4.14), one can write $F'$ in terms of chaos expansion and use Wick's formula to get the expression

$$(4.16) \qquad \mathbf{E}\big(F'(Y)F'(Z) - 4a^2YZ\big) = \sum_{n\geq 1}\frac{c_n^2}{(2n+1)!}(\mathbf{E}YZ)^{2n+1},$$

where $c_n = \mathbf{E}(F^{(2n+2)}(Y))$. Let $\sigma^2 = \mathbf{E}Y^2 = \mathbf{E}Z^2$. Since

$$\mathbf{E}\big(F'(Y)\big)^2 = \sum_{n\geq 0}\frac{c_n^2\sigma^{4n+2}}{(2n+1)!} < +\infty,$$

and $\mathbf{E}YZ \leq \sigma^2$, we can replace the term $(\mathbf{E}YZ)^{2n+1}$ in (4.16) by $\sigma^{2(2n-\beta')} \times (\mathbf{E}YZ)^{1+\beta'}$ and still get a convergent series. This implies the bound (4.14), and hence completes the proof. $\square$

LEMMA 4.12.  *There exist $\beta, N > 0$ such that*

$$\left|\mathbf{E}\big(F''(X)F''(Y)F''(Z)\big) - 8a^3\right|$$
$$\lesssim \delta^{-N}\big((\mathbf{E}XY)^2 + (\mathbf{E}YZ)^2 + (\mathbf{E}ZX)^2\big) + \delta^\beta.$$

PROOF. Again, we decompose the quantity into a main smooth part and a small remainder by

$$\mathbf{E}(F''(X)F''(Y)F''(Z)) - 8a^3$$
$$= \mathbf{E}(F_\delta''(X)F_\delta''(Y)F_\delta''(Z)) - 8a_\delta^3 + \text{remainder},$$

where $a_\delta$ is given at the beginning of this subsection. The remainder is easily seen to be bounded by $\delta^\beta$ for some $\beta > 0$. As for the main smooth part, we write it as

$$\mathbf{E}(F_\delta''(X)F_\delta''(Y)F_\delta''(Z)) - 8a_\delta^3$$
$$= \mathbf{E}(\{F_\delta''(X)\}\{F_\delta''(Y)\}\{F_\delta''(Z)\})$$
$$+ 2a_\delta\mathbf{E}(\{F_\delta''(X)\}\{F_\delta''(Y)\} + \{F_\delta''(Y)\}\{F_\delta''(Z)\} + \{F_\delta''(Z)\}\{F_\delta''(X)\})$$
$$+ 4a_\delta^2\mathbf{E}(\{F_\delta''(X)\} + \{F_\delta''(Y)\} + \{F_\delta''(Z)\}).$$

The last term above vanishes by definition of $\{\cdot\}$. For the second term, the same argument via chaos expansion as in the proof for (4.14) gives the bound $(\mathbf{E}XY)^2 + (\mathbf{E}YZ)^2 + (\mathbf{E}ZX)^2$ uniformly in $\delta$. As for the first one, we repeat the same argument as for the term $\langle T_\delta, \mathbf{E}\Phi_\varepsilon \rangle$ in Lemma 4.11 and apply (6.6) to obtain the bound

$$\mathbf{E}(\{F_\delta''(X)\}\{F_\delta''(Y)\}\{F_\delta''(Z)\})$$
$$\lesssim \delta^{-N}(\mathbf{E}XY)(\mathbf{E}YZ)(\mathbf{E}ZX)$$
$$\lesssim \delta^{-N}((\mathbf{E}XY)^2 + (\mathbf{E}YZ)^2 + (\mathbf{E}ZX)^2).$$

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

LEMMA 4.13. *We have the decomposition*

$$(4.17) \qquad \mathbf{E}(F'(X)F'(Y)\{F(Z)\} - 4a^3XYZ^{\diamond 2}) = H^{(1)} + H^{(2)},$$

*where*

$$H^{(1)} = \mathbf{E}(\mathscr{T}_{(1)}(F'(X))F'(Y)\{F(Z)\}),$$
$$H^{(2)} = 2a\mathbf{E}(X\mathscr{T}_{(1)}(F'(Y))\{F(Z)\}),$$

*and we use the notation $\mathscr{T}_{(j)}(\cdot)$ from Section 6.1 to denote the random variable $\cdot$ obtained by removing its components belonging to the homogeneous Wiener chaoses of order up to (and including) $j$. Furthermore, $H^{(1)}$ and $H^{(2)}$ satisfy the following bounds and properties. For $H^{(1)}$, there exist $\beta, N > 0$ such that*

$$(4.18) \qquad |H^{(1)}| \lesssim \delta^{-N}(\mathbf{E}XY)(\mathbf{E}XZ)^2 + \delta^\beta(\mathbf{E}XZ)(\mathbf{E}YZ)$$

*for every $\delta > 0$. $H^{(2)}$ has the further decomposition*

$$H^{(2)} = (\mathbf{E}XY)\mathbf{E}(\{F(Y)\}\{F(Z)\})$$
$$(4.19) \qquad\qquad + (\mathbf{E}XZ)\mathbf{E}(\mathscr{T}_{(1)}(F'(Y))\mathscr{T}_{(1)}(F'(Z))),$$

M. HAIRER AND W. XU

*and there exists $\beta' > 0$ such that*

$$(4.20) \qquad \big|H^{(2)}\big| \lesssim (\mathbf{E}XY + \mathbf{E}XZ)(\mathbf{E}YZ)^{1+\beta'}.$$

PROOF. With the above definition of $H^{(j)}$, the difference between the left- and right-hand sides of (4.17) is $4a^2\mathbf{E}(XY\,\mathscr{T}_{(2)}(F(Z)))$, which is $0$ since $\mathscr{T}_{(2)}(F(Z))$ has components only in fourth and higher order chaos, and hence is orthogonal to $XY$. This verifies the decomposition (4.17).

The decomposition (4.19) can be verified by chaos expanding $F'(Y)$ and $F(Z)$ and applying Wick's formula. The bound for $H^{(2)}$ then follows in the same way as the argument for proving (4.14).

As for $H^{(1)}$, we let $\Upsilon = \widehat{F'} \otimes \widehat{F'} \otimes \widehat{F}$, and $\Upsilon_\delta$ be the distribution that replaces every $F^{(\ell)}$ in $\Upsilon$ by $F_\delta^{(\ell)}$. Similar as before, we write

$$H^{(1)} = \langle \Upsilon_\delta, \mathbf{E}\boldsymbol{\phi}_\varepsilon \rangle + \langle \Upsilon - \Upsilon_\delta, \mathbf{E}\boldsymbol{\phi}_\varepsilon \rangle,$$

where

$$\boldsymbol{\phi}_\varepsilon(\boldsymbol{\theta}; x, y, z) = \mathscr{T}_{(1)}\big(\sin(\theta_{\mathfrak{x}}X)\big)\sin(\theta_{\mathfrak{y}}Y)\{\!\{\cos(\theta_{\mathfrak{z}}Z)\}\!\}.$$

For the term with $\Upsilon_\delta$, the function $\boldsymbol{\phi}_\varepsilon$ does not fall exactly into the assumption of the general bounds in Section 6, but one can still get the bound

$$\sup_{\boldsymbol{\theta} \in \mathfrak{R}_{\mathbf{K}}} \big|\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\mathbf{E}\boldsymbol{\phi}_\varepsilon\big| \lesssim (1 + |\mathbf{K}|)^{N'}(\mathbf{E}XY)(\mathbf{E}XZ)^2$$

for some $N' > 0$. This can be achieved either by performing the same clustering argument as in Section 6 or by exact computation of $\mathbf{E}\boldsymbol{\phi}_\varepsilon$ using trigonometric identities (this is possible since there are only three terms in the product). Combining this bound with (4.8), we get

$$\big|\langle \Upsilon_\delta, \mathbf{E}\boldsymbol{\phi}_\varepsilon \rangle\big| \lesssim \delta^{-N}(\mathbf{E}XY)(\mathbf{E}XZ)^2.$$

As for the term $\langle \Upsilon - \Upsilon_\delta, \mathbf{E}\boldsymbol{\phi}_\varepsilon \rangle$, we let

$$\Phi_\varepsilon = \sin(\theta_{\mathfrak{x}}X)\sin(\theta_{\mathfrak{y}}Y)\{\!\{\cos(\theta_{\mathfrak{z}}Z)\}\!\},$$

and write

$$\boldsymbol{\phi}_\varepsilon = \Phi_\varepsilon + (\boldsymbol{\phi}_\varepsilon - \Phi_\varepsilon).$$

Note that $\Phi_\varepsilon$ precisely falls within the assumption of Theorem 6.4, and $\Phi_\varepsilon - \boldsymbol{\phi}_\varepsilon$ has the same form as the corresponding term appearing in $H^{(2)}$. We can then estimate the two terms separately. More precisely, applying Proposition 4.10 together with Theorem 6.4 and then (4.9), we get

$$\big|\langle \Upsilon - \Upsilon_\delta, \Phi_\varepsilon \rangle\big| \lesssim \delta^\beta(\mathbf{E}XZ)(\mathbf{E}YZ).$$

Using the same procedure for $H^{(2)}$ and applying (4.9), the other term can be bounded by $\delta^\beta(\mathbf{E}XZ)(\mathbf{E}YZ)^2$. Since we always have $\mathbf{E}YZ \lesssim 1$, it gives the same bound as $|\langle \Upsilon - \Upsilon_\delta, \Phi_\varepsilon \rangle|$. We have thus completed the proof of the lemma. $\quad\square$

**5. The main convergence theorem.** In this section, we will assume the bounds stated in Theorems 6.2 and 6.4, and use them to establish the convergence of our models to the limiting model which describes the KPZ equation. By stationarity of the input $\Psi_\varepsilon$, we can use 0 as the base point of our models without loss of generality. Again, $X$, $Y$ and $Z$ denote the value of the field $\varepsilon^{\frac{1}{2}}\Psi_\varepsilon(\cdot)$ at the point $x$, $y$ and $z$, respectively.

5.1. *Convergence of the models.* Recall that the renormalised model defined in (3.2) and (3.3) is built from $\Psi_\varepsilon$ with the renormalisation constants

$$
\begin{aligned}
C_\bullet^{(\varepsilon)} &= \frac{1}{a\varepsilon}\mathbf{E}F\big(\varepsilon^{\frac{1}{2}}\Psi_\varepsilon\big) = \frac{\widehat{a}}{\varepsilon}, \\
C_{\bullet\!\bullet}^{(\varepsilon)} &= \frac{1}{a^2\varepsilon^2}\int K(x-y)K(x-z)\mathbf{E}\big(\{F(Y)\}\{F(Z)\}\big)\,dy\,dz, \\
C_{\circ\!\circ}^{(\varepsilon)} &= \frac{1}{4a^3\varepsilon^2}\int K(x-y)K(y-z)\mathbf{E}\big(F'(X)F'(Y)\{F(Z)\}\big)\,dy\,dz, \\
C_{\!\Psi}^{(\varepsilon)} &= \frac{1}{2a^3\varepsilon^2}\int K(x-y)K(x-z)\mathbf{E}\big(\{F''(X)\}\{F(Y)\}\{F(Z)\}\big)\,dy\,dz.
\end{aligned}
$$

(5.1)

Here, $C_\tau^{(\varepsilon)}$ is precisely the mean of $\mathbf{\Pi}^\varepsilon\tau$. By stationarity of $\Psi_\varepsilon$, the above definitions of these constants are all independent of the location of $x$.

REMARK 5.1. Note that since $\mathbf{E}\{F''(X)\} = 0$, we also have

$$
C_{\!\Psi}^{(\varepsilon)} = \frac{1}{2a^3\varepsilon^2}\int K(x-y)K(x-z)\mathbf{E}\big(\{F''(X)\}\big(\{F(Y)\}\{F(Z)\}\big)\big)\,dy\,dz,
$$

which does indeed coincide with the mean of $\widehat{\Pi}^\varepsilon \mathbf{V}^\bullet \cdot \widehat{\Pi}^\varepsilon\square$, as mentioned just after (3.3).

We also note that the symbols $\bullet_\circ$ and $\bullet_\square$ are both of negative homogeneities. The reason why they do not exhibit renormalisations is because of the spatial anti-symmetry of $P'$, which leads the corresponding expectation to vanish.

With the above choice of the constants $C_\tau^{(\varepsilon)}$, we can show the convergence of our models to the KPZ model. This is the content of the following theorem.

THEOREM 5.2. *Let* $\Psi_\varepsilon = P' * \xi_\varepsilon$, *and* $\widehat{\Pi}^\varepsilon = \widehat{\mathcal{L}}_\varepsilon(\Psi_\varepsilon)$ *be the renormalised model defined in* (3.2) *and* (3.3) *with input* $\psi = \Psi_\varepsilon$ *and constants* $C_j^{(\varepsilon)}$ *defined in* (5.1). *Let* $\Pi^{\mathrm{KPZ}}$ *be the KPZ model described in the Appendix. Then there exists* $\zeta > 0$ *such that for every* $\tau$ *in* (3.1) *with* $|\tau| < 0$, *we have*

$$
\big(\mathbf{E}|\langle\widehat{\Pi}_z^\varepsilon\tau - \Pi_z^{\mathrm{KPZ}}\tau, \varphi_z^\lambda\rangle|^{2n}\big)^{\frac{1}{2n}} \lesssim_n \varepsilon^\zeta\lambda^{|\tau|+\zeta},
$$

(5.2)

*where the bound holds uniformly over all* $\varepsilon \in (0, 1)$, *all* $\lambda \in (0, 1)$ *and all space–time points* $z$ *in compact sets. As a consequence, we have* $\|\widehat{\Pi}^\varepsilon; \Pi^{\mathrm{KPZ}}\|_{\varepsilon,0} \to 0$ *in probability as* $\varepsilon \to 0$.

As long as the bound (5.2) holds for all $\tau$ with negative homogeneities, we can proceed as in [25], Proposition 6.3, to conclude the convergence of $\widehat{\Pi}^\varepsilon$ to the limiting KPZ model. From now on, we will focus on proving (5.2) for symbols with $|\tau| < 0$.

5.2. *Proof of Theorem* 5.2. According to Table (3.1), there are ten basis elements with negative homogeneities, and hence we need to check the bound (5.2) for all of them. The bounds on $\mathbf{\mathsf{?}}$ are a consequence of those for $\bullet$. Thus, there are nine essentially different ones to check.

For the sake of conciseness of the presentation, we provide details for three of them: $\circ$, $\overset{\bullet}{\underset{\circ}{\circ}}$ and $\overset{\bullet\,\bullet}{\sqcup}$. The element $\circ$ is simple but still illustrative enough to explain the general procedure. On the other hand, the symbol $\overset{\bullet}{\underset{\circ}{\circ}}$ is much more complicated, but contains all the subtleties that appear when dealing with the other symbols. Finally, $\overset{\bullet\,\bullet}{\sqcup}$ is the one whose convergence (to 0) requires the strongest differentiability assumption on $F$ ($\mathcal{C}^{7+}$), so we also include details for it.

Before we start giving details for the above three symbols, we list the requirement on the decay of local norm of $\|\widehat{F}\|$ that guarantees the convergence for each symbol:

(5.3)

| $\|\widehat{F}\|_{M+2,\mathfrak{R}_K}$ | 3 | 3 | 3 | 3 | 4 | 5 | 6 | 7 | 7 |
|---|---|---|---|---|---|---|---|---|---|
| $\tau$ | $\square$ | $\circ$ | $\bullet$ | $\overset{\bullet}{\underset{\circ}{\diagdown}}$ | $\overset{\bullet}{\underset{\circ}{\diagup}}$ | $\overset{\bullet\,\bullet}{\vee}$ | $\overset{\bullet}{\underset{\circ}{\circ}}$ | $\overset{\bullet}{\underset{\square}{\diagdown}}$ | $\overset{\bullet\,\bullet}{\sqcup}$ |

Here, the number $n$ in the first row indicates that we need the decay $\|\widehat{F}\|_{M+2,\mathfrak{R}_K} = \mathcal{O}(|K|^{-n-})$ in order for the corresponding process to converge. We will see that for the three symbols which we give detailed arguments below, the requirements listed above are indeed sufficient. The decay requirement for the other symbols can also be easily deduced in the same way. In the rest of the section, we use the notation

$$\mathbf{x} = (x_1, \ldots, x_{2n}), \qquad \boldsymbol{\varphi}^\lambda(\mathbf{x}) = \prod_{k=1}^{2n} \varphi^\lambda(x_k).$$

Also, $F_\delta^{(\ell)}$ is $F^{(\ell)}$ regularised by a symmetric mollifier at scale $\delta$. Hence, $F_\delta^{(\ell)}$ is even if $\ell$ is, and it is odd if $\ell$ is odd. Finally, by translation invariance, it suffices to check the bound (5.2) with $z = 0$.

5.2.1. *The case* $\tau = \circ$. We start with $\tau = \circ$. In what follows, we will always write $\tau_\varepsilon = \widehat{\Pi}_0^\varepsilon \tau$ for simplicity. Recall that $X = \varepsilon^{\frac{1}{2}} \Psi_\varepsilon(x)$. By (3.2), we have

$$\tau_\varepsilon(x) = \frac{1}{2a\sqrt{\varepsilon}} F'(X) = \frac{1}{2a\sqrt{\varepsilon}} \mathscr{T}_{(1)}(F'(X)) + \Psi_\varepsilon(x)$$

$$=: \tau_\varepsilon^{(1)}(x) + \tau_\varepsilon^{(2)}(x),$$

where $\mathcal{T}_{(1)}(F'(X))$ is $F'(X)$ with the first chaos removed. Since $\Psi_\varepsilon \to \Psi$ in probability in $\mathcal{C}^{-\frac{1}{2}-\kappa}$, it suffices to check $\tau_\varepsilon^{(1)}$ vanishes in the same topology as $\varepsilon \to 0$.

Following the notation in Section 6, we let $\Phi_\varepsilon(\theta, x) = \sin(\theta X)$. The dependence of $\Phi$ on $\varepsilon$ is via $X = \varepsilon^{\frac{1}{2}}\Psi_\varepsilon(x)$. Also recall that for every test function $\varphi$ on $\mathbf{R}^+ \times \mathbf{T}$ and $\lambda > 0$, we write $\varphi^\lambda(x) = \lambda^{-3}\varphi(x/\lambda)$. For every function $\mathcal{F} : \mathbf{R} \times (\mathbf{R}^+ \times \mathbf{T}) \mapsto \mathbf{R}$, let

$$(\mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{F})(\theta) := \frac{1}{2a\sqrt{\varepsilon}} \int_{\mathbf{R}^+ \times \mathbf{T}} \mathcal{F}(\theta, x)\varphi^\lambda(x)\,dx,$$

where we have omitted in notation the dependence of $\mathcal{A}$ on $\varphi$ for simplicity. We split $F'$ into a regular part $F_\delta'$ and a small remainder $F' - F_\delta'$. Since $F'$ and $F_\delta'$ are both odd, by the definition of the Fourier transform in (1.6), we have

$$(5.4) \qquad \langle \tau_\varepsilon^{(1)}, \varphi^\lambda \rangle = i\langle \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\Phi_\varepsilon) \rangle_\theta + i\langle \widehat{F'} - \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\Phi_\varepsilon) \rangle_\theta,$$

where we have used Fubini to change the order of integration on the right-hand side, and the notation $\langle \cdot, \cdot \rangle_\theta$ refers to integration in the $\theta$ variable. We deal with the two terms in (5.4) separately. For the first one, by Proposition 4.10, we have

$$
\begin{aligned}
&\left| \langle \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\Phi_\varepsilon) \rangle_\theta \right| \\
(5.5) \qquad &\lesssim \sum_{K \in \mathbf{Z}} \|\widehat{F_\delta'}\|_{M+2, \mathfrak{R}_K} \cdot \sup_{r \le M+2} \sup_{\theta \in \mathfrak{R}_K} \left| (\mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\Phi_\varepsilon))^{(r)}(\theta) \right|.
\end{aligned}
$$

Taking the $2n$th moment on both sides and using Lemma 4.3 to interchange the supremum and expectation for each term on the right-hand side, we get

$$
\begin{aligned}
&\left\| \langle \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\Phi_\varepsilon) \rangle_\theta \right\|_{2n} \\
&\lesssim \sum_{K \in \mathbf{Z}} \|\widehat{F_\delta'}\|_{M+2, \mathfrak{R}_K} \cdot \sup_{r \le M+3} \sup_{\theta \in \mathfrak{R}_K} \left\| (\mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\partial_\theta^r \Phi_\varepsilon))(\theta) \right\|_{2n},
\end{aligned}
$$

where we have used the notation $\|\cdot\|_{2n} = (\mathbf{E}|\cdot|^{2n})^{\frac{1}{2n}}$. The maximum in the number of derivatives is now taken over the range $r \le M + 3$ since one pays one more derivative from Lemma 4.3. The term inside the supremum above, when raised to the $2n$th power, has the expression

$$
\begin{aligned}
&\mathbf{E} \left| (\mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\partial_\theta^r \Phi_\varepsilon))(\theta) \right|^{2n} \\
(5.6) \qquad &= (4a^2\varepsilon)^{-n} \int \varphi^\lambda(\mathbf{x}) \left[ \mathbf{E} \prod_{k=1}^{2n} \mathcal{T}_{(1)}(\partial_\theta^r \Phi_\varepsilon(\theta, x_k)) \right] d\mathbf{x}.
\end{aligned}
$$

Applying Theorem 6.2 to the object in the bracket above, we get the bound

$$
\begin{aligned}
&\mathbf{E} \left| (\mathcal{A}_{\varepsilon,\lambda}^\circ \mathcal{T}_{(1)}(\partial_\theta^r \Phi_\varepsilon))(\theta) \right|^{2n} \\
(5.7) \qquad &\lesssim \varepsilon^{-n}(1 + |\theta|)^N \int |\varphi^\lambda(\mathbf{x})| \left( \mathbf{E} \prod_{k=1}^{2n} P_N(X_k) \right) d\mathbf{x}
\end{aligned}
$$

for some $N \geq 1$, where $P_N(X_k) = \sum_{j=1}^{N} X_k^{\diamond(2j+1)}$. Since $\varepsilon^{-\frac{1}{2}} X_k^{\diamond(2j+1)} = \varepsilon^j \times \Psi_\varepsilon^{\diamond(2j+1)}(x_k)$ and the sum in $P_N$ starts from the third-order term, we see that the right-hand side is a linear combination of the $2n$-moments of the quantities $\langle \varepsilon^j \Psi_\varepsilon^{\diamond(2j+1)}, |\varphi^\lambda| \rangle$ with $j = 1, \dots, N$. These are higher order version of $\langle \Psi_\varepsilon, |\varphi^\lambda| \rangle$ with additional Wick powers in $\Psi_\varepsilon$ and $\varepsilon$ balanced, so we have the bound

$$\|(\mathcal{A}_{\varepsilon,\lambda}^{\circ} \mathcal{T}_{(1)}(\partial_\theta^r \Phi_\varepsilon))(\theta)\|_{2n} \lesssim \varepsilon^{\kappa'} \lambda^{-\frac{1}{2}-\kappa'} (1+|\theta|)^N$$

for some $N$ and all $\kappa' < \kappa$. Note that here we have relaxed the upper bound by replacing the exponent $\frac{N}{2n}$ with $N$, but this would not affect our result. Plugging the above bound into (5.5) and applying (4.8) to $\|\widehat{F_\delta'}\|_{M+2;\mathfrak{R}_K}$ with $N$ replaced by $2N$, we get the bound

$$(5.8) \qquad \|\langle \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^{\circ} \mathcal{T}_{(1)}(\Phi_\varepsilon) \rangle_\theta \|_{2n} \lesssim \varepsilon^{\kappa'} \delta^{-2N} \lambda^{-\frac{1}{2}-\kappa'}.$$

We now turn to the second term in (5.4). Write

$$(\mathcal{A}_{\varepsilon,\lambda}^{\circ} \mathcal{T}_{(1)}(\Phi_\varepsilon))(\theta)$$

$$(5.9) \qquad = (\mathcal{A}_{\varepsilon,\lambda}^{\circ} \Phi_\varepsilon)(\theta) - \frac{\theta}{2a} e^{-\frac{\theta^2 \sigma^2}{2}} \int \Psi_\varepsilon(x) \varphi^\lambda(x)\, dx,$$

where $\sigma^2 = \mathbf{E} X^2$, and we estimate the action of $\widehat{F'} - \widehat{F_\delta'}$ on these two terms separately. For the first one, similar as before, we have

$$\|\langle \widehat{F'} - \widehat{F_\delta'}, \mathcal{A}_{\varepsilon,\lambda}^{\circ} \Phi_\varepsilon \rangle_\theta \|_{2n}$$

$$(5.10) \qquad \lesssim \sum_{K \in \mathbf{Z}} \|\widehat{F'} - \widehat{F_\delta'}\|_{M+2,\mathfrak{R}_K} \sup_{r \leq M+3} \sup_{\theta \in \mathfrak{R}_K} \|(\mathcal{A}_{\varepsilon,\lambda}^{\circ} \Phi_\varepsilon)^{(r)}(\theta)\|_{2n}.$$

The quantity $\| \cdot \|_{2n}$ inside the supremum above (raised to the $2n$th power) has the expression

$$\mathbf{E} |(\mathcal{A}_{\varepsilon,\lambda}^{\circ} \Phi_\varepsilon)^{(r)}(\theta)|^{2n} = (4a^2\varepsilon)^{-n} \int \boldsymbol{\varphi}^\lambda(\mathbf{x}) \left[ \mathbf{E} \prod_{k=1}^{2n} \partial_\theta^r \Phi_\varepsilon(\theta, x_k) \right] d\mathbf{x}.$$

The difference between here and (5.6) is that the first chaos component of $\partial_\theta^r \Phi_\varepsilon$ is not removed. We can then apply the bound (6.7) to the expression inside the bracket above with $\tilde{\mathcal{T}} = \tilde{\mathcal{O}} = \{1, \dots, 2n\}$ so that

$$\left| \mathbf{E} \prod_{k=1}^{2n} \partial_\theta^r \Phi_\varepsilon(\theta, x_k) \right| \lesssim (1+|\theta|)^{2n} \left( \mathbf{E} \prod_{k=1}^{2n} X_k \right).$$

Here, the power $2n$ on $(1+|\theta|)$ is precisely $|\tilde{\mathcal{O}}|$ in (6.6). This immediately gives

$$\mathbf{E} |(\mathcal{A}_{\varepsilon,\lambda}^{\circ} \Phi_\varepsilon)^{(r)}(\theta)|^{2n} \lesssim \varepsilon^{-n} (1+|\theta|)^{2n} \int |\boldsymbol{\varphi}^\lambda(\mathbf{x})| \left( \mathbf{E} \prod_{k=1}^{2n} X_k \right) d\mathbf{x}.$$

Note that the right-hand side (without $\theta$) is precisely the $2n$th moment of $\int \Psi_\varepsilon(x)|\varphi^\lambda(x)|\,dx$, so we have the bound

$$\big\|\big(\mathcal{A}^\circ_{\varepsilon,\lambda}\Phi_\varepsilon\big)^{(r)}(\theta)\big\|_{2n} \lesssim (1+|\theta|)\big\|\langle\Psi_\varepsilon,|\varphi^\lambda|\rangle\big\|_{2n} \lesssim (1+|\theta|)\lambda^{-\frac{1}{2}}.$$

By (4.9), we have

$$\big\|\widehat{F'} - \widehat{F'_\delta}\big\|_{M+2,\mathfrak{R}_K} \lesssim \delta^\beta(1+|K|)^{-6+\beta}$$

for every $\beta \in (0,1)$. Plugging this, together with the above bound for $\|(\mathcal{A}^\circ_{\varepsilon,\lambda}\Phi_\varepsilon)^{(r)}(\theta)\|_{2n}$, back into (5.10), we obtain

$$\big\|\langle\widehat{F'} - \widehat{F'_\delta}, \mathcal{A}^\circ_{\varepsilon,\lambda}\Phi_\varepsilon\rangle_\theta\big\|_{2n} \lesssim \delta^\beta\lambda^{-\frac{1}{2}}$$

for some $\beta \in (0,1)$. The same bound holds for the second term in (5.9) but the procedure is simpler. Thus, the remainder part of $\tau_\varepsilon^{(1)}$ satisfies the bound

$$(5.11) \qquad \big\|\langle\widehat{F'} - \widehat{F'_\delta}, \mathcal{A}^\circ_{\varepsilon,\lambda}\mathcal{T}_{(1)}(\Phi_\varepsilon)\rangle_\theta\big\|_{2n} \lesssim \delta^\beta\lambda^{-\frac{1}{2}}.$$

Now, choosing $\delta = \varepsilon^{\frac{\kappa'}{4N}}$, applying the two bounds (5.8) and (5.11) back to (5.4), and recalling that $\kappa'$ can be arbitrarily small, we have thus proved (5.2) for $\tau = \circ$.

5.2.2. *The case $\tau = \overset{\bullet}{\underset{\circ}{\bullet}}\circ$.* We now turn to the case $\tau = \overset{\bullet}{\underset{\circ}{\bullet}}\circ$. Again, we write $\tau_\varepsilon = \widehat{\Pi}^\varepsilon_0\tau$. By (3.2), (3.3) and (5.1), we have

$$\tau_\varepsilon(x) = \frac{1}{4a^3\varepsilon^2}\int \big(K(x-y) - K(-y)\big)K(y-z)F'(X)F'(Y)\{F(Z)\}\,dy\,dz$$

$$- \frac{1}{4a^3\varepsilon^2}\int K(x-y)K(y-z)\mathbf{E}\big(F'(X)F'(Y)\{F(Z)\}\big)\,dy\,dz.$$

For simplicity, we write $K(x,y) = K(x-y) - K(-y)$. Following the notation in Section 6, we let $\mathcal{T} = \{\mathfrak{x}, \mathfrak{y}, \mathfrak{z}\}$ be the type space with categories $\mathcal{O} = \{\mathfrak{x}, \mathfrak{y}\}$ and $\mathcal{E} = \{\mathfrak{z}\}$. These reflect the roles of the fields $X$, $Y$ and $Z$ appearing in $\tau_\varepsilon$.

We write $\boldsymbol{\theta} = (\theta_{\mathfrak{x}}, \theta_{\mathfrak{y}}, \theta_{\mathfrak{z}}) \in \mathbf{R}^3$. For each multi-index $\mathbf{m} = (m_{\mathfrak{x}}, m_{\mathfrak{y}}, m_{\mathfrak{z}}) \in \mathbf{N}^\mathcal{T}$, let $C_\mathbf{m}$ denote the coefficient of the term $X^{\diamond m_{\mathfrak{x}}} \diamond Y^{\diamond m_{\mathfrak{y}}} \diamond Z^{\diamond m_{\mathfrak{z}}}$ in the chaos expansion of $\mathcal{H}(X, Y, Z) \overset{\mathrm{def}}{=} F'(X)F'(Y)\{F(Z)\}$. Note that these coefficients do depend on $(x, y, z)$ and are given by the formula

$$(5.12) \qquad C_\mathbf{m}(x, y, z) = \frac{1}{\mathbf{m}!}\mathbf{E}(\partial^\mathbf{m}\mathcal{H})(X, Y, Z).$$

Finally, let $\mathcal{M} \subset \mathbf{N}^\mathcal{T}$ be

$$\mathcal{M} = \big\{(0, 0, 0), (1, 0, 1), (0, 1, 1), (0, 0, 2), (1, 1, 2)\big\},$$

and let $\mathscr{T}_{\mathcal{M}}(\mathcal{H}(X, Y, Z))$ be the chaos expansion of $\mathcal{H}(X, Y, Z)$ with the components in $\mathcal{M}^6$ removed. With this notation, we now decompose $\tau_\varepsilon$ into

$$\tau_\varepsilon(x) = \sum_{j=1}^{6} \tau_\varepsilon^{(j)}(x),$$

where the six terms are given by

$$\tau_\varepsilon^{(1)}(x) = \frac{1}{4a^3\varepsilon^2} \int K(x, y)K(y-z)\mathscr{T}_{\mathcal{M}}\big(F'(X)F'(Y)\{F(Z)\}\big)\,dy\,dz,$$

$$\tau_\varepsilon^{(2)}(x) = \frac{1}{4a^3\varepsilon^2} \int K(x, y)K(y-z) \cdot C_{1,0,1}(x, y, z)X \diamond Z\,dy\,dz,$$

$$\tau_\varepsilon^{(3)}(x) = \frac{1}{4a^3\varepsilon^2} \int K(x, y)K(y-z) \cdot C_{0,1,1}(x, y, z)Y \diamond Z\,dy\,dz,$$

$$\tau_\varepsilon^{(4)}(x) = \frac{1}{4a^3\varepsilon^2} \int K(x, y)K(y-z) \cdot C_{0,0,2}(x, y, z)Z^{\diamond 2}\,dy\,dz,$$

$$\tau_\varepsilon^{(5)}(x) = \frac{1}{4a^3\varepsilon^2} \int K(x, y)K(y-z) \cdot C_{1,1,2}(x, y, z)X \diamond Y \diamond Z^{\diamond 2}\,dy\,dz,$$

$$\tau_\varepsilon^{(6)}(x) = -\frac{1}{4a^3\varepsilon^2} \int K(-y)K(y-z) \cdot C_{0,0,0}(x, y, z)\,dy\,dz,$$

with the constants $C_{\mathbf{m}}$ given by (5.12).

In what follows, we will prove the convergence of each $\tau_\varepsilon^{(j)}$ to their corresponding limit, which altogether give the limiting KPZ model for $\overset{\bullet}{\underset{\circ\circ}{}}$. We start with $\tau_\varepsilon^{(1)}$. For $\boldsymbol{\theta} = (\theta_{\mathfrak{x}}, \theta_{\mathfrak{y}}, \theta_{\mathfrak{z}})$, let

$$\Phi_\varepsilon(\boldsymbol{\theta}, x, y, z) = \sin(\theta_{\mathfrak{x}}X)\sin(\theta_{\mathfrak{y}}Y)\{\cos(\theta_{\mathfrak{z}}Z)\}.$$

We also define the operator $\mathcal{A}_{\varepsilon,\lambda}^\tau$ for $\tau = \overset{\bullet}{\underset{\circ\circ}{}}$ by

$$(\mathcal{A}_{\varepsilon,\lambda}^\tau \mathcal{F})(\boldsymbol{\theta}) = \int \varphi^\lambda(x)K(x, y)K(y-z)\mathcal{F}(\boldsymbol{\theta}, x, y, z)\,dz\,dy\,dx.$$

Here and below, we omit the symbol $\tau$ in $\mathcal{A}$ for simplicity. With this notation, we have the expression

(5.13)　　　$\langle \tau_\varepsilon^{(1)}, \varphi^\lambda \rangle = \langle \Upsilon_\delta, \mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon) \rangle_{\boldsymbol{\theta}} + \langle \Upsilon - \Upsilon_\delta, \mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon) \rangle_{\boldsymbol{\theta}},$

where $\Upsilon = \widehat{F'} \otimes \widehat{F'} \otimes \widehat{F}$, $\Upsilon_\delta$ is the distribution that replaces every appearance of $F^{(\ell)}$ in $\Upsilon$ by $F_\delta^{(\ell)}$, and $\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon)$ is the chaos expansion of $\Phi_\varepsilon$ in terms of $X$, $Y$ and $Z$ with components in $\mathcal{M}$ removed. For the first term in (5.13), similar as before,

---

[6]Components in $\mathcal{M}$ are terms of the form $X^{m_{\mathfrak{x}}} \diamond Y^{m_{\mathfrak{y}}} \diamond Z^{m_{\mathfrak{z}}}$ with $\mathbf{m} = (m_{\mathfrak{x}}, m_{\mathfrak{y}}, m_{\mathfrak{z}}) \in \mathcal{M}$.

we have

$$\big\|\langle \Upsilon_\delta, \mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon)\rangle_{\boldsymbol{\theta}}\big\|_{2n}$$

$$(5.14) \qquad \lesssim \sum_{\mathbf{K}\in\mathbf{Z}^3} \|\Upsilon_\delta\|_{M+2,\mathfrak{R}_{\mathbf{K}}} \sup_{\mathbf{r}:|\mathbf{r}|_\infty\le M+3} \sup_{\boldsymbol{\theta}\in\mathfrak{R}_{\mathbf{K}}} \big\|(\mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon))(\boldsymbol{\theta})\big\|_{2n}.$$

We first control $\|(\mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon))(\boldsymbol{\theta})\|_{2n}$. Raising it to the $2n$th power, we have

$$\mathbf{E}\big|(\mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon))(\boldsymbol{\theta})\big|^{2n} = \int \boldsymbol{\varphi}^\lambda(\mathbf{x}) \prod_{k=1}^{2n}(K(x_k,y_k)\cdot K(y_k-z_k))$$

$$\times \left[\mathbf{E}\prod_{k=1}^{2n}\mathscr{T}_{\mathcal{M}}\big(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon(\theta,x_k,y_k,z_k)\big)\right] d\mathbf{z}\, d\mathbf{y}\, d\mathbf{x}.$$

It is straightforward to check that the set $\mathcal{M}$ satisfies $\mathcal{B}(\mathbf{n})\cap\mathcal{M} = $ for all $\mathbf{n}\in\mathcal{M}^c$, so we apply Theorem 6.2 to the expectation above to get

$$\mathbf{E}\big|(\mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon))(\boldsymbol{\theta})\big|^{2n}$$

$$\lesssim (1+\theta)^N \int |\boldsymbol{\varphi}^\lambda(\mathbf{x})| \prod_{k=1}^{2n}\big(|K(x_k,y_k)K(y_k-z_k)|\big)$$

$$\times \left[\mathbf{E}\prod_{k=1}^{2n}\mathscr{T}_{\mathcal{M}}\big(P_N(X_k)P_N(Y_k)Q_N(Z_k)\big)\right] d\mathbf{z}\, d\mathbf{y}\, d\mathbf{x},$$

which holds for some $N\ge 2$, and $P_N(X) = \sum_{j=1}^N X^{\diamond(2j-1)}$, $Q_N(Z) = \sum_{j=1}^N Z^{\diamond(2j)}$. It is also straightforward to check that the right-hand side above corresponds to a higher order version of the object
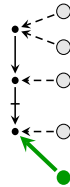


in the sense that every additional two Wick powers of any of these variables are accompanied by one power of $\varepsilon$. It is also strict so that there is at least one additional power of $\varepsilon$. This then implies deduce the bound

$$\big\|(\mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon))(\boldsymbol{\theta})\big\|_{2n} \lesssim (1+\theta)^N \cdot \varepsilon^{\kappa'}\lambda^{-\kappa'}$$

for all $\kappa' < \kappa$. Plugging this bound into (5.14) and employing (4.8) to control $\|\Upsilon_\delta\|_{M+2,\mathfrak{R}_{\mathbf{K}}}$, we then obtain

$$(5.15) \qquad \big\|\langle \Upsilon_\delta, \mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon)\rangle_{\boldsymbol{\theta}}\big\|_{2n} \lesssim \varepsilon^{\kappa'}\delta^{-2N}\lambda^{-\kappa'}.$$

As for the second term in (5.13), we write

$$\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon) = \Phi_\varepsilon - \sum_{\mathbf{m}\in\mathcal{M}} C_{\mathbf{m}} \cdot X^{\diamond m_{\mathfrak{x}}} \diamond Y^{\diamond m_{\mathfrak{y}}} \diamond Z^{\diamond m_{\mathfrak{z}}}, \qquad \mathbf{m} = (m_{\mathfrak{x}}, m_{\mathfrak{y}}, m_{\mathfrak{z}}),$$

and estimate each term above separately. For the action of $\mathcal{A}_{\varepsilon,\lambda}$ on $\Phi_\varepsilon$, we have

$$\mathbf{E}\big|(\mathcal{A}_{\varepsilon,\lambda}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon)(\boldsymbol{\theta})\big|^{2n} = \int \varphi^\lambda(\mathbf{x}) \prod_{k=1}^{2n}\big(K(x_k,y_k)K(y_k-z_k)\big)$$

$$\times \left[\mathbf{E}\prod_{k=1}^{2n}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon)(\boldsymbol{\theta},x_k,y_k,z_k)\right] d\mathbf{z}\, d\mathbf{y}\, d\mathbf{x}.$$

Now, let $\tilde{\mathcal{T}} = \mathcal{T} \times \{1,\ldots,2n\}$ where the types of the elements in $\tilde{\mathcal{T}}$ are the same as their projections onto $\mathcal{T}$. We then apply (6.6) to the right-hand side above to get

$$\mathbf{E}\big|(\mathcal{A}_{\varepsilon,\lambda}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon)(\boldsymbol{\theta})\big|^{2n} \lesssim \theta^{8n} \int |\varphi^\lambda(\mathbf{x})| \prod_{k=1}^{2n}|K(x_k,y_k)K(y_k-z_k)|$$

(5.16)
$$\times \left[\mathbf{E}\prod_{k=1}^{2n}(X_k Y_k Z_k^{\diamond 2})\right] d\mathbf{z}\, d\mathbf{y}\, d\mathbf{x},$$

where $\theta = 1 + |\boldsymbol{\theta}|$. Here, the power of $\theta$ is $8n$ since there are $4n$ points in $\tilde{\mathcal{O}}$ (each contributing 1 power) and $2n$ points in $\tilde{\mathcal{E}}$ (each contributing 2). Note that the right-hand side (without $\theta$) is precisely the $2n$th moment of the quantity $\langle \Pi_0^{\text{KPZ};\varepsilon}, \varphi_0^\lambda \rangle$ without logarithmic renormalisation (i.e., one sets $C_2^{(\varepsilon)} = 0$ in the model), and hence it is bounded by $|\log\varepsilon|^n \lambda^{-2n\kappa'}$. Thus, we deduce the bound

$$\big\|(\mathcal{A}_{\varepsilon,\lambda}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi_\varepsilon)(\boldsymbol{\theta})\big\|_{2n} \lesssim (1+|\boldsymbol{\theta}|)^4 |\log\varepsilon|\lambda^{-\kappa'}.$$

A similar bound, both in terms of the powers of $\boldsymbol{\theta}$ and of the powers of $\varepsilon$ and $\lambda$, holds for the action of $\mathcal{A}_{\varepsilon,\lambda}$ on $C_{\mathbf{m}}(\boldsymbol{\theta},x,y,z)X^{\diamond m_{\mathfrak{x}}}\diamond Y^{\diamond m_{\mathfrak{y}}}\diamond Z^{\diamond m_{\mathfrak{z}}}$. Since (4.9) implies

$$\|\Upsilon - \Upsilon_\delta\|_{M+2;\mathfrak{R}_{\mathbf{K}}} \lesssim \delta^\beta (1+|K_{\mathfrak{x}}|)(1+|K_{\mathfrak{y}}|)(1+|\mathbf{K}|)^{-7-\alpha+\beta},$$

we then have

(5.17)
$$\big\|\langle \Upsilon - \Upsilon_\delta, \mathcal{A}_{\varepsilon,\lambda}\mathscr{T}_{\mathcal{M}}(\Phi_\varepsilon)\rangle_{\boldsymbol{\theta}}\big\|_{2n} \lesssim \delta^\beta|\log\varepsilon|\lambda^{-\kappa'}.$$

Now, combining (5.15) and (5.17), taking $\delta = \varepsilon^{\frac{\kappa'}{4N}}$, and noting that both $\kappa'$ and $\beta \in (0,1)$ can be arbitrarily small, we can then deduce that there exists $\zeta > 0$ such that

$$\big\|\langle 1\tau_\varepsilon^{(1)}, \varphi^\lambda\rangle\big\|_{2n} \lesssim \varepsilon^\zeta \lambda^{|\tau|+\zeta}.$$

We now turn to $\tau_\varepsilon^{(2)}$. Let

$$C_{1,0,1}^{(1)} = \mathbf{E}\big(\{F''(X)\{F'(Y)F'(Z)\big), \qquad C_{1,0,1}^{(2)} = 2a\mathbf{E}\big(F'(Y)F'(Z) - 4a^2 YZ\big),$$

and $C_{1,0,1}^{(3)} = 8a^3 \mathbf{E} YZ$. We then write $\tau_\varepsilon^{(2)} = \sum_{j=1}^3 \tau_\varepsilon^{(2;j)}$, where

$$\tau_\varepsilon^{(2;j)}(x) = \frac{1}{4a^3\varepsilon} \int K(x,y)K(y-z) \cdot C_{1,0,1}^{(j)}(x,y,z)\Psi_\varepsilon(x) \diamond \Psi_\varepsilon(z) \, dy \, dz.$$

There is only $\varepsilon^{-1}$ left since the other negative power is combined with $X \diamond Z$ so that we have $\Psi_\varepsilon(x) \diamond \Psi_\varepsilon(z)$ in the expression. We analyse the three terms separately. For $\tau_\varepsilon^{(2;3)}$, we have



(5.18)

where we have identified



This is because the correlation function $\mathbf{E}(\Psi_\varepsilon(x)\Psi_\varepsilon(y))$, represented by the two dashed arrows above, is symmetric under the reflection of its space variable with respect to the origin. Since the kernel $K$ is anti-symmetric in the space variable, the whole kernel on the left-hand side above is anti-symmetric, and hence we can identify it with its renormalised version. The right-hand side of (5.18) is precisely one of the terms in $\langle \Pi_0^{\mathrm{KPZ};\varepsilon} \, \mathord{\otimes}, \varphi_0^\lambda \rangle$ in the Appendix, so it does converge to the desired limit with the error bound (5.2). For $\langle \tau_\varepsilon^{(2;1)}, \varphi_0^\lambda \rangle$, according to the bound (4.13), it can be decomposed into two parts. The first one, corresponding to the first term on the right-hand side of (4.13), can be represented by



Here, the highlights on the edges mean that they are upper bounds for the corresponding parts in the object itself. By the bounds in [25], Section 6.2.6, the $L^{2n}$th

moment of the above graph is controlled by $\varepsilon^{\kappa'}\delta^{-N}\lambda^{-2\kappa'}$. The second one, corresponding to the second term on the right-hand side of (4.13), carries a logarithmic divergence but multiplied by $\delta^\beta$. Altogether, we have the bound

$$\|\langle \tau_\varepsilon^{(2;1)}, \varphi_0^\lambda \rangle\|_{2n} \lesssim \left(\varepsilon^{\kappa'}\delta^{-N} + \delta^\beta |\log \varepsilon|\right)\lambda^{-2\kappa'}.$$

We then choose $\delta = \varepsilon^{\frac{\kappa'}{4N}}$ so that $\|\langle \tau_\varepsilon^{(2;1)}, \varphi_0^\lambda \rangle\|_{2n} \lesssim \varepsilon^{\kappa'}\lambda^{-2\kappa'}$. Finally, for $\tau_\varepsilon^{(2;2)}$, we note that the kernel $K(y - z)\mathbf{E}(F'(Y)F'(Z) - 4a^2YZ)$ is anti-symmetric in the space variable, and hence can be identified with its renormalised kernel. Hence, it can be controlled by the graph
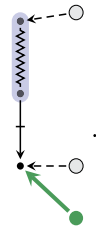


Here, the highlight on the renormalised kernel means that we have the bound $\|\mathscr{R}Q_\varepsilon\|_{3+\beta'} \lesssim \varepsilon^{\beta'}$ for some $\beta' > 0$. This is the same $\beta'$ as in (4.14). Hence, we have $\|\langle \tau_\varepsilon^{(2;2)}, \varphi_0^\lambda \rangle\|_{2n} \lesssim \varepsilon^{\kappa'}\lambda^{-2\kappa'}$ as well. This completes the proof for $\tau_\varepsilon^{(2)}$.

The situations for $\tau_\varepsilon^{(3)}$ and $\tau_\varepsilon^{(4)}$ are similar. One can show that their main parts are given by
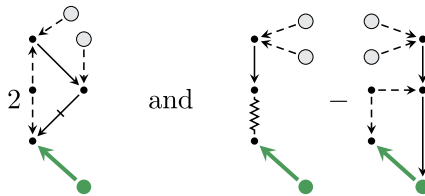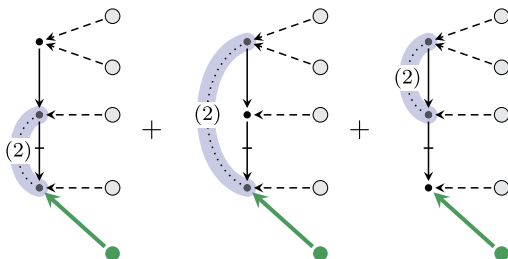


respectively, and their other parts vanish with the correct order just as in the case for $\tau_\varepsilon^{(2;1)}$ and $\tau_\varepsilon^{(2;2)}$. This implies that both $\tau_\varepsilon^{(3)}$ and $\tau_\varepsilon^{(4)}$ converge to the right limit.

For $\tau_\varepsilon^{(5)}$, its main part (when tested against $\varphi_0^\lambda$) is given by



(5.19)

which does converge to the right limit. By Lemma 4.12, we see that the remainder can be decomposed into two parts. The first one can be controlled by the graph



multiplied by $\delta^{-N}$. The other one from the remainder is the same as (5.19) but accompanied by a positive power of $\delta$. Hence, the $\|\cdot\|_{2n}$ norm of the remainder (after testing against $\varphi_0^\lambda$) is bounded by $\varepsilon^{\kappa'}(\delta^{-N}+\delta^\beta)\lambda^{-2\kappa'}$. One can again choose $\delta$ to be a small positive power of $\varepsilon$ so that the it vanishes in the correct order.

Finally for $\tau_\varepsilon^{(6)}$, according to Lemma 4.13, we write

$$C_{0,0,0} = H^{(1)} + H^{(2)} + 4a^3\mathbf{E}\big(XYZ^{\diamond 2}\big).$$

The part with $4a^3\mathbf{E}(XYZ^{\diamond 2})$ is precisely
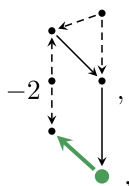


while the other two parts can be shown (with a similar argument as for $\tau_\varepsilon^{(2)}$) to vanish with the correct order in $\lambda$. This completes the proof for $\tau = $ .

5.2.3. *The case* $\tau = $ . We now turn to the last symbol $\tau = $ . By (3.2), (3.3) and (5.1), we have

$$\tau_\varepsilon(x) = \bar\tau_\varepsilon(x) - \frac{C_{\textbf{v}}^{(\varepsilon)}}{2a}\cdot \{\!\!\{F''(X)\}\!\!\},$$

where

$$\bar\tau_\varepsilon(x) = \frac{1}{2a^3\varepsilon^2}\int K(x-y)K(x-z)\{\!\!\{F''(X)\}\!\!\}\{\!\!\{F(Y)\}\!\!\}\{\!\!\{F(Z)\}\!\!\}\,dy\,dz.$$

We deal with the second term first. Since $\mathbf{E}\{\!\!\{F(Y)\}\!\!\}\{\!\!\{F(Z)\}\!\!\} \lesssim (\mathbf{E}YZ)^2$, it is not hard to see that $C_{\textbf{v}}^{(\varepsilon)} = \mathcal{O}(|\log\varepsilon|)$. The same procedure as above also implies $\|\langle\{\!\!\{F''(X)\}\!\!\}, \varphi^\lambda\rangle\|_{2n} \lesssim \varepsilon^{\kappa'}\lambda^{-\kappa'}$. Thus, the term $\frac{1}{2}C_{\textbf{v}}^{(\varepsilon)}\{\!\!\{F''(X)\}\!\!\}$ vanishes at the correct order. It then suffices to show that $\bar\tau_\varepsilon$ vanishes and satisfies the bound (5.2).

If we brutally implement the above procedure with the general bound (6.6), then because of the two derivatives on $F$ (with $X$), we will end up with requiring

9+ differentiability of $F$ in order for $\bar{\tau}_\varepsilon$ to satisfy the desired bound.[7] However, a more careful observation reveals that, up to the subtraction of $C_{\bullet\bullet}^{(\varepsilon)}$, $\bar{\tau}_\varepsilon$ is the product of three almost bounded processes: $\widehat{\Pi}^\varepsilon\square$ and $(\widehat{\Pi}^\varepsilon \curlywedge)^2$. By writing $F'' = F_\delta'' + (F'' - F_\delta'')$, we can ignore the bad effect of two derivatives on $F_\delta$. As for the small remainder, we can separate the product of the three "almost bounded" processes into subproducts with less terms. This will reduce the requirement on the regularity of $F$.

More precisely, we write $\bar{\tau}_\varepsilon(x) = \bar{\tau}_{\varepsilon,\delta}^{(1)}(x) + \bar{\tau}_{\varepsilon,\delta}^{(2)}(x)$, where

$$\bar{\tau}_{\varepsilon,\delta}^{(1)}(x) = \frac{1}{2a^3\varepsilon^2} \int K(x-y)K(x-z) \{\!\!\{ F_\delta''(X) \}\!\!\} \{\!\!\{ F(Y) \}\!\!\} \{\!\!\{ F(Z) \}\!\!\} \, dy\, dz,$$

$$\bar{\tau}_{\varepsilon,\delta}^{(2)}(x) = \frac{1}{2a^3\varepsilon^2} \int K(x-y)K(x-z) \{\!\!\{ F''(X) - F_\delta''(X) \}\!\!\} \{\!\!\{ F(Y) \}\!\!\} \{\!\!\{ F(Z) \}\!\!\} \, dy\, dz.$$

For $\bar{\tau}_{\varepsilon,\delta}^{(1)}$, similar as before, we get a power $(1 + |\mathbf{K}|)^6$ from the norm of "test functions" localised in $\mathfrak{R}_\mathbf{K}$. As for the norm of distributions, by Assumption 1.1 and Lemma 4.5, we have

$$\| \widehat{F_\delta''} \otimes \widehat{F} \otimes \widehat{F} \|_{M+2,\mathfrak{R}_K} \lesssim_N \delta^{-N} (1 + |K_{\mathfrak{r}}|)^{-N} \big( (1 + |K_{\mathfrak{y}}|)(1 + |K_{\mathfrak{z}}|) \big)^{-7-\alpha}$$

for all large $N$. By choosing $\delta = \varepsilon^{\frac{\kappa'}{4N}}$ as before, we obtain the bound

(5.20) $$\| \langle \bar{\tau}_{\varepsilon,\delta}^{(1)}, \varphi^\lambda \rangle \|_{2n} \lesssim \varepsilon^{\frac{\kappa'}{2}} \lambda^{-\kappa'}.$$

As for $\tau_{\varepsilon,\delta}^{(2)}$, we have the expression

$$\langle \tau_{\varepsilon,\delta}^{(2)}, \varphi^\lambda \rangle = \int \{\!\!\{ F''(X) - F_\delta''(X) \}\!\!\} \cdot \left( \varepsilon^{-1} \int K(x-y) \{\!\!\{ F(Y) \}\!\!\} \, dy \right)^2 \varphi^\lambda(x) \, dx.$$

We first separate the two terms in the integrand so that

(5.21) $$|\langle \tau_{\varepsilon,\delta}^{(2)}, \varphi^\lambda \rangle| \lesssim \sup_x |\{\!\!\{ F''(X) - F_\delta''(X) \}\!\!\}| \cdot \langle (\curlywedge_\varepsilon)^2, |\varphi^\lambda| \rangle,$$

where $\curlywedge_\varepsilon = \widehat{\Pi}^\varepsilon \curlywedge = K * \widehat{\Pi}^\varepsilon \bullet$. Considering the $L^{2n}$ norm of both sides and applying Hölder's inequality, we get

$$\| \langle \tau_{\varepsilon,\delta}^{(2)}, \varphi^\lambda \rangle \|_{2n} \lesssim \left( \mathbf{E} \sup_x |\{\!\!\{ F''(X) - F_\delta''(X) \}\!\!\}|^{4n} \right)^{\frac{1}{4n}} \cdot \| \langle (\curlywedge_\varepsilon)^2, |\varphi^\lambda| \rangle \|_{4n}.$$

The first term is easily seen to be bounded by some positive power of $\varepsilon$ since exchanging the supremum with the expectation costs an arbitrary small power of

---

[7]The 9+ differentiability arises as follows. The product $\{\!\!\{ \cos(\theta_{\mathfrak{r}} X) \}\!\!\} \{\!\!\{ \cos(\theta_{\mathfrak{y}} Y) \}\!\!\} \{\!\!\{ \cos(\theta_{\mathfrak{z}} Z) \}\!\!\}$ gives rise to $\theta^6$, and another two powers of $\theta$ come from $F''$. Thus, if we brutally bound things in this way, we need $\| \widehat{F} \|$ to decay faster than $\theta^{-9-}$.

$\varepsilon$, but we have $\delta = \varepsilon^{\frac{\kappa'}{4N}}$. For the second term, we have $\|\langle(\overset{\bullet}{\underset{\varepsilon}{\shortmid}})^2, |\varphi^\lambda|\rangle\|_{4n} \lesssim |\log \varepsilon| \cdot \lambda^{-\kappa'}$. This logarithmic factor can be killed by the positive power of $\varepsilon$ from the first term, so that we get

$$(5.22) \qquad \|\langle\tau^{(2)}_{\varepsilon,\delta}, \varphi^\lambda\rangle\|_{2n} \lesssim \varepsilon^{\kappa''}\lambda^{-\kappa'}.$$

The bound for $\bar{\tau}_\varepsilon$, and hence the symbol $\overset{\bullet\bullet}{\underset{\sqcup}{}}$ then follows immediately by combining (5.20) and (5.22).

REMARK 5.3. It is clear from the above argument that the $7+$ differentiability of $F$ exactly guarantees the bound (5.20) for $\bar{\tau}^{(1)}_{\varepsilon,\delta}$. The readers may wonder why we only separate $F''_\delta$ from $F''$ but not the other two $F$'s. The reason is that in (5.21) when we apply $L^\infty$ and $L^1$ bounds to separate the integrands, it is essential to have the stochastic part of the latter integrand positive (like $(\overset{\bullet}{\underset{\varepsilon}{\shortmid}})^2$) so that we can keep the structure of the stochastic objects without adding absolute value to them.

5.3. *Behaviour of renormalisation constants.* We now explore behaviour of the renormalisation constants $C^{(\varepsilon)}_\tau$'s defined in (5.1). It is clear that $C^{(\varepsilon)}_\bullet$ diverges at order $\varepsilon^{-1}$. As for the other constants, we will see that both $C^{(\varepsilon)}_{\bullet\bullet}$ and $C^{(\varepsilon)}_{\infty}$ diverge logarithmically. However, the two logarithmic divergences actually cancel each other out, as already shown in [21, 25] for polynomial $F$. Finally, $C^{(\varepsilon)}_{\bullet\bullet}$ is uniformly bounded in $\varepsilon$.

We first show the cancellation of the two logarithmic divergences for general $F$. This follows from that for the polynomial $F$ and Lemma 4.13. We give it in the following proposition.

PROPOSITION 5.4. *We have*

$$\sup_{\varepsilon \in (0,1)} (C^{(\varepsilon)}_{\bullet\bullet} + 4C^{(\varepsilon)}_{\infty}) < +\infty.$$

PROOF. Let

$$\tilde{C}^{(\varepsilon)}_{\bullet\bullet} = \int K(x-y)K(x-z)\mathbf{E}\big(\Psi^{\diamond 2}_\varepsilon(y)\Psi^{\diamond 2}_\varepsilon(z)\big)\,dy\,dz,$$

$$\tilde{C}^{(\varepsilon)}_{\infty} = \int K(x-y)K(y-z)\mathbf{E}\big(\Psi_\varepsilon(x)\Psi_\varepsilon(y)\Psi^{\diamond 2}_\varepsilon(z)\big)\,dy\,dz.$$

We first show that these two quantities contain all the logarithmic divergence for $C^{(\varepsilon)}_{\bullet\bullet}$ and $C^{(\varepsilon)}_{\infty}$ in the sense that both $C^{(\varepsilon)}_{\bullet\bullet} - \tilde{C}^{(\varepsilon)}_{\bullet\bullet}$ and $C^{(\varepsilon)}_{\infty} - \tilde{C}^{(\varepsilon)}_{\infty}$ are uniformly bounded in $\varepsilon$. To see this, note that

$$C^{(\varepsilon)}_{\bullet\bullet} - \tilde{C}^{(\varepsilon)}_{\bullet\bullet} = \frac{1}{a^2\varepsilon^2}\int K(x-y)K(x-z)\mathbf{E}\big(\{F(Y)\}\{F(Z)\} - a^2 Y^{\diamond 2}Z^{\diamond 2}\big)\,dy\,dz.$$

One can directly perform a chaos expansion and show as in the proof of Lemma 4.13 that

$$\left| \mathbf{E}\big(\{F(Y)\}\{F(Z)\} - a^2 Y^{\diamond 2} Z^{\diamond 2}\big) \right| \lesssim (\mathbf{E} Y Z)^4.$$

This implies the bound

$$\left| C_{\bullet\bullet}^{(\varepsilon)} - \widetilde{C}_{\bullet\bullet}^{(\varepsilon)} \right| \lesssim \varepsilon \int \left| K(x-y)K(y-z) \right| \cdot \mathbf{E}\big(\Psi_\varepsilon^{\diamond 4}(y)\Psi_\varepsilon^{\diamond 4}(z)\big)\, dy\, dz.$$

By the bounds for polynomial models in [25] (second part of Theorem 6.5), the right-hand side above converges to a finite limit as $\varepsilon \to 0$, and hence $C_{\bullet\bullet}^{(\varepsilon)} - \widetilde{C}_{\bullet\bullet}^{(\varepsilon)}$ stays finite. As for $C_{\circ\!\!\circ}^{(\varepsilon)}$, we have

$$C_{\circ\!\!\circ}^{(\varepsilon)} - \widetilde{C}_{\circ\!\!\circ}^{(\varepsilon)} = \frac{1}{4a^3\varepsilon^2} \sum_{j=1}^{2} \int K(x-y)K(y-z)H^{(j)}(x,y,z)\, dy\, dz,$$

where $H^{(j)}$ are given by Lemma 4.13. The bound for $H^{(2)}$ implies that the integral with $H^{(2)}$ stays finite as $\varepsilon \to 0$. As for $H^{(1)}$, one can choose $\delta$ to be a small positive power of $\varepsilon$ so that the bound in Lemma 4.13 is sufficient to guarantee that the integral with $H^{(1)}$ vanishes in the limit as $\varepsilon \to 0$. Hence, $C_{\circ\!\!\circ}^{(\varepsilon)} - \widetilde{C}_{\circ\!\!\circ}^{(\varepsilon)}$ is also uniformly bounded in $\varepsilon$.

Since $\widetilde{C}_{\bullet\bullet}^{(\varepsilon)}$ and $\widetilde{C}_{\circ\!\!\circ}^{(\varepsilon)}$ are precisely the logarithmically divergent quantities from the KPZ equation (with $F(u) = u^2$), it then follows from [25], Theorem 6.5, that $\widetilde{C}_{\bullet\bullet}^{(\varepsilon)} + \widetilde{C}_{\circ\!\!\circ}^{(\varepsilon)}$ converges to a finite limit. This completes the proof of the proposition. □

We have the following proposition on the behaviour of $C_{\bullet\!\!\sqcup\!\!\bullet}^{(\varepsilon)}$ as $\varepsilon \to 0$.

PROPOSITION 5.5. *The constant $C_{\bullet\!\!\sqcup\!\!\bullet}^{(\varepsilon)}$ defined in* (5.1) *is uniformly bounded in $\varepsilon \in (0,1)$.*

PROOF. We would like to have the control

$$(5.23) \qquad \mathbf{E}\big(\{F''(X)\}\{F(Y)\}\{F(Z)\}\big) \lesssim \mathbf{E}\big(X^{\diamond 2} Y^{\diamond 2} Z^{\diamond 2}\big).$$

This will immediately reduce the problem to the polynomial case and gives the claim. However, similar as in the case for the symbol $\bullet\!\!\sqcup\!\!\bullet$, if we Fourier expand $F''$ and $F$ and brutally apply Theorem 6.4, then we would need $F \in \mathcal{C}^{9+}$ to get the bound (5.23). Fortunately, there are only three terms in the product, and hence we can do exact computations with the trigonometric identities. This will give us the desired bound (5.23) under Assumption 1.1, and hence the claim follows. □

REMARK 5.6. With some extra effort, it is possible to show that both $C_{\bullet\bullet}^{(\varepsilon)} + 4C_{\circ\!\!\circ}^{(\varepsilon)}$ and $C_{\bullet\!\!\sqcup\!\!\bullet}^{(\varepsilon)}$ converge to a finite limit as $\varepsilon \to 0$. But the convergence does not really matter here, so we omit those details and only claim uniform boundedness.

5.4. *Identification of the limit.* We now have all the ingredients in place to prove the main result of the article.

THEOREM 5.7. *Let* $\gamma \in (\frac{3}{2}, \frac{5}{3})$ *and* $\eta \in (\frac{1}{2} - \frac{1}{M+4}, \frac{1}{2})$. *Let* $\{h_0^{(\varepsilon)}\}$ *be a sequence of functions in* $\mathcal{C}_\varepsilon^{\gamma,\eta}$ *such that* $\|h_0^{(\varepsilon)}; h_0\|_{\gamma,\eta;\varepsilon} \to 0$ *for some* $h_0 \in \mathcal{C}^\eta$. *Then there exists* $C_\varepsilon \to +\infty$ *such that for every* $T > 0$, *the solution* $h_\varepsilon$ *defined in* (1.5) *with initial data* $h_0^{(\varepsilon)}$ *converge in probability in* $\mathcal{C}^\eta([0,T] \times \mathbf{T})$ *to the solution to the KPZ equation with parameter* $a$ *and initial condition* $h_0$.

PROOF. Let $\Pi^{\mathrm{KPZ}}$ be the standard KPZ model described in the Appendix. Let $\widehat{\Pi}^\varepsilon$ be the sequence of models defined in (3.2) and (3.3) with input $\psi = \Psi_\varepsilon$ and the constants $C_j^{(\varepsilon)}$ specified in (5.1). Let $u_0^{(\varepsilon)} = h_0^{(\varepsilon)} - Z_\varepsilon(0, \cdot)$ where $Z_\varepsilon = P * \xi_\varepsilon$. By the convergence of $Z_\varepsilon$ to $Z = P * \xi$ and the convergence of $h_0^{(\varepsilon)}$, we have $\|u_0^{(\varepsilon)}; u_0\|_{\gamma,\eta;\varepsilon} \to 0$ in probability where $u_0 = h_0 - Z(0, \cdot)$.

Now, let $U^{(\varepsilon)} \in \mathcal{D}_\varepsilon^{\gamma,\eta}$ and $U \in \mathcal{D}^{\gamma,\eta}$ denote the solutions to the fixed-point problem with models $\widehat{\Pi}^\varepsilon$ and $\Pi^{\mathrm{KPZ}}$ and initial datum $u_0^{(\varepsilon)}$ and $u_0$, respectively. Let

$$u_\varepsilon := \widehat{\mathcal{R}}^\varepsilon U^{(\varepsilon)}, \qquad u := \mathcal{R}^{\mathrm{KPZ}} U,$$

where $\widehat{\mathcal{R}}^\varepsilon$ and $\mathcal{R}^{\mathrm{KPZ}}$ are the reconstruction maps associated to the corresponding models. By Theorem 3.8, $u_\varepsilon$ solves the remainder equation 2.2 with initial data $u_0^{(\varepsilon)}$ and constant $C_\varepsilon$ as in (3.5), so that $h_\varepsilon = u_\varepsilon + Z_\varepsilon$ solves (1.5) with initial condition $h_0^{(\varepsilon)}$. Similarly, combining its definition with [25], Theorem 1.2 (see also [14], Theorem 15.1), it follows that there exists a constant $c$ such that $h := u + Z - ct$ solves the KPZ equation (in the sense of Hopf–Cole) with parameter $a$ and initial condition $h_0$. Since $Z_\varepsilon \to Z$ in probability in $\mathcal{C}^\eta([0,T] \times \mathbf{T})$, the convergence of $h_\varepsilon$ to $h$ is established if we show that $u_\varepsilon \to u$ in the same space. The additional term $ct$ can of course easily be generated for $h_\varepsilon$ as well by adding $c$ to $C_\varepsilon$.

By Theorem 5.2, we have $\|\widehat{\Pi}^\varepsilon; \Pi^{\mathrm{KPZ}}\|_{\varepsilon,0} \to 0$. Also, one can easily check that

$$\lim_{\varepsilon \to 0} \mathbf{E} \sup_{z \in [0,T] \times \mathbf{T}} \varepsilon^{\frac{1}{2}+\kappa} |\Psi_\varepsilon(z)| = 0.$$

Furthermore, we know from [22], Proposition 7.11, and the fact that the Hopf–Cole solutions to the KPZ equation are global in time almost surely, that we can find $U$ solving (3.8) up to any fixed time $T > 0$. Hence, Theorem 3.7 and continuity of the reconstruction maps imply that $u_\varepsilon \to u$ in probability in $\mathcal{C}^\eta([0,T] \times \mathbf{T})$ as required. $\square$

REMARK 5.8. Note that for any fixed $\varepsilon > 0$, there may be a set $\Omega_T$ of strictly positive probability such that for realisations $\xi_\varepsilon \in \Omega_T$, the solution to (1.5) explodes before time $T$. This is not a problem, convergence in probability should be interpreted as stating that the probability of this event converges to 0 and that the solution conditioned on survival converges in probability.

**6. A general pointwise bound.** In this section, we state and prove two general bounds that control the correlation functions of trigonometric polynomials of Gaussian processes by those of suitable polynomials. These bounds, combined with the convergence of polynomial models in [25], are the main ingredient for the convergence result in Theorem 5.2.

6.1. *The statement.* For every finite set $\mathscr{A}$, let $\mathbf{N}^{\mathscr{A}}$ be the set of multi-indices on $\mathscr{A}$. We define the length of a multi-index by $|\mathbf{n}| = \sum_{\alpha \in \mathscr{A}} n^{\alpha}$. For $\mathscr{B} \subset \mathscr{A}$, we write $\mathbf{N}^{\mathscr{A}}_{\mathscr{B}} \subset \mathbf{N}^{\mathscr{A}}$ for those multi-indices that vanish on $\mathscr{B}$. For every collection of joint Gaussian random variables $\mathbf{X} = (X_{\alpha})_{\alpha \in \mathscr{A}}$ and every $\mathbf{n} \in \mathbf{N}^{\mathscr{A}}$, we write $\mathbf{X}^{\diamond \mathbf{n}} = \diamondsuit_{\alpha} X_{\alpha}^{\diamond n^{\alpha}}$.

For $H \colon \mathbf{R}^{\mathscr{A}} \to \mathbf{R}$ a continuous function with at most exponential growth, $\mathbf{X} = (X_{\alpha})_{\alpha \in \mathscr{A}}$ a collection of jointly Gaussian random variables, and $\mathcal{M} \subset \mathbf{N}^{\mathscr{A}}$ finite, we define

$$(6.1) \qquad \mathscr{T}_{\mathcal{M}}\big(H(\mathbf{X})\big) = H(\mathbf{X}) - \sum_{\mathbf{n} \in \mathcal{M}} C_{\mathbf{n}} \mathbf{X}^{\diamond \mathbf{n}},$$

where $C_{\mathbf{n}} = \mathbf{E}(\partial^{\mathbf{n}} H)(\mathbf{X})/\mathbf{n}!$ is the coefficient of the term $\mathbf{X}^{\diamond \mathbf{n}}$ in the chaos expansion of $H(\mathbf{X})$.

Let $\mathcal{T}$ be a finite set of "types", which comes equipped with a partition $\mathcal{T} = \mathcal{O} \sqcup \mathcal{E}$, with $\mathcal{O}$ denoting "odd" points and $\mathcal{E}$ standing for "even" points. We now consider $\mathcal{T}$, $\mathcal{O}$ and $\mathcal{E}$ as being fixed. For every $\mathbf{n} \in \mathbf{N}^{\mathcal{T}}$, we define the following sets:

$$(6.2) \qquad \begin{aligned} \mathcal{O}_1(\mathbf{n}) &= \{t \in \mathcal{O} : n^t \text{ odd}\}, \\ \mathcal{O}_2(\mathbf{n}) &= \{t \in \mathcal{O} : n^t \text{ even}\}, \\ \mathcal{E}_1(\mathbf{n}) &= \{t \in \mathcal{E} : n^t \text{ odd}\}, \\ \mathcal{E}_2(\mathbf{n}) &= \{t \in \mathcal{E} : n^t \geq 2 \text{ even}\}, \\ \mathcal{E}_0(\mathbf{n}) &= \{t \in \mathcal{E} : n^t = 0\}. \end{aligned}$$

Note that $\mathcal{O}_2(\mathbf{n})$ includes those $t \in \mathcal{O}$ for which $n^t = 0$, while for $t \in \mathcal{E}$ we separate 0 and strictly positive even indices. We will sometimes omit the argument $\mathbf{n}$ from the sets $\mathcal{O}_j$ and $\mathcal{E}_j$ and we use $|\cdot|$ to denote the cardinality of a set. We also define

$$(6.3) \qquad \mathcal{B}(\mathbf{n}) = \big\{2\mathbf{k} + \mathbf{n} : \mathbf{k} \in \mathbf{N}^{\mathcal{T}}_{\mathcal{E}_0(\mathbf{n})}\big\},$$

and we call this the "branching" of $\mathbf{n}$. Note that in $\mathcal{B}(\mathbf{n})$, we add positive even integers only to the components $t$ which are *not* in $\mathcal{E}_0(\mathbf{n})$.

Consider a $\mathcal{T}$-tuple of space–time points $(x_t)_{t \in \mathcal{T}}$ in $\mathbf{R}^+ \times \mathbf{T}$, and recall the notation $X_t = \varepsilon^{\frac{1}{2}} \Psi_{\varepsilon}(x_t)$. Also, let $\theta_t \in \mathbf{R}$ for every $t$. For convenience, we define $\mathrm{trig}_t = \sin$ if $t \in \mathcal{O}$ and $\mathrm{trig}_t = \cos$ if $t \in \mathcal{E}$. Since $\mathbf{E}\sin(\theta_t X_t) = 0$, we have

$$(6.4) \qquad \Phi(\boldsymbol{\theta}, \mathbf{X}) := \prod_{t \in \mathcal{O}} \sin(\theta_t X_t) \prod_{t \in \mathcal{E}} \wr\cos(\theta_t X_t)\wr = \prod_{t \in \mathcal{T}} \wr\mathrm{trig}_t(\theta_t X_t)\wr,$$

where we have used the shorthand notation $\boldsymbol{\theta} = (\theta_t)_{t \in \mathcal{T}}$ and $\mathbf{X} = (X_t)_{t \in \mathcal{T}}$. For every $\mathbf{n} \in \mathbf{N}^{\mathcal{T}}$, we write $C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X})$ for the coefficient of $\mathbf{X}^{\diamond \mathbf{n}}$ in the chaos expansion of $\Phi(\boldsymbol{\theta}, \mathbf{X})$, defined as in (6.1). In this section, we will be interested in the correlation functions for quantities of the form $\mathscr{T}_{\mathcal{M}}(\Phi(\boldsymbol{\theta}, \mathbf{X}))$. The set $\mathcal{M}$ of removed chaos components will always be chosen such that it satisfies the following assumption.

ASSUMPTION 6.1. $\mathcal{M} \subset \mathbf{N}^{\mathcal{T}}$ is finite and there exists a finite set of "roots" $\mathcal{R} \subset \mathcal{M}^c$ such that $\bigcup_{\mathbf{m} \in \mathcal{R}} \mathcal{B}(\mathbf{m}) = \mathcal{M}^c$.

For every $t \in \mathcal{T}$ and $K \geq 1$, fix $K$ space–time points $\{x_{k,t} : k \in [K]\}$. Here and below, we use $[K]$ to denote the set $\{1, \ldots, K\}$. For every $k \in [K]$, we write $\mathbf{x}_k = (x_{k,t})_{t \in \mathcal{T}}$ and $\mathbf{X}_k = (X_{k,t})_{t \in \mathcal{T}}$, where as before $X_{k,t} = \varepsilon^{\frac{1}{2}} \Psi_\varepsilon(x_{k,t})$. For every $N \geq 1$ and every Gaussian random variable $X$, we also write $P_N$ and $Q_N$ for the random variables

$$P_N(X) = \sum_{k=1}^{N} X^{\diamond(2k-1)}, \qquad Q_N(X) = \sum_{k=1}^{N} X^{\diamond(2k)}.$$

Note that $Q_N$ has no constant term. The main statement is the following.

THEOREM 6.2. *Let $\mathcal{M} \subset \mathbf{N}^{\mathcal{T}}$ satisfy Assumption 6.1. Let $\mathbf{r} \in \mathbf{N}^{\mathcal{T}}$. Then there exists $N$ depending on $K$, $|\mathcal{T}|$ and $\mathcal{M}$ only such that*

$$\left| \mathbf{E} \prod_{k=1}^{K} \mathscr{T}_{\mathcal{M}} \left( \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_k) \right) \right|$$

(6.5)
$$\lesssim \theta^N \mathbf{E} \prod_{k=1}^{K} \mathscr{T}_{\mathcal{M}} \left[ \left( \prod_{t \in \mathcal{O}} P_N(X_{k,t}) \right) \left( \prod_{t \in \mathcal{E}} Q_N(X_{k,t}) \right) \right],$$

*where $\theta = 1 + \max_t |\theta_t|$, and the proportionality constant depends on $K$, $\mathbf{r}$ and $|\mathcal{T}|$ only. In particular, the bound is uniform over all locations of the points $\{x_{k,t}\}$ and all $\varepsilon \in (0, 1)$.*

REMARK 6.3. We do not claim that this is a general statement about arbitrary collections of Gaussian random variables. Instead, this only holds for collections of the type $X_t = \varepsilon^{\frac{1}{2}} \Psi_\varepsilon(x_t)$ as always considered in this article. In particular, we exploit the fact that the $X_t$ are all positively correlated and the fact that their covariances are related in a specific way to the distances between the corresponding space–time locations, which themselves satisfy the triangle inequality. Both of these ingredients are crucial in the proof.

We will mainly use the above bound for arbitrary even $K$ and for $K = 1$. In the case $K = 1$, if $\mathbf{0} \in \mathcal{M}$, then both sides of (6.5) vanish and the bound holds trivially.

If $\mathbf{0} \notin \mathcal{M}$ (still with $K = 1$), then both sides of (6.5) are the same as if $\mathcal{M} = \varnothing$, and hence one can remove $\mathcal{M}$. In this case, we have an improved bound in terms of a better power of $\theta$. It can be stated in a more general form as follows.

THEOREM 6.4. *Let $\tilde{\mathcal{T}}$ be a finite set of types with even and odd categories $\tilde{\mathcal{E}}$ and $\tilde{\mathcal{O}}$. Then we have*

$$
(6.6) \qquad |\mathbf{E}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Phi(\boldsymbol{\theta}, \mathbf{X})| \lesssim \theta^{|\tilde{\mathcal{O}}|+2|\tilde{\mathcal{E}}|}\mathbf{E}\Big[\Big(\prod_{t\in\tilde{\mathcal{O}}} X_t\Big)\Big(\prod_{t\in\tilde{\mathcal{E}}} X_t^{\diamond 2}\Big)\Big],
$$

*where $\theta = 1 + \max_t |\theta_t|$. If we do not recenter the factors of $\Phi$, then we have the bound*

$$
(6.7) \qquad \Big|\mathbf{E}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}\Big(\prod_{t\in\tilde{\mathcal{T}}} \mathrm{trig}_t(\theta_t X_t)\Big)\Big| \lesssim \theta^{|\tilde{\mathcal{O}}|}\mathbf{E}\prod_{t\in\tilde{\mathcal{O}}} X_t.
$$

*In particular, if $|\tilde{\mathcal{O}}|$ is odd, then both sides of (6.7) vanish.*

REMARK 6.5.    Theorem 6.4 will be applied in two places. We will first use it to get a fine control of the coefficients $C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X}_k)$ (see the expression (6.8)) right after we prove it in Section 6.7. In this case, we take $\tilde{\mathcal{T}} = \mathcal{T}$. Second, we apply it to the small "remainder" part of the nonlinearity $F$ to reduce the regularity assumption for the convergence in Section 5. In this case, we will take $\tilde{\mathcal{T}} = [K] \times \mathcal{T}$, and each point $x_{k,t}$ should be thought of as having a unique type in $\tilde{\mathcal{T}}$. See also [32] for an improvement of Theorem 6.4 in the special case when $\theta_t$ are equal for all $t \in \tilde{T}$.

In the remaining part of this section, we will first set up a general method to prove Theorem 6.4. Since the bound (6.5) is trivial when $K = 1$ and $\mathbf{0} \in \mathcal{M}$, and is indifferent between all other $\mathcal{M}$ that do not contain $\mathbf{0}$ (including the empty set), this already establishes (6.5) for $K = 1$. We will then use Theorem 6.4 to prove Theorem 6.2 for $K \geq 2$. Also note that when both $K$ and $|\mathcal{O}|$ are odd (or just $|\tilde{\mathcal{O}}|$ being odd in the case of Theorem 6.4), all of the above bounds are trivial since both sides of the inequality are 0.

6.2. *Factorial decay of the coefficients.*    In this subsection, we give a bound on the decay of $C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X})$, as defined in (6.4) and (6.1). Recall (6.2) and in particular that $\mathcal{O}_i$ and $\mathcal{E}_i$ depend on $\mathbf{n}$. The definition of $C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X})$ immediately yields the identity

$$
C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X}) = (-1)^{\frac{1}{2}(|\mathbf{n}|+|\mathcal{E}_1|-|\mathcal{O}_1|)} \cdot \frac{\boldsymbol{\theta}^{\mathbf{n}}}{\mathbf{n}!}
$$

$$
(6.8) \qquad \times \mathbf{E}\Big[\Big(\prod_{t\in\mathcal{O}_2\cup\mathcal{E}_1} \sin(\theta_t X_t)\Big)\Big(\prod_{t\in\mathcal{O}_1\cup\mathcal{E}_2} \cos(\theta_t X_t)\Big)\Big(\prod_{t\in\mathcal{E}_0} \langle\!\langle\cos(\theta_t X_t)\rangle\!\rangle\Big)\Big].
$$

The following proposition is then an easy consequence.

PROPOSITION 6.6. *There exists $C > 0$ depending on the multi-index $\mathbf{r}$ only, such that for every $\mathbf{n} \in \mathbf{N}^{\mathcal{T}}$ with $|\mathbf{n}|$ having the same parity as $|\mathcal{O}|$, every $(\theta_t)_{t \in \mathcal{T}}$ and every $(x_t)_{t \in \mathcal{T}}$, one has*

$$(6.9) \qquad \left| \partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{x}) \right| \leq \frac{(C\theta)^{|\mathbf{n}|}}{\mathbf{n}!},$$

*where $\theta = 1 + \max_t |\theta_t|$. As a consequence, we have*

$$(6.10) \qquad \sum_{|\mathbf{n}| = N} \left| \partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{x}) \right| \leq \frac{(C\theta)^N}{N!}$$

*if $N$ has the same parity as $|\mathcal{O}|$, and $0$ otherwise. The constant $C$ in (6.10) depends on $\mathbf{r}$ and $|\mathcal{T}|$.*

PROOF. The parity of $|\mathcal{O}|$ determines whether $\Phi(\boldsymbol{\theta}, \mathbf{X})$ expands into an odd or even chaos series. Thus, $C_{\mathbf{n}}(\boldsymbol{\theta}, \mathbf{X}) = 0$ if $|\mathbf{n}|$ has a different parity as $|\mathcal{O}|$. As for the bounds, the first claim is an immediate consequence of (6.8), and the second claim follows from the multinomial theorem. □

REMARK 6.7. Proposition 6.6 gives the factorial decay of the coefficients, and (a simple variant of it) is enough to prove Theorem 6.4. On the other hand, the proof of Theorem 6.2 requires finer control of the $C_{\mathbf{n}}$'s beyond factorial decay, namely the dependence on $\mathbf{x}$. We will obtain this control from Theorem 6.4.

6.3. *Clustering.* Let $\mathcal{T} = \mathcal{O} \sqcup \mathcal{E}$ be the type space described in Section 6.1, and denote the whole collection of $K|\mathcal{T}|$ points by $(x_u)_{u \in \mathscr{I}}$ with $\mathscr{I} = [K] \times \mathcal{T}$. We also fix a finite set $\mathcal{M} \subset \mathbf{N}^{\mathcal{T}}$ satisfying Assumption 6.1 as well as a "root set" $\mathcal{R}$.

Let $(\theta_t)_{t \in \mathcal{T}}$ be a fixed collection of "frequencies", and $\theta = 1 + \max_t |\theta_t|$. Our aim is to obtain the bound in Theorem 6.2 uniformly over all locations of the points as well as the $\theta_t$'s, except of course for the polynomial dependence in $\theta$ that is already explicit in the statement.

To do this, we divide the points into clusters of size of order $\theta^2 \varepsilon$ in the following way. Let $\sim$ be the transitive closure of the relation on $\mathscr{I}$ given by $u \sim u'$ if $|x_u - x_{u'}| \leq \theta^2 \varepsilon$. We write $\mathscr{C}$ for the partition of $\mathscr{I}$ into clusters obtained in this way. For $\mathcal{U} \subset \mathscr{I}$, we also write $\mathscr{C}(\mathcal{U})$ for the collection of clusters in $\mathscr{C}$ that contain at least one point in $\mathcal{U}$.

Let $\mathcal{S}_{\max}$ denote the set of indices $s$ such that for every cluster $\mathcal{C} \in \mathscr{C}$, $\mathcal{C} \cap (\{s\} \times \mathcal{T}) = \mathcal{C}$ or $\varnothing$. In other words, indices $s$ in $\mathcal{S}_{\max}$ are such that, starting from a point $x_{(s,t)}$ with any $t \in \mathcal{T}$ and performing steps of size at most $\theta^2 \varepsilon$, it is only possible to reach other points $x_{(s,t')}$. Recall the definitions of $\mathscr{T}_{\mathcal{M}}$ and $\Phi$ from (6.1) and (6.4).

For any $\mathcal{S} \subset K$ (but we will always choose $\mathcal{S} \subset \mathcal{S}_{\max}$), we then have

$$
\prod_{k=1}^{K} \mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_k)) = \sum_{\mathcal{I} \sqcup \mathcal{J} = [K] \setminus \mathcal{S}} (-1)^{|\mathcal{I}|} \left[ \left( \prod_{s \in \mathcal{S}} \mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_s)) \right) \right.
$$
$$
(6.11) \qquad \left. \times \left( \prod_{i \in \mathcal{I}} \sum_{\mathbf{p} \in \mathcal{M}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}}^{(i)} \mathbf{X}_i^{\diamond \mathbf{p}} \right) \left( \prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_j) \right) \right],
$$

where the sum is taken over all disjoint subsets $\mathcal{I}$ and $\mathcal{J}$ such that $\mathcal{I} \cup \mathcal{J} = [K] \setminus \mathcal{S}$. Here, $C_{\mathbf{p}}^{(i)}$ denotes the coefficient of the term $\mathbf{X}_i^{\diamond \mathbf{p}}$ in the chaos expansion of $\Phi(\boldsymbol{\theta}, \mathbf{X}_i)$, and hence $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}}^{(i)}$ denotes the coefficient for the same term in $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_i)$. We keep the superscript $i$ to emphasise its dependence on the locations $\mathbf{x}_i = (x_{i,t})_t$. Note that each term on the right-hand side above depends on $\mathcal{S}$, while the left-hand side does not.

REMARK 6.8. While (6.11) holds for any subset $\mathcal{S} \subset [K]$, only $\mathcal{S} \subset \mathcal{S}_{\max}$ gives us useful control on its right-hand side. Such sets $\mathcal{S}$ have the property that for every $s \in \mathcal{S}$, every $k \neq s$ and all $t, t' \in \mathcal{T}$, one has

$$
(6.12) \qquad |x_{s,t} - x_{k,t'}| \geq \theta^2 \varepsilon.
$$

This is the only property that the analysis in Sections 6.4 and 6.6 is based on, and $\mathcal{S}_{\max}$ is the maximal subset of $[K]$ with this property. On the other hand, the final proofs of Theorems 6.4 and 6.2 require more specific choice of $\mathcal{S}$ (either $\varnothing$ or $\mathcal{S}_{\max}$), but this happens only from Section 6.7. Hence, for this moment, we assume $\mathcal{S} \subset \mathcal{S}_{\max}$. When we reach the final stage of the proof of the main theorems, we will make clear the specific choice of $\mathcal{S}$ in those situations.

We now proceed with an arbitrary (fixed) $\mathcal{S} \subset \mathcal{S}_{\max}$. Each term in the sum on the right-hand side of (6.11) is a product of three terms, each being a product over a (possibly empty) subset of $[K]$. For $s \in \mathcal{S}$, considering the corresponding factor in the first term, let $\mathcal{R}$ be any set of "roots" for $\mathcal{M}$ described above, so we have the chaos expansion

$$
(6.13) \qquad \mathscr{T}_{\mathcal{M}}(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_s)) = \sum_{\mathbf{m} \in \mathcal{R}} \sum_{\mathbf{k} \in \mathbf{N}_{\mathcal{E}_0(\mathbf{m})}^{\mathcal{T}}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}+\mathbf{m}}^{(s)} \mathbf{X}_s^{\diamond 2\mathbf{k}+\mathbf{m}},
$$

where we set

$$
(6.14) \qquad \bar{C}_{\mathbf{n}}^{(s)} = C_{\mathbf{n}}^{(s)} / |\{(\mathbf{k}, \mathbf{m}) : \mathbf{m} \in \mathcal{R}, \mathbf{k} \in \mathbf{N}_{\mathcal{E}_0(\mathbf{m})}^{\mathcal{T}}, 2\mathbf{k} + \mathbf{m} = \mathbf{n}\}|.
$$

Now we turn to the term involving the product over $j \in \mathcal{J}$. For every $\mathcal{J} \subset [K] \setminus \mathcal{S}$, let

$$
\mathscr{U}^{\mathcal{J}} := \{ \mathcal{C} \cap (\mathcal{J} \times \mathcal{T}) : \mathcal{C} \in \mathscr{C}, |\mathcal{C} \cap (\mathcal{J} \times \mathcal{T})| \geq 1 \}.
$$

We write $\mathscr{U}^{\mathcal{J}} = \mathscr{U}_0^{\mathcal{J}} \sqcup \mathscr{U}_1^{\mathcal{J}} \sqcup \mathscr{U}_2^{\mathcal{J}}$, where

$$
\begin{aligned}
\mathscr{U}_2^{\mathcal{J}} &= \{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}} : |\mathfrak{u}| = 1, \mathfrak{u} \subset \mathcal{J} \times \mathcal{E}\}, \\
\mathscr{U}_1^{\mathcal{J}} &= \{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}} : |\mathfrak{u} \cap (\mathcal{J} \times \mathcal{O})| \text{ is odd}\}, \\
\mathscr{U}_0^{\mathcal{J}} &= \{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}} : |\mathfrak{u}| \geq 2, |\mathfrak{u} \cap (\mathcal{J} \times \mathcal{O})| \text{ is even}\}.
\end{aligned}
$$

(6.15)

For $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$, we furthermore define subsets $\mathcal{N}_i(\mathfrak{u}) \subset \mathbf{N}^{\mathfrak{u}}$ by

$$
\begin{aligned}
\mathcal{N}_2 &= \{\mathbf{n} : |\mathbf{n}| \geq 2 \text{ is even}\}, \qquad \mathcal{N}_1 = \{\mathbf{n} : |\mathbf{n}| \text{ is odd}\}, \\
\mathcal{N}_0 &= \{\mathbf{n} : |\mathbf{n}| \geq 0 \text{ is even}\},
\end{aligned}
$$

and we set $\bar{\mathcal{N}}(\mathfrak{u}) = \mathcal{N}_i$ for $\mathfrak{u} \in \mathscr{U}_i^{\mathcal{J}}$. These sets do depend on $\mathfrak{u}$ since we only consider $\mathbf{n} \in \mathbf{N}^{\mathfrak{u}}$. Note that in $\mathcal{N}_2$, the "multi-indices" $\mathbf{n}$ are really positive even integers since $|\mathfrak{u}| = 1$. We still write it in $\mathbf{n}$ in order to keep the same notation for all $\mathcal{N}_i$.

Introduce now variables $\boldsymbol{\beta} = (\beta_{j,t})_{j \in [K], t \in \mathcal{T}}$ and write $\boldsymbol{\beta}_j = (\beta_{j,t})_{t \in \mathcal{T}}$. Also, let $\mathbf{a} = (\mathbf{a}_{j,t})_{j \in [K], t \in \mathcal{T}}$ be given by $\mathbf{a}_j = \mathbf{r}$, so that the third product on the right-hand side of (6.11) can be written as

$$
\prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_j) = \prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\beta}_j}^{\mathbf{a}_j} \Phi(\boldsymbol{\beta}_j, \mathbf{X}_j)\Big|_{\boldsymbol{\beta}_j = \boldsymbol{\theta}}.
$$

For any $\mathfrak{u} \subset [K] \times \mathcal{T}$, write $\mathbf{a}_{\mathfrak{u}}$ for the restriction of $\mathbf{a}$ to $\mathfrak{u}$, and similarly for $\boldsymbol{\beta}$, so that

$$
\prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\beta}_j}^{\mathbf{a}_j} \Phi(\boldsymbol{\beta}_j, \mathbf{X}_j) = \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \partial_{\boldsymbol{\beta}_{\mathfrak{u}}}^{\mathbf{a}_{\mathfrak{u}}} \Phi_{\mathfrak{u}}(\boldsymbol{\beta}_{\mathfrak{u}}, \mathbf{X}_{\mathfrak{u}}),
$$

where we further used the notation

$$
\Phi_{\mathfrak{u}}(\boldsymbol{\beta}_{\mathfrak{u}}, \mathbf{X}_{\mathfrak{u}}) = \prod_{(j,t) \in \mathfrak{u}} \{\!\!\{\operatorname{trig}_t(\beta_{j,t} X_{j,t})\}\!\!\}.
$$

Writing similar to before $C_{\mathbf{n}}^{(\mathfrak{u})}(\boldsymbol{\beta}_{\mathfrak{u}})$ for the coefficient of $\mathbf{X}_{\mathfrak{u}}^{\diamond \mathbf{n}}$ in the Wiener chaos expansion of $\Phi_{\mathfrak{u}}(\boldsymbol{\beta}_{\mathfrak{u}}, \mathbf{X}_{\mathfrak{u}})$, we conclude that

(6.16) $$
\prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_j) = \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \sum_{\mathbf{n} \in \bar{\mathcal{N}}(\mathfrak{u})} \partial_{\boldsymbol{\beta}_{\mathfrak{u}}}^{\mathbf{a}_{\mathfrak{u}}} C_{\mathbf{n}}^{(\mathfrak{u})}(\boldsymbol{\beta}_{\mathfrak{u}}) \mathbf{X}_{\mathfrak{u}}^{\diamond \mathbf{n}}\Big|_{\boldsymbol{\beta}_j = \boldsymbol{\theta}, \forall j \in \mathcal{J}}.
$$

In principle, one would think that the sum on the right-hand side of (6.16) should be taken over the whole set $\mathbf{N}^{\mathfrak{u}}$, but since $C_{\mathbf{n}}^{(\mathfrak{u})} = 0$ for $\mathbf{n} \in \mathbf{N}^{\mathfrak{u}} \setminus \bar{\mathcal{N}}(\mathfrak{u})$, we can write the sum as above. We will also write $\partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}}^{(\mathfrak{u})}(\boldsymbol{\theta}_{\mathfrak{u}}) \mathbf{X}_{\mathfrak{u}}^{\diamond \mathbf{n}}$ for each corresponding term on the right-hand side of (6.16), so that we have

(6.17) $$
\prod_{j \in \mathcal{J}} \partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_j) = \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \sum_{\mathbf{n} \in \bar{\mathcal{N}}(\mathfrak{u})} \partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}}^{(\mathfrak{u})}(\boldsymbol{\theta}_{\mathfrak{u}}) \mathbf{X}_{\mathfrak{u}}^{\diamond \mathbf{n}}.
$$

Note that the right-hand side of (6.17) is a *notation* for that of (6.16). Plugging (6.13) and (6.17) back into (6.11) and expanding the product, we obtain the expression

$$(6.18) \quad \mathbf{E} \prod_{k=1}^{K} \mathscr{T}_{\mathcal{M}}\big(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X}_k)\big) = \sum_{\mathcal{I} \sqcup \mathcal{J} = [K] \setminus \mathcal{S}} (-1)^{|\mathcal{I}|} \sum_{\substack{\mathbf{m} \in \mathcal{R}^{\mathcal{S}} \\ \mathbf{p} \in \mathcal{M}^{\mathcal{I}}}} \mathcal{G}_{\mathcal{I}, \mathcal{J}, \mathbf{m}, \mathbf{p}}(\boldsymbol{\theta}, \mathbf{X}),$$

where the range of the outer sum is as before and the terms in the inner sum are given by

$$
\begin{aligned}
&\mathcal{G}_{\mathcal{I}, \mathcal{J}, \mathbf{m}, \mathbf{p}}(\boldsymbol{\theta}, \mathbf{X}) \\
&\quad = \sum_{\mathbf{k}, \mathbf{n}} \bigg( \prod_{s \in \mathcal{S}} \big(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s + \mathbf{m}_s}^{(s)}\big) \bigg) \bigg( \prod_{i \in \mathcal{I}} \big(\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}\big) \bigg) \bigg( \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \big(\partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}_{\mathfrak{u}}}^{(\mathfrak{u})}\big) \bigg) \\
(6.19) &\qquad \times \mathbf{E} \bigg[ \bigg( \prod_{s \in \mathcal{S}} \mathbf{X}_s^{\diamond(2\mathbf{k}_s + \mathbf{m}_s)} \bigg) \bigg( \prod_{i \in \mathcal{I}} \mathbf{X}_i^{\diamond \mathbf{p}_i} \bigg) \bigg( \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \mathbf{X}_{\mathfrak{u}}^{\diamond \mathbf{n}_{\mathfrak{u}}} \bigg) \bigg].
\end{aligned}
$$

Here, the sum is taken over $\mathbf{k} = (\mathbf{k}_s)_{s \in \mathcal{S}}$ where each $\mathbf{k}_s$ runs over $\mathbf{N}_{\mathcal{E}_0(\mathbf{m}_s)}^{\mathcal{T}}$, and $\mathbf{n} = (\mathbf{n}_{\mathfrak{u}})_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}}$ where each $\mathbf{n}_{\mathfrak{u}}$ runs over $\bar{\mathcal{N}}(\mathfrak{u})$. We also use the usual convention that empty products equal 1.

As for Theorem 6.2, for $K \geq 2$ and $\mathcal{S} = \mathcal{S}_{\max}$, we will show that each $\mathcal{G}$ in (6.19) is bounded by the right-hand side of (6.5) (note that this termwise bound is in general false for $\mathcal{S} \neq \mathcal{S}_{\max}$ unless $\mathcal{M} = \varnothing$). Since $\mathcal{S}_{\max}$, $\mathcal{M}$ and $\mathcal{R}$ are all finite, Theorem 6.2 (for $K \geq 2$) then immediately follows from (6.18). Also, as we will obtain the same bound for each choice of $\mathcal{I}$, $\mathcal{J}$, $\mathbf{m}$ and $\mathbf{p}$, we drop these subscripts and simply write $\mathcal{G}(\boldsymbol{\theta}, \mathbf{X})$ for (6.19). We will also drop the dependence on these subsets in other situations whenever no confusion may arise (e.g., in (6.15)).

In the case $K = 1$, we always have $\mathcal{S}_{\max} = \{1\}$ whatever the location of the points. Taking $\mathcal{S} = \mathcal{S}_{\max}$ in this case simply gives the chaos expansion of the left-hand side of (6.18), and no information of the clustering can be used with that expansion. As mentioned in Section 6.1, the only interesting situation under $K = 1$ is when $\mathbf{0} \notin \mathcal{M}$, which is exactly the same as $\mathcal{M} = \varnothing$. Hence, it is reduced to Theorem 6.4, which can be shown by choosing $\mathcal{S} = \varnothing$.

Note that $\mathcal{I}$, $\mathcal{J}$, $\mathbf{m}$, $\mathbf{p}$ (and also $\mathcal{H}$) all depend on $\mathcal{S}$. Since most of the intermediate bounds below are true for all $\mathcal{S} \subset \mathcal{S}_{\max}$, and there are only finitely many choices of $\mathcal{S}$, varying $\mathcal{S}$ only changes the proportionality constants in these bounds. Thus, we will still keep the notation and proceed with general $\mathcal{S} \subset \mathcal{S}_{\max}$. We will specify which $\mathcal{S}$ we use only when it becomes necessary. Also, although $\mathcal{S}_{\max}$, and hence any $\mathcal{S} \subset \mathcal{S}_{\max}$ depends on the location of the points $\{x_{k,t}\}$, the bounds below will be uniform over all locations.

*Notation.* In what follows, we use the notation

$$\mathbf{m}_s = (m_s^t)_{t\in\mathcal{T}}, \qquad \mathbf{m} = (\mathbf{m}_s)_{s\in\mathcal{S}}, \qquad |\mathbf{m}_s| = \sum_{t\in\mathcal{T}}|m_s^t|, \qquad |\mathbf{m}| = \sum_{s\in\mathcal{S}}|\mathbf{m}_s|,$$

and similarly for other multi-indices $\mathbf{k} = (\mathbf{k}_s)_{s\in\mathcal{S}}$, $\mathbf{p} = (\mathbf{p}_i)_{i\in\mathcal{I}}$ and $\mathbf{n} = (\mathbf{n}_h)_{h\in\mathcal{H}}$.

6.4. *The representative point.* We now start to develop ingredients that are needed to bound the right-hand side of (6.19). For this, we choose in an arbitrary way one representative point $u^*(\mathfrak{u})$ from each cluster $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ and we write $X_\mathfrak{u}$ instead of $X_{u^*(\mathfrak{u})}$ for the corresponding random variable. We then show that all of the Wick powers $\mathbf{X}_\mathfrak{u}^{\diamond\mathbf{n}_\mathfrak{u}}$ appearing in (6.19) can be replaced by $X_\mathfrak{u}^{\diamond|\mathbf{n}_\mathfrak{u}|}$, at the cost of a polynomial factor in $\theta$, independently of the specific choice of $u^*(\mathfrak{u})$. The precise statement is the following.

PROPOSITION 6.9. *Let $\mathcal{S} \subset \mathcal{S}_{\max}$ and $\mathscr{U}^{\mathcal{J}}$ be as described above. Let $|\mathbf{n}| = \sum_\mathfrak{u} |\mathbf{n}_\mathfrak{u}|$ and $|\mathbf{p}| = \sum_i |\mathbf{p}_i|$. Then there exists $C > 0$ such that*

$$\mathbf{E}\left[\left(\prod_{s\in\mathcal{S}}\mathbf{X}_s^{\diamond(2\mathbf{k}_s+\mathbf{m}_s)}\right)\left(\prod_{i\in\mathcal{I}}\mathbf{X}_i^{\diamond\mathbf{p}_i}\right)\left(\prod_{\mathfrak{u}\in\mathscr{U}^{\mathcal{J}}}\mathbf{X}_\mathfrak{u}^{\diamond\mathbf{n}_\mathfrak{u}}\right)\right]$$

$$(6.20) \qquad \leq C^{|\mathbf{n}|}\theta^{2|\mathbf{p}|}\mathbf{E}\left[\left(\prod_{s\in\mathcal{S}}\mathbf{X}_s^{\diamond(2\mathbf{k}_s+\mathbf{m}_s)}\right)\left(\prod_{i\in\mathcal{I}}\mathbf{X}_i^{\diamond\mathbf{p}_i}\right)\left(\prod_{\mathfrak{u}\in\mathscr{U}^{\mathcal{J}}}X_\mathfrak{u}^{\diamond|\mathbf{n}_\mathfrak{u}|}\right)\right],$$

*uniformly over the exponents $(\mathbf{k}_s)$, $(\mathbf{p}_i)$, $(\mathbf{n}_\mathfrak{u})$, and all locations of points satisfying the constraints enforced by the definitions of $\mathcal{S}$, $\mathcal{I}$ and $\mathcal{J}$. The constant $C$ can be taken $C = 2\Lambda^2 \max_\mathfrak{u} |\mathfrak{u}|$, where $\Lambda$ is as in (4.1).*

REMARK 6.10. Note that on the left-hand side of (6.20), the boldface letter $\mathbf{X}_\mathfrak{u}$ refers to the collection of random variables $(X_u)_{u\in\mathfrak{u}}$, so

$$\mathbf{X}_\mathfrak{u}^{\diamond\mathbf{n}_\mathfrak{u}} = \bigdiamond_{u\in\mathfrak{u}} X_u^{\diamond n^u}.$$

On the other hand, $X_\mathfrak{u}$ with the normal capital $X$ on the right-hand side refers to the random variable $X_{u^*(\mathfrak{u})}$. Hence, the Wick power $|\mathbf{n}_\mathfrak{u}|$ is an integer instead of a multi-index.

PROOF. When the sum of the exponents $|2\mathbf{k} + \mathbf{m}| + |\mathbf{p}| + |\mathbf{n}|$ is odd, then both sides of (6.20) are 0, so we only consider the case when it is even.

By Wick's formula, both sides of (6.20) are a sum over products of pairwise expectations, and the number of pairings for the pairwise products are equal. Furthermore, there is a natural one-to-one correspondence between the pairings on the two sides in that every factor of $X_u$ on the left-hand side with $u \in \mathfrak{u}$ for some $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ is replaced by $X_\mathfrak{u}$ on the right-hand side. Since all correlations are positive, it thus suffices to control the effect of such replacements. Consider $(i, t) \in [K] \times \mathcal{T}$,

$\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$, and $u, u^* \in \mathfrak{u}$. We then distinguish two cases. In the first case, one has $|x_{(i,t)} - x_u| \geq \theta^2 \varepsilon$, in which case the triangle inequality and the definition of a cluster imply that

$$|x_{(i,t)} - x_{u^*}| \leq |x_{(i,t)} - x_u| + |x_u - x_{u^*}| \leq |\mathfrak{u}||x_{(i,t)} - x_u|.$$

It then follows from Lemma 4.2 that in this case

$$\mathbf{E}(X_{i,t}X_u) \leq \Lambda^2 |\mathfrak{u}| \mathbf{E}(X_{i,t}X_{u^*}).$$

If instead $|x_{(i,t)} - x_u| < \theta^2 \varepsilon$, then $|x_{(i,t)} - x_{u^*}| \leq |\mathfrak{u}|\theta^2 \varepsilon$. We then conclude from Lemma 4.1 that

$$\mathbf{E}(X_{i,t}X_{u^*}) \geq \frac{1}{\Lambda(1 + |\mathfrak{u}|\theta^2)} \geq \frac{1}{2|\mathfrak{u}|\Lambda^2\theta^2} \mathbf{E}(X_{(i,t)}X_u).$$

Note now that by the properties of the Wick product, correlations of the type $\mathbf{E}(X_{(i,t)}X_u)$ only ever show up with either $(i, t) \in \mathfrak{u}'$ for some $\mathfrak{u}' \neq \mathfrak{u}$ in $\mathscr{U}^{\mathcal{J}}$, or $i \in \mathcal{S}$, or $i \in \mathcal{I}$. In the first two cases, it follows from the definitions of $\mathcal{S}$ and our clusters that we are necessarily in the situation $|x_{(i,t)} - x_u| \geq \theta^2 \varepsilon$, so the converse can only arise for $i \in \mathcal{I}$.

In order to conclude, it suffices to note that the number of factors with $i \in \mathcal{I}$ is precisely $|\mathbf{p}|$, while the total number of factors that require a substitution is $|\mathbf{n}| = \sum_{\mathfrak{u}} |\mathbf{n}_{\mathfrak{u}}|$. □

REMARK 6.11. The proposition says that the replacement by a single representative point for all $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ costs $2|\mathbf{p}|$ powers of $\theta$. Since by (6.11), we are only interested in those $p$ with each $\mathbf{p}_i \in \mathcal{M}$, the total cost of powers in $\theta$ is always finite and depends on $\mathcal{M}$ and $K$ only. In particular, for fixed $\mathcal{M}$, it is linear in $K$.

Now, for every $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ and $i = 0, 1, 2$, let $q_{\mathfrak{u}} = i$ if $\mathfrak{u} \in \mathscr{U}_i^{\mathcal{J}}$. We have the following proposition.

PROPOSITION 6.12. Recall that $|\mathbf{p}| = \sum_i |\mathbf{p}_i|$. There exists $C > 0$ depending on $\Lambda$, $\mathbf{r}$ and the total number of points such that for every $\mathcal{G}$ in (6.19), we have

$$\left| \mathcal{G}(\boldsymbol{\theta}, \mathbf{X}) \right|$$

$$\leq \theta^{2|\mathbf{p}|} \sum_{\mathbf{k}, \boldsymbol{\ell}} \left( \prod_{s \in \mathcal{S}} |\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s + \mathbf{m}_s}^{(s)}| \right) \left( \prod_{i \in \mathcal{I}} |\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}| \right) \left( \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \frac{(C\theta)^{2\ell_{\mathfrak{u}} + q_{\mathfrak{u}}}}{(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})!} \right)$$

(6.21)   $$\times \mathbf{E} \left[ \left( \prod_{s \in \mathcal{S}} \mathbf{X}_s^{\diamond(2\mathbf{k}_s + \mathbf{m}_s)} \right) \left( \prod_{i \in \mathcal{I}} \mathbf{X}_i^{\diamond \mathbf{p}_i} \right) \left( \prod_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} X_{\mathfrak{u}}^{\diamond(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})} \right) \right],$$

where the sum is taken over $\mathbf{k} = (\mathbf{k}_s)_{s \in \mathcal{S}}$ where every $\mathbf{k}_s$ runs through $\mathcal{N}_{\mathcal{E}_0(\mathbf{m}_s)}^{\mathcal{T}}$, and $\boldsymbol{\ell} = (\ell_{\mathfrak{u}})_{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}} \in \mathbf{N}^{\mathscr{U}^{\mathcal{J}}}$. The bound is uniform over all locations of points.

PROOF. Applying Proposition 6.9 to (6.19), we get

$$|\mathcal{G}(\boldsymbol{\theta}, \mathbf{X})| \leq \theta^{2|\mathbf{p}|} \sum_{\mathbf{k}, \mathbf{n}} \left( \prod_{s \in \mathcal{S}} |\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s + \mathbf{m}_s}^{(s)}| \right) \left( \prod_{i \in \mathcal{I}} |\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}| \right) \left( \prod_{\mathfrak{u} \in \mathcal{U}^{\mathcal{J}}} C^{|\mathbf{n}_{\mathfrak{u}}|} |\partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}_{\mathfrak{u}}}^{(\mathfrak{u})}| \right)$$

$$\times \mathbf{E}\left[ \left( \prod_{s \in \mathcal{S}} \mathbf{X}_s^{\diamond (2\mathbf{k}_s + \mathbf{m}_s)} \right) \left( \prod_{i \in \mathcal{I}} \mathbf{X}_i^{\diamond \mathbf{p}_i} \right) \left( \prod_{\mathfrak{u} \in \mathcal{U}^{\mathcal{J}}} X_{\mathfrak{u}}^{\diamond |\mathbf{n}_{\mathfrak{u}}|} \right) \right],$$

where we have used $|\mathbf{n}| = \sum_{\mathfrak{u}} |\mathbf{n}_{\mathfrak{u}}|$ to decompose the constant $C^{|\mathbf{n}|}$ from Proposition 6.9 into factors of $C^{|\mathbf{n}_{\mathfrak{u}}|}$ and distributed them into the corresponding terms in the product. The sum over $\mathbf{k}$ has the same range as in (6.19), while $\mathbf{n} = (\mathbf{n}_{\mathfrak{u}})_{\mathfrak{u} \in \mathcal{U}^{\mathcal{J}}}$ with each $\mathbf{n}_{\mathfrak{u}}$ running through

$$\bar{\mathcal{N}}(\mathfrak{u}) = \{\mathbf{n}_{\mathfrak{u}} : |\mathbf{n}_{\mathfrak{u}}| = 2\ell_{\mathfrak{u}} + q_{\mathfrak{u}}, \ell_{\mathfrak{u}} \geq 0\}.$$

By Proposition 6.6, the coefficient $\partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}_{\mathfrak{u}}}^{(\mathfrak{u})}$ satisfies the bound

$$|\partial_{\boldsymbol{\theta}_{\mathfrak{u}}}^{\mathbf{r}_{\mathfrak{u}}} C_{\mathbf{n}_{\mathfrak{u}}}^{(\mathfrak{u})}| \leq \frac{(C\theta)^{|\mathbf{n}_{\mathfrak{u}}|}}{\mathbf{n}_{\mathfrak{u}}!}.$$

We then sum these coefficients over the level sets $\{\mathbf{n}_{\mathfrak{u}} : |\mathbf{n}_{\mathfrak{u}}| = 2\ell_{\mathfrak{u}} + q_{\mathfrak{u}}\}$ for every fixed $\boldsymbol{\ell}$, and the claim follows from the multinomial theorem. $\square$

Propositions 6.9 and 6.12 hold for all possible choices of $u^*(\mathfrak{u}) \in \mathfrak{u}$. But in the context below, it will be convenient to make more specific choices based on the cluster $\mathfrak{u}$. More precisely, we let $u^*(\mathfrak{u})$ be any point in $\mathfrak{u} \cap (\mathcal{J} \times \mathcal{O})$ if $\mathfrak{u} \in \mathcal{U}_1^{\mathcal{J}}$, and arbitrary otherwise. Note that for $\mathfrak{u} \in \mathcal{U}_2^{\mathcal{J}}$, since the only point there belongs to $\mathcal{J} \times \mathcal{E}$, so $u^*$ has to be even. We fix this choice throughout the rest of this section.

The parities of the chosen points will only be used in Sections 6.7 and 6.9 below. Hence, we will still use the notation $x_{\mathfrak{u}}$ or $X_{\mathfrak{u}}$ when we do not use those properties.

6.5. *The graphic representation.* Our aim now is to bound the right-hand side of (6.21), which contains the expectation of products of arbitrarily high Wick powers of Gaussians. Such an expectation can be written as a sum over products of pairwise expectations. In order to describe our objects and bounds in a convenient way, we introduce graphical notation to describe products of pairwise expectations, and we perform most operations at the graphical level.

Given a set $\mathbb{V}$, write $\mathbb{V}_2$ for the set of all subsets $\{u, v\} \subset \mathbb{V}$ with exactly two elements. A (generalised) graph is a pair $\Gamma = (\mathbb{V}, \mathbb{E})$, where $\mathbb{V}$ is a set of vertices and $\mathbb{E} \colon \mathbb{V}_2 \to \mathbf{N}$ is the set of edges with multiplicities. More precisely, each edge $\{u, v\} \in \mathbb{V}_2$ has a multiplicity $\mathbb{E}(u, v) = \mathbb{E}(v, u)$. We do not allow self-loops, so $\mathbb{E}(u, u) = 0$ for all $u \in \mathbb{V}$.

Given a graph $\Gamma = (\mathbb{V}, \mathbb{E})$, for every $u \in \mathbb{V}$, we define the degree of $u$ by

$$\deg(u) := \sum_{v \in \mathbb{V}} \mathbb{E}(u, v),$$

and the total degree of the graph $\Gamma$ is defined by

$$\deg(\Gamma) := \sum_{u \in \mathbb{V}} \deg(u) = 2 \sum_{\{u,v\} \subset \mathbb{V}_2} \mathbb{E}(u, v).$$

In this article, the vertex set $\mathbb{V}$ will always be a subset of $[K] \times \mathcal{T}$, and we identify vertices of $\Gamma$ with a finite collection of space–time points $\{x_u\}$. Recall that $X_u = \varepsilon^{\frac{1}{2}} \Psi_\varepsilon(x_u)$. Defining $R : \mathbb{V}_2 \to \mathbf{R}^+$ by $R(u, v) = \mathbf{E} X_u X_v$, we assign to $\Gamma$ a (positive) value $|\Gamma| = \prod_{e \in \mathbb{V}_2} (R(e))^{\mathbb{E}(e)}$. It is then clear that $|\Gamma|$ is one of the terms appearing in the expectation $\mathbf{E}(\prod_u X_u^{\diamond \ell_u})$. On the other hand, in order to encode Wick products between $X_u$ and $X_v$ for $u \neq v$, we introduce the notion of admissible graphs.

DEFINITION 6.13. Let $\mathcal{A}, \mathcal{B}$ be finite index sets, let $\pi : \mathcal{A} \to \mathcal{B}$, and let $\mathbf{x} = \{x_a\}_{a \in \mathcal{A}}$ be a collection of points. For every $\mathbf{m} \in \mathbf{N}^{\mathcal{A}}$ and every $k \in \mathcal{B}$, let $\mathbf{m}_k$ denote the restriction of $\mathbf{m}$ to $\pi^{-1}(k)$. Then a graph $\Gamma$ with vertex set $\mathcal{A}$ is admissible with respect to $(\pi, \mathbf{m})$ if $\deg a = m_a$ and if $\mathbb{E}(a, b) = 0$ as soon as $\pi(a) = \pi(b)$.

REMARK 6.14. We will use Definition 6.13 in the context of bounding products of the type

$$\prod_{b \in \mathcal{B}} \left( \bigdiamond_{a \in \pi^{-1}(b)} \mathbf{X}_a^{\diamond m_a} \right). \tag{6.22}$$

In this case, we also say that "$\Gamma$ is admissible with respect to (6.22)", with $\pi$ and $\mathbf{m}$ implied from the expression appearing on the right-hand side.

Given a graph of total degree $N$, we would like to bound its value by that of a new graph with smaller degree, obtained by an explicit operation on the edges of the original graph. We introduce a few more notions to better describe these operations.

Since we will always ignore isolated vertices, we say that $\Gamma' = (\mathbb{V}', \mathbb{E}')$ is a subgraph of $\Gamma = (\mathbb{V}, \mathbb{E})$ if $\mathbb{V}' = \mathbb{V}$ and $\mathbb{E}' \leq \mathbb{E}$. We say that the collection $\{\Gamma_k = (\mathbb{V}, \mathbb{E}_k)\}$ of subgraphs of $\Gamma$ adds up to $\Gamma = (\mathbb{V}, \mathbb{E})$ if $\sum_k \mathbb{E}_k = \mathbb{E}$. The definition of $|\Gamma|$ then implies that $|\Gamma| = \prod_k |\Gamma_k|$ if $\{\Gamma_k\}$ adds up to $\Gamma$. Lemma 4.1 also simply translates into the following lemma.

LEMMA 6.15. *For $a, b, c \geq 0$, we have the bound*



$$\left| \ \vcenter{} \ \right| \leq \frac{2\Lambda^3 \varepsilon}{\min\{|x - y|, |x - z|\} + \varepsilon} \cdot \left| \ \vcenter{} \ \right|,$$

*where $\Lambda$ is the same as in (4.1).*

Note that there is a factor $\varepsilon$ in the numerator since an edge between two points $u$ and $v$ stands for the correlation $\mathbf{E}(X_u X_v) = \varepsilon \mathbf{E}(\Psi_\varepsilon(u)\Psi_\varepsilon(v))$.

REMARK 6.16. Strictly speaking, the value of a graph depends on $\varepsilon$ since the random field itself does, so a more proper notation is $|\Gamma|_\varepsilon$ instead of $|\Gamma|$. However, as we shall see later, all the bounds we obtain in this section are uniform in $\varepsilon$, so we omit it in notation for simplicity.

6.6. *Backward induction.* Proposition 6.12 reduces proving Theorem 6.2 to bounding the right-hand side of (6.21). It involves a sum over indices $\mathbf{k}_s$ and $\ell_\mathfrak{u}$ all the way to infinity. In order to control it by polynomials up to a fixed degree to match the right-hand side of (6.5), we control correlation functions of products of Wick polynomials with high degrees by those with lower degrees via a backward induction argument.

Recall the index sets $\mathcal{S}, \mathcal{I}, \mathcal{J}$ and collection of subclusters $\mathscr{U}^\mathcal{J}$ described above. We fix as before an arbitrary $u^*(\mathfrak{u}) \in \mathfrak{u}$ for every $\mathfrak{u} \in \mathscr{U}^\mathcal{J}$ and write $x_\mathfrak{u} = x_{u^*(\mathfrak{u})}$, $X_\mathfrak{u} = X_{u^*(\mathfrak{u})}$, etc. Also recall the notation $\mathbf{x}_s = (x_{s,t})_{t \in \mathcal{T}}$ and $\mathbf{x}_i = (x_{i,t})_{t \in \mathcal{T}}$. In this subsection, we fix the vertex set to be

$$(6.23) \qquad \mathbb{V} = ((\mathcal{S} \cup \mathcal{I}) \times \mathcal{T}) \cup \{u^*(\mathfrak{u}) : \mathfrak{u} \in \mathscr{U}^\mathcal{J}\}.$$

Henceforth, for $i \in \mathcal{S} \cup \mathcal{J}$, $\mathbf{x}_i$ may denote either a $\mathcal{T}$-tuple $(x_{i,t})_{t \in \mathcal{T}}$ or a collection of $|\mathcal{T}|$ points $\{x_{i,t}\}_{t \in \mathcal{T}}$. Fix $\mathbf{m} = (\mathbf{m}_s)_{s \in \mathcal{S}} \in \mathcal{R}^\mathcal{S}$, $\mathbf{p} = (\mathbf{p}_i)_{i \in \mathcal{I}} \in \mathcal{M}^\mathcal{I}$ and $\mathbf{q} = (q_\mathfrak{u})_{\mathfrak{u} \in \mathscr{U}^\mathcal{J}} \in \mathbf{N}^{\mathscr{U}^\mathcal{J}}$. Here in this subsection, we do not use any of the properties of $\mathcal{M}$, $\mathcal{R}$ or $\mathbf{q}$ described before, so they can be any subset of $\mathbf{N}^\mathcal{T}$ (or any integer for $q_\mathfrak{u}$). In particular, Proposition 6.20 below does not depend on $\mathcal{M}$ satisfying Assumption 6.1.

For every pair $(\mathbf{k}, \boldsymbol{\ell})$ such that

$$\mathbf{k} = (\mathbf{k}_s)_{s \in \mathcal{S}} \in \underset{s \in \mathcal{S}}{\times} \mathbf{N}^\mathcal{T}_{\mathcal{E}_0(\mathbf{m}_s)} \quad \text{and} \quad \boldsymbol{\ell} = (\ell_\mathfrak{u})_{\mathfrak{u} \in \mathscr{U}^\mathcal{J}} \in \mathbf{N}^{\mathscr{U}^\mathcal{J}},$$

let $\Omega_{\mathbf{k}, \boldsymbol{\ell}}$ be the set of admissible graphs, in the sense of Remark 6.14, for the product

$$(6.24) \qquad \left(\prod_{s \in \mathcal{S}} \mathbf{X}_s^{\diamond(2\mathbf{k}_s + \mathbf{m}_s)}\right)\left(\prod_{i \in \mathcal{I}} \mathbf{X}_i^{\diamond \mathbf{p}_i}\right)\left(\prod_{\mathfrak{u} \in \mathscr{U}^\mathcal{J}} X_\mathfrak{u}^{\diamond(2\ell_\mathfrak{u} + q_\mathfrak{u})}\right).$$

Our aim is to control the value of graphs in $\Omega_{\mathbf{k}, \boldsymbol{\ell}}$ for large $(\mathbf{k}, \boldsymbol{\ell})$ by those in $\Omega_{\mathbf{k}', \boldsymbol{\ell}'}$ with smaller $(\mathbf{k}', \boldsymbol{\ell}')$. To be more precise about the type of graphs which control those in $\Omega_{\mathbf{k}, \boldsymbol{\ell}}$, we introduce the following definition.

DEFINITION 6.17. Let $\Omega^*$ be the set of graphs $\Gamma$ with vertex set $\mathbb{V}$ given in (6.23) such that all of the following hold:

1. $\Gamma \in \Omega_{\mathbf{k}, \boldsymbol{\ell}}$ for some $(\mathbf{k}, \boldsymbol{\ell})$.

3536 M. HAIRER AND W. XU

2. If $k_s^t \geq 1$ for some $(s, t) \in \mathcal{S} \times \mathcal{T}$, then precisely one of the following is true for the corresponding point $x_{s,t}$:

   - There exists $s' \in \mathcal{S} \setminus \{s\}$ such that $\mathbb{E}(x_{s,t}, x_v) = 0$ whenever $v \notin \{s'\} \times \mathcal{T}$. Furthermore, if $\mathbb{E}(x_{s,t}, x_{s',t'}) \geq 2$ for that $s'$ and for some $t'$, then $k_{s'}^{t'} = 0$.
   - There exists $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ with $\ell_\mathfrak{u} = 0$ such that $\mathbb{E}(x_{s,t}, x_v) = 0$ whenever $v \neq \mathfrak{u}$.
   - There exists $i \in \mathcal{I}$ such that $\mathbb{E}(x_{s,t}, x_v) = 0$ whenever $v \notin \{i\} \times \mathcal{T}$.

3. If $\ell_\mathfrak{u} \geq 1$ for some $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$, then precisely one of the following is true for the corresponding point $x_\mathfrak{u}$:

   - There exists $s \in \mathcal{S}$ such that $\mathbb{E}(x_\mathfrak{u}, x_v) = 0$ whenever $v \notin (\{s\} \cup \mathcal{I}) \times \mathcal{T}$. Furthermore, if $\mathbb{E}(x_\mathfrak{u}, x_{s,t}) \geq 1$ for that $s$ and some $t \in \mathcal{T}$, then $k_s^t = 0$.
   - There exists $\mathfrak{u}' \in \mathscr{U}^{\mathcal{J}} \setminus \{\mathfrak{u}\}$ such that $\mathbb{E}(x_\mathfrak{u}, x_v) = 0$ whenever $v \notin \{\mathfrak{u}'\} \cup (\mathcal{I} \times \mathcal{T})$. Furthermore, if $\mathbb{E}(x_\mathfrak{u}, x_{\mathfrak{u}'}) \geq 2$, then $\ell_{\mathfrak{u}'} = 0$.

REMARK 6.18. The second and third conditions above impose some constraints on the pair $(\mathbf{k}, \boldsymbol{\ell})$ in the first condition. In particular, $\Omega^* \cap \Omega_{\mathbf{k}, \boldsymbol{\ell}} \neq \varnothing$ only when

$$
\begin{aligned}
2k_s^t + m_s^t &\leq \max_{\tilde{s}, i, \tilde{\mathfrak{u}}} \{|\mathbf{m}_{\tilde{s}}| + |\mathcal{T}|, |\mathbf{p}_i|, q_{\tilde{\mathfrak{u}}}\} \quad \text{and} \\
2\ell_\mathfrak{u} + q_\mathfrak{u} &\leq \max_{\tilde{s}, \tilde{\mathfrak{u}}} \{|\mathbf{m}_{\tilde{s}}|, q_{\tilde{\mathfrak{u}}} + 1\} + |\mathbf{p}|,
\end{aligned}
\tag{6.25}
$$

for all $(s, t) \in \mathcal{S} \times \mathcal{T}$ and all $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$, where the maximum over $\tilde{s}$ and $\tilde{\mathfrak{u}}$ are taken over $\mathcal{S}$ and $\mathscr{U}^{\mathcal{J}}$, respectively. As a consequence, there can be only finitely many graphs (depending on $|\mathbf{m}|$, $|\mathbf{p}|$, $|\mathbf{q}|$ and $|\mathcal{T}|$) in $\Omega^*$.

REMARK 6.19. The main difference between the second and third constraints is that in the third one, we do not remove "extra" edges between $x_\mathfrak{u}$ and $x_{i,t}$. The main reason is that there is no assumption on the distance between $\mathbf{x}_i$ and the clusters $\{\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}\}$. Reducing the situation to one with the third constraint analogous to the second one would cause some additional powers of $\theta$ in the bound, which compensates the decrease of the degrees. It also complicates the argument, so we leave it as it is for simplicity. This will not make a difference in the final statement since the regularity we require for $F$ follows from Theorem 6.4, which does not involve any point in $\mathcal{I} \times \mathcal{T}$.

PROPOSITION 6.20. *There exists $C \geq 1$ depending only on the value $\Lambda$ in* (4.1) *such that*

$$
\max_{\Gamma \in \Omega_{\mathbf{k}, \boldsymbol{\ell}}} \theta^{\deg(\Gamma)} |\Gamma| \leq C^{|\mathbf{k}| + |\boldsymbol{\ell}|} \max_{\Gamma^* \in \Omega^*} \theta^{\deg(\Gamma^*)} |\Gamma^*|,
\tag{6.26}
$$

*for every pair $(\mathbf{k}, \boldsymbol{\ell})$.*

PROOF. We claim that whenever $\Gamma \in \Omega_{\mathbf{k},\boldsymbol{\ell}} \setminus \Omega^*$, there exists a $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}},\bar{\boldsymbol{\ell}}}$ for some $\bar{\mathbf{k}} \leq \mathbf{k}$ and $\bar{\boldsymbol{\ell}} \leq \boldsymbol{\ell}$ with $\bar{\mathbf{k}} + \bar{\boldsymbol{\ell}} < \mathbf{k} + \boldsymbol{\ell}$ such that

$$(6.27) \qquad |\Gamma| \leq C^{|\mathbf{k}-\bar{\mathbf{k}}|+|\boldsymbol{\ell}-\bar{\boldsymbol{\ell}}|} \theta^{\deg(\bar{\Gamma})-\deg(\Gamma)} |\bar{\Gamma}|.$$

This bound can then be iterated until one reaches $\bar{\Gamma} \in \Omega^*$. Since furthermore $\Omega_{\mathbf{0},\mathbf{0}} \subset \Omega^*$, this necessarily happens after at most $|\mathbf{k} + \boldsymbol{\ell}|$ steps, which then concludes the proof. It remains to exhibit a $\bar{\Gamma}$ as above for every $\Gamma$. We distinguish four different cases which cover the set $\Omega_{\mathbf{k},\boldsymbol{\ell}} \setminus \Omega^*$.

*Case* 1: There exists $(s,t) \in \mathcal{S} \times \mathcal{T}$ such that $k_s^t \geq 1$ and such that we can find two other vertices $u$ and $v$ with $\mathbb{E}((s,t),u) \wedge \mathbb{E}((s,t),v) \geq 1$ and such that $\bar{\Gamma} \in \Omega_{\mathbf{k},\boldsymbol{\ell}}$ does *not* imply $\mathbb{E}(u,v) = 0$. In this case, we define $\bar{\Gamma}$ so that it differs from $\Gamma$ solely by decreasing $\mathbb{E}((s,t),u)$ and $\mathbb{E}((s,t),v)$ by 1, while at the same time increasing $\mathbb{E}(u,v)$ by 1. It is then immediate that $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}},\bar{\boldsymbol{\ell}}}$ with $\bar{\mathbf{k}} = \mathbf{k} - \mathbf{1}_{(s,t)}$ and $\bar{\boldsymbol{\ell}} = \boldsymbol{\ell}$. By Lemma 6.15 and the fact that $\min\{|x_{s,t} - x_u|, |x_{s,t} - x_v|\} \geq \theta^2 \varepsilon$ by the definition of $\mathcal{S}$ (and the fact that neither $u$ nor $v$ belong to $\{s\} \times \mathcal{T}$ by the definition of $\Omega_{\mathbf{k},\boldsymbol{\ell}}$), one has the bound



Since $\deg \bar{\Gamma} = \deg \Gamma - 2$, this bound is indeed of the from (6.27) as required.

*Case* 2: There exists $(s,t) \in \mathcal{S} \times \mathcal{T}$ such that $k_s^t \geq 1$ and one of the following two conditions hold:

- There exists $u = (s',t') \in \mathcal{S} \times \mathcal{T}$ with $s' \neq s$ and $k_{s'}^{t'} \geq 1$ such that $\mathbb{E}((s,t),u) \geq 2$.
- There exists $u = \mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ with $\ell_{\mathfrak{u}} \geq 1$ and $\mathbb{E}((s,t),u) \geq 2$.

In this case, we define $\bar{\Gamma}$ so that it differs from $\Gamma$ solely by decreasing $\mathbb{E}((s,t),u)$ by 2, so that $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}},\bar{\boldsymbol{\ell}}}$ with $\bar{\mathbf{k}} = \mathbf{k} - \mathbf{1}_{(s,t)} - \mathbf{1}_{(s',t')}$ and $\bar{\boldsymbol{\ell}} = \boldsymbol{\ell}$ in the first case, while $\bar{\mathbf{k}} = \mathbf{k} - \mathbf{1}_{(s,t)}$ and $\bar{\boldsymbol{\ell}} = \boldsymbol{\ell} - \mathbf{1}_{\mathfrak{u}}$ in the second case. By Lemma 4.1 and the fact that $|x_{s,t} - x_u| \geq \theta^2 \varepsilon$, we then obtain the bound

$$(6.28) \qquad \left| \overset{(s,t)}{\bullet}\!\!-\!\!-_{a+2}\!\!-\!\!-\overset{u}{\bullet} \right| \leq \Lambda^2 \theta^{-4} \left| \overset{(s,t)}{\bullet}\!\!-\!\!-_{a}\!\!-\!\!-\overset{u}{\bullet} \right|.$$

Since this time $\deg \bar{\Gamma} = \deg \Gamma - 4$, this is again of the from (6.27) as required.

*Case* 3: There exists $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ with $\ell_{\mathfrak{u}} \geq 1$ and such that we can find two other vertices $u$ and $v$ with $\mathbb{E}(\mathfrak{u},u) \wedge \mathbb{E}(\mathfrak{u},v) \geq 1$ such that $u, v \notin \mathcal{I} \times \mathcal{T}$ and such that $\bar{\Gamma} \in \Omega_{\mathbf{k},\boldsymbol{\ell}}$ does *not* imply $\mathbb{E}(u,v) = 0$. In this case, we proceed exactly as in Case 1, with $(s,t)$ replaced by $\mathfrak{u}$. Here, the fact that $u, v \notin \mathcal{I} \times \mathcal{T}$ is crucial to guarantee that both points are at distance at least $\theta^2 \varepsilon$ from $x_{s,t}$.

*Case* 4: There exist $\mathfrak{u}, \mathfrak{u}' \in \mathscr{U}^{\mathcal{J}}$ such that $\ell_{\mathfrak{u}} \geq 1$ and $\mathbb{E}(x_{\mathfrak{u}}, x_{\mathfrak{u}'}) \geq 2$. In this case, we proceed as in Case 2, noting that $\bar{\Gamma} \in \Omega_{\bar{\mathbf{k}}, \bar{\boldsymbol{\ell}}}$ with $\bar{\mathbf{k}} = \mathbf{k}$ and $\bar{\boldsymbol{\ell}} = \boldsymbol{\ell} - \mathbf{1}_{\mathfrak{u}} - \mathbf{1}_{\mathfrak{u}'}$ and that we have again the bound (6.28), but with $(s, t)$ replaced by $\mathfrak{u}$, which is again of the from (6.27) as required.

It remains to show that the four cases above do indeed cover all of $\Omega_{\mathbf{k}, \boldsymbol{\ell}} \setminus \Omega^*$. Comparing these cases to Definition 6.17, the only way in which $\Gamma$ could possibly fail to belong to $\Omega^*$ which is not obviously covered by these cases is to have points $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$ and $(s, t) \in \mathcal{S} \times \mathcal{T}$ such that $\mathbb{E}((s, t), \mathfrak{u}) = 1$, $\ell_{\mathfrak{u}} \geq 1$, and $k_s^t \geq 1$. However, this case must either be covered by Case 1, or $(s, t)$ is only connected to $\mathfrak{u}$, in which case one must have $\mathbb{E}((s, t), \mathfrak{u}) = 2k_s^t + m_s^t \geq 2$ by the definition of $\Omega_{\mathbf{k}, \boldsymbol{\ell}}$, implying that it is covered by Case 4.  $\square$

6.7. *Proof of Theorem* 6.4.  We are now ready to prove Theorem 6.4. We will show it for our fixed-type space $\mathcal{T} = \mathcal{O} \sqcup \mathcal{E}$, and the statement of the theorem for $\tilde{\mathcal{T}}$ is just a change of notation. We will mainly focus on the bound (6.6), and briefly explain how one can obtain (6.7) by slightly modifying the argument.

We assume $|\mathcal{O}|$ is even, for otherwise both sides of (6.6) and (6.7) vanish and there is nothing to prove. Let $\boldsymbol{\theta} = (\theta_t)_{t \in \mathcal{T}}$ be a collection of frequencies and $\theta = 1 + \max_t |\theta_t|$ as before. The left-hand side of (6.6) corresponds to the case $K = 1$ and $\mathcal{M} = \varnothing$ in the identity (6.18). We choose $\mathcal{S} = \varnothing$, so there are two terms on its right-hand side. If $\mathcal{I}$ is not empty, then the corresponding term vanishes since $\mathcal{M}$ is empty. Hence, we only need to control (6.19) in the case $\mathcal{I} = \varnothing$ and $\mathcal{J} = [K] = \{1\}$. We will then simply drop the notation involving $\mathcal{J}$ or elements from it. For example, we will write $\mathscr{U}$ instead of $\mathscr{U}^{\mathcal{J}}$, and $u \in \mathcal{E}$ instead of $u \in \{1\} \times \mathcal{E}$, etc.

For $j \in \{0, 1, 2\}$, recall the definition of $\mathscr{U}_j$ from (6.15). By Proposition 6.12, we then have the bound

$$(6.29) \qquad |\mathbf{E}\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X})| \leq \sum_{\boldsymbol{\ell} \in \mathbf{N}^{\mathscr{U}}} \mathbf{E}\Bigg[\prod_{\mathfrak{u} \in \mathscr{U}} \frac{(C\theta)^{2\ell_{\mathfrak{u}} + q_{\mathfrak{u}}}}{(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})!} X_{\mathfrak{u}}^{\diamond(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})}\Bigg],$$

where $C$ depends on $\Lambda$, $|\mathcal{T}|$ and $\mathbf{r}$ only, and $q_{\mathfrak{u}} = j$ if $\mathfrak{u} \in \mathscr{U}_j$ ($j = 0, 1, 2$). As before, we write $|\boldsymbol{\ell}| = \sum_{\mathfrak{u}} \ell_{\mathfrak{u}}$ and $|\mathbf{q}| = \sum_{\mathfrak{u}} q_{\mathfrak{u}}$. For $\boldsymbol{\ell} \in \mathbf{N}^{\mathscr{U}}$, let $\Omega_{\boldsymbol{\ell}}$ denote the collection of admissible graphs for the product $\prod_{\mathfrak{u} \in \mathscr{U}} X_{\mathfrak{u}}^{\diamond(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})}$, which is consistent with the notation in the previous section since $\mathcal{S} = \mathcal{I} = \varnothing$. Since there are at most $(2|\boldsymbol{\ell}| + |\mathbf{q}| - 1)!!$ graphs in $\Omega_{\boldsymbol{\ell}}$, we can further control (6.29) by

$$|\mathbf{E}\partial_{\boldsymbol{\theta}}^{\mathbf{r}}(\boldsymbol{\theta}, \mathbf{X})|$$

$$(6.30) \qquad \leq \sum_{\boldsymbol{\ell} \in \mathbf{N}^{\mathscr{U}}} \Bigg(\frac{C^{2|\boldsymbol{\ell}| + |\mathbf{q}|}}{(2\boldsymbol{\ell} + \mathbf{q})!} \cdot (2|\boldsymbol{\ell}| + |\mathbf{q}| - 1)!! \cdot \max_{\Gamma \in \Omega_{\boldsymbol{\ell}}} (\theta^{2|\boldsymbol{\ell}| + |\mathbf{q}|} |\Gamma|)\Bigg),$$

where we have used the shorthand notation $(2\boldsymbol{\ell} + \mathbf{q})! = \prod_{\mathfrak{u}}(2\ell_{\mathfrak{u}} + q_{\mathfrak{u}})!$. Applying Proposition 6.20 to control the term $\theta^{2|\boldsymbol{\ell}| + |\mathbf{q}|} |\Gamma|$ and using the multinomial theorem for the sum over $\boldsymbol{\ell}$, we get

$$(6.31) \qquad |\mathbf{E}\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \Phi(\boldsymbol{\theta}, \mathbf{X})| \leq C \max_{\Gamma \in \Omega^*} (\theta^{\deg(\Gamma)} |\Gamma|).$$

We want to show that the right-hand side above can be controlled by the right-hand side of (6.6). In order to show this, we note first that since we are in the case $\mathcal{S} = \mathcal{I} = \varnothing$, the nodes of the graphs $\Gamma \in \Omega^*$ are indexed by elements of $\mathscr{U}$. Furthermore, these graphs $\Gamma$ are such that nodes indexed by $\mathfrak{u} \in \mathscr{U}_j$ have degree $j$, with the exception of nodes in $\mathscr{U}_0$ that can have degree either 0 or 2. In other words, there exists a (unique) $A \subset \mathscr{U}_0$ such that $\Gamma$ is admissible for the product

$$(6.32) \qquad \left( \prod_{\mathfrak{u} \in \mathscr{U}_1} X_{\mathfrak{u}} \right) \cdot \left( \prod_{\mathfrak{u} \in \mathscr{U}_2 \cup A} X_{\mathfrak{u}}^{\diamond 2} \right).$$

We then show that for any $\Gamma \in \Omega^*$ there exists a graph $\mathrm{Enh}(\Gamma)$ with nodes indexed by $\mathcal{T}$ such that $\mathrm{Enh}(\Gamma)$ is admissible for the product appearing on the right-hand side of (6.6) and such that

$$(6.33) \qquad |\Gamma| \lesssim \theta^{\deg(\mathrm{Enh}(\Gamma)) - \deg \Gamma} |\mathrm{Enh}(\Gamma)|.$$

Since $\deg(\mathrm{Enh}(\Gamma)) = |\mathcal{O}| + 2|\mathcal{E}|$ by the definition of admissibility, the claimed bound then follows immediately from the fact that the covariances of the random variables $X_t$ are all positive.

As previously, $\mathrm{Enh}(\Gamma)$ is built iteratively and it suffices to show that (6.33) holds at each step of the iteration. At each step, one of the nodes of $\Gamma$ indexed by $\mathfrak{u}$ is replaced by a collection of nodes indexed by the elements $u \in \mathfrak{u}$. We again describe our enhancement procedure in graphic notation. We use $\circ$ to represent a point in $\mathcal{O}$, and $\bullet$ to represent a point in $\mathcal{E}$. A point $\circ$ in a grey area represents a point outside the cluster under consideration, and the parity of that point does not matter. Also, a grey area is not necessarily a cluster—it just means an area outside the cluster in consideration. In particular, two $\circ$'s drawn in the same grey area may belong to different clusters.

For any cluster $\mathfrak{u} \in \mathscr{U}$, we write $\mathcal{O}_{\mathfrak{u}} = \mathcal{O} \cap \mathfrak{u}$ and $\mathcal{E}_{\mathfrak{u}} = \mathcal{E} \cap \mathfrak{u}$.[8] We also let $u^* = u^*(\mathfrak{u})$ denote the representative point of $\mathfrak{u}$. Fix an arbitrary $\Gamma \in \Omega^*$, and let $A \subset \mathscr{U}_0$ be the set in (6.32). For this $\Gamma$, we have $\deg(u^*(\mathfrak{u})) = 1$ for every $\mathfrak{u} \in \mathscr{U}_1$, $\deg(u^*(\mathfrak{u})) = 2$ for every $\mathfrak{u} \in \mathscr{U}_2 \cup A$, and 0 otherwise. We now iterate over all clusters and construct $\mathrm{Enh}(\Gamma)$ in the following way:

1. If $\mathfrak{u} \in \mathscr{U}_2$, then $\mathfrak{u}$ consists of a single point $u = u^*$ with $\deg(u) = \deg(\mathfrak{u}) = 2$, and we do not change anything, except for relabeling $\mathfrak{u}$ by $u$.

2. For $\mathfrak{u} \in \mathscr{U}_1$, $u^*$ belongs to $\mathcal{O}$. By definition of $\Omega^*$, $\deg(u^*) = 1$ and this edge connects to a point outside $\mathfrak{u}$. Then, depending on whether $|\mathcal{E}_{\mathfrak{u}}| = 1$ or not, we do

---

[8]These are different notation from $\mathcal{O}_j$ and $\mathcal{E}_j$ defined at the beginning of the section.

the following operations:



(6.34)

When $|\mathcal{E}_{\mathfrak{u}}| \neq 1$, we pair all points in $\mathfrak{u} \cap \mathcal{O}$ except $u^*$ (this is possible since there are an even number of them), and link all points in $\mathfrak{u} \cap \mathcal{E}$ cyclically. When $|\mathcal{E}_{\mathfrak{u}}| = 1$, we connect $u^*$ with the unique even point, then move the previously existing edge attached to $u^*$ to that even point as well, and finally pair all the remaining "odd" points. The corresponding bounds are immediate from the fact that the covariance between points belonging to the same cluster is greater than $\theta^{-2}$ by Lemma 4.1, as well as Lemma 4.2 which allows us to move endpoints of edges within a given cluster at the cost of some fixed multiplicative constant.
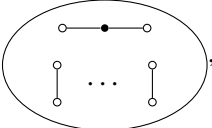
3. For $\mathfrak{u} \in A$, the situation is more complicated. By (6.32), $u^*$ has two edges and both of them are connected to points outside $\mathfrak{u}$. There are four possibilities depending on whether $u^* \in \mathcal{O}$ or $\mathcal{E}$, and whether $|\mathcal{E}_{\mathfrak{u}}| = 1$ or not. If $u^* \in \mathcal{O}$ and $|\mathcal{E}_{\mathfrak{u}}| = 1$, we perform the operation



(6.35)

Here, we have moved the two edges from $u$ to the "even" point, and paired all the odd points (since $A \subset \mathcal{U}_0$, so every cluster in $A$ contains an even number of them). For the other three situations, we can "move edges around" in a similar way, pair all the odd points, and cyclically connect the even ones as in one of the situations from (6.34). In this way, we get bounds with the same power of $\theta$ as (6.35).

4. For $\mathfrak{u} \in \mathscr{U}_0 \setminus A$, all the points in this cluster have degree 0 and $|\mathcal{O}_\mathfrak{u}|$ is even, so we have

(6.36)
$$1 \leq C\theta^{|\mathcal{O}_\mathfrak{u}|+2|\mathcal{E}_\mathfrak{u}|} \left( \underset{\cdots}{\circ\!-\!\bullet\!-\!\circ} \right), \qquad |\mathcal{E}_\mathfrak{u}| = 1.$$

The bound for the case $|\mathcal{E}_\mathfrak{u}| \neq 1$ is the same.

Iterating over all $\mathfrak{u} \in \mathscr{U}$, we obtain indeed a graph with nodes indexed by $\mathcal{T}$ and such that all nodes in $\mathcal{O}$ have degree 1, while nodes in $\mathcal{E}$ have degree 2, so that it is admissible for the right-hand side of (6.6) as required, and such that (6.33) holds, thus concluding the proof of (6.6).

We now briefly explain how one can prove (6.7). The argument would be a simple modification of that for (6.6). We first note that after the same clustering and backward reduction procedure, we arrive at the bound of type (6.31), with its left-hand side replaced by that in (6.7). However, the difference here is that $\Omega^*$ in this case contains different set of graphs than those admissible for (6.32).

In fact, since the cosines are not recentered, the Wick power series for those points in $\mathscr{U}_2$ starts at 0th order rather than 2nd. As a consequence, for the $q_\mathfrak{u}$ in (6.21), we will have $q_\mathfrak{u} = 0$ for $\mathfrak{u} \in \mathscr{U}_2$. Hence, it is natural in this case to define $\Omega_{\mathbf{k},\boldsymbol{\ell}}$ as before, but setting now $q_\mathfrak{u} = 0$ for $\mathfrak{u} \in \mathscr{U}_2$ in (6.24). Besides this change in the definition of $\Omega_{\mathbf{k},\boldsymbol{\ell}}$, we keep Definition 6.17 for $\Omega^*$. By this definition, we see that $\Omega^*$ in this case simply consists of graphs admissible for the product $\prod_{\mathfrak{u} \in \mathscr{U}_1} X_\mathfrak{u}$.

Then, in the enhancement procedure for $\Gamma \in \Omega^*$, we only pair the remaining points in $\mathcal{O}$ within each cluster but do not add any edges to points in $\mathcal{E}$. This will produce a graph admissible for the right-hand side of (6.7), together with the correct power of $\theta$.

REMARK 6.21. It transpires from the proof that if we had considered the Wiener chaos expansion of $\Phi(\boldsymbol{\theta}, \mathbf{X})$ directly without clustering, then the dependence on $\theta$ we would have obtained on the right-hand side of (6.26) would have been $\theta^{\deg(\Gamma)}$ rather than $\theta^{\deg(\Gamma^*)}$. Since $\deg(\Gamma) \geq 2(|\mathbf{k}| + |\boldsymbol{\ell}|)$ for $\Gamma \in \Omega_{\mathbf{k},\boldsymbol{\ell}}$, the $\theta$-dependence in Theorem 6.4 would be $e^{C\theta^2}$ for some large $C$ depending on $K$ and $|\mathcal{T}|$, and the same would be true for Theorem 6.2. In this case, we would have had to require $\widehat{F}$ to decay faster than Gaussian in order to ensure the convergence of the stochastic objects.

6.8. *Simplifying the interim bound for $\mathcal{G}$.* The bound (6.6) provides us with finer controls on the coefficients $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s+\mathbf{m}_s}^{(s)}$ and $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}$ in (6.21). Recall the definitions of $\mathcal{O}_j$ and $\mathcal{E}_j$ from (6.2). For $s \in \mathcal{S}_{\max}$, $i \in \mathcal{I}$, and $j = 1, 2$, we let $\mathcal{O}_j^s = \mathcal{O}_j(\mathbf{m}_s)$ and $\mathcal{O}_j^i = \mathcal{O}_j(\mathbf{p}_i)$. We define $\mathcal{E}_j^s$ and $\mathcal{E}_j^i$ similarly for $j = 0, 1, 2$. We have the following lemma on the controls of the coefficients.

LEMMA 6.22. *The coefficients* $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s+\mathbf{m}_s}^{(s)}$ *and* $\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}$ *satisfy the bound*

$$|\partial_{\boldsymbol{\theta}}^{\mathbf{r}} \bar{C}_{2\mathbf{k}_s+\mathbf{m}_s}^{(s)}|$$

$$\leq C\theta^{2|\mathcal{T}|}$$

$$\times \frac{(C\theta)^{|2\mathbf{k}_s+\mathbf{m}_s|}}{(2\mathbf{k}_s+\mathbf{m}_s)!} \cdot \mathbf{E}\Big[\Big(\prod_{t\in\mathcal{O}_2^s\cup\mathcal{E}_1^s} X_t\Big)\Big(\prod_{t\in\mathcal{O}_1^s\cup\mathcal{E}_2^s}(1+X_t^{\diamond 2})\Big)\Big(\prod_{t\in\mathcal{E}_0^s} X_t^{\diamond 2}\Big)\Big],$$

$$|\partial_{\boldsymbol{\theta}}^{\mathbf{r}} C_{\mathbf{p}_i}^{(i)}|$$

$$\leq C\theta^{2|\mathcal{T}|}\theta^{|\mathbf{p}_i|}\mathbf{E}\Big[\Big(\prod_{t\in\mathcal{O}_2^i\cup\mathcal{E}_1^i} X_t\Big)\Big(\prod_{t\in\mathcal{O}_1^i\cup\mathcal{E}_2^i}(1+X_t^{\diamond 2})\Big)\Big(\prod_{t\in\mathcal{E}_0^i} X_t^{\diamond 2}\Big)\Big].$$

PROOF. In the expression (6.8), for $k \in \mathcal{S} \cup \mathcal{I}$ and $t \in \mathcal{O}_1^k \cup \mathcal{E}_2^k$, we write

$$\cos(\theta_t X_t) = \{\!\!\{\cos(\theta_t X_t)\}\!\!\} + \mathbf{E}\cos(\theta_t X_t).$$

We then expand the product and apply (6.6) to each term in the sum. The lemma then follows immediately. Note that for the sum in the first term, we have $\mathcal{O}_j^s = \mathcal{O}_j(2\mathbf{k}_s + \mathbf{m}_s)$ and the same for $\mathcal{E}_j^s$. $\square$

We can now further simplify the bound for $\mathcal{G}$ given in Proposition 6.12. The statement is as follows.

PROPOSITION 6.23. *Fix* $\mathcal{M} \subset \mathbf{N}^{\mathcal{T}}$ *as well as a root set* $\mathcal{R}$ *satisfying Assumption* 6.1. *Let* $\mathcal{S}, \mathcal{I}, \mathcal{J}$ *be any disjoint subsets of* $[K]$ *such that their union equals* $[K]$. *Fix* $\mathbf{m} \in \mathcal{M}^{\mathcal{S}}$ *and* $\mathbf{p} \in \mathcal{M}^{\mathcal{I}}$. *Let* $\Omega^*$ *be the collection of graphs characterised by Definition* 6.17 *with the vertex set* $\mathbb{V}$ *in* (6.23). *Then we have*

$$|\mathcal{G}(\boldsymbol{\theta}, \mathbf{X})| \leq C\theta^{2|\mathbf{p}|+2|\mathcal{T}|(|\mathcal{S}|+|\mathcal{I}|)+\deg(\Omega^*)}\Big(\sum_{\Gamma^*\in\Omega^*} |\Gamma^*|\Big)$$

$$(6.37) \qquad \times \Big[\prod_{k\in\mathcal{S}\cup\mathcal{I}} \mathbf{E}\Big(\Big(\prod_{t\in\mathcal{O}_2^k\cup\mathcal{E}_1^k} X_t\Big)\Big(\prod_{t\in\mathcal{O}_1^k\cup\mathcal{E}_2^k}(1+X_t^{\diamond 2})\Big)\Big(\prod_{t\in\mathcal{E}_0^k} X_t^{\diamond 2}\Big)\Big)\Big],$$

*where* $\deg(\Omega^*) = \max_{\Gamma^*\in\Omega^*} \deg(\Gamma^*)$. *The constant* $C$ *depends on* $|\mathcal{T}|$, $K$, $\mathbf{r}$ *and* $\Lambda$ *only.*

PROOF. Applying Lemma 6.22 to (6.21), we get

$$|\mathcal{G}(\boldsymbol{\theta}, \mathbf{X})| \leq C\theta^{2|\mathbf{p}|+2|\mathcal{T}|(|\mathcal{S}|+|\mathcal{I}|)}\Big[\sum_{\mathbf{k},\boldsymbol{\ell}}\Big(\frac{C^{|2\mathbf{k}+\mathbf{m}|+|2\boldsymbol{\ell}+\mathbf{q}|}}{(2\mathbf{k}+\mathbf{m})!(2\boldsymbol{\ell}+\mathbf{q})!} \cdot \sum_{\Gamma\in\Omega_{\mathbf{k},\boldsymbol{\ell}}} \theta^{\deg(\Gamma)}|\Gamma|\Big)\Big]$$

$$\times \Big[\prod_{k\in\mathcal{S}\cup\mathcal{I}} \mathbf{E}\Big(\Big(\prod_{t\in\mathcal{O}_2^k\cup\mathcal{E}_1^k} X_t\Big)\Big(\prod_{t\in\mathcal{O}_1^k\cup\mathcal{E}_2^k}(1+X_t^{\diamond 2})\Big)\Big(\prod_{t\in\mathcal{E}_0^k} X_t^{\diamond 2}\Big)\Big)\Big],$$

where the range of the sum of $(\mathbf{k}, \boldsymbol{\ell})$ is the same as that in (6.21). Note that the factor $\theta^{2|\mathbf{p}|}$ comes from the bound (6.21), while the other $|2\mathbf{k} + \mathbf{m}| + |\mathbf{p}|$ powers of $\theta$ from Lemma 6.22 are in the term $\theta^{\deg(\Gamma)}$. By Proposition 6.20 and the multinomial theorem, we can control the term in the bracket in the first line by $\sum_{\Gamma^* \in \Omega^*} \theta^{\deg(\Gamma^*)} |\Gamma^*|$. The claim then follows. $\quad\square$

6.9. *Enhancement and proof of Theorem* 6.2. We now start to develop the final ingredients to prove Theorem 6.2. From now on, we take $\mathcal{S} = \mathcal{S}_{\max}$. We also assume $K \geq 2$ for otherwise the bound (6.5) is either trivial ($\mathcal{M} \neq \varnothing$) or implied by (6.6) ($\mathcal{M} = \varnothing$).

Let the vertex set $\mathbb{V}$ be the whole collection of $K|\mathcal{T}|$ points $\{(k, t)\}_{k \in [K], t \in \mathcal{T}}$. Fix any $\Gamma^* \in \Omega^*$ as given by Definition 6.17. For every $k \in \mathcal{S} \cup \mathcal{I}$, we also fix an arbitrary graph $\Gamma^{(k)}$ that is admissible for the product[9]

$$(6.38) \qquad \left( \prod_{t \in \mathcal{O}_2^k \cup \mathcal{E}_1^k} X_t \right) \left( \prod_{t \in \mathcal{O}_1^k \cup \mathcal{E}_2^k} (1 + X_t^{\diamond 2}) \right) \left( \prod_{t \in \mathcal{E}_0^k} X_t^{\diamond 2} \right).$$

Let

$$(6.39) \qquad \bar{\Gamma} = \left( \bigcup_{k \in \mathcal{S}_{\max} \cup \mathcal{I}} \Gamma^{(k)} \right) \cup \Gamma^*.$$

Note that some of the points in $\bar{\Gamma}$ may have degree 0. In view of (6.37), it suffices to control $|\bar{\Gamma}|$, possibly with a few extra powers of $\theta$, by the right-hand side of (6.5). To do this, we perform an enhancement procedure to $\bar{\Gamma}$ similar to the one in Section 6.7. The enhanced graph $\overline{\mathrm{Enh}}(\bar{\Gamma})$ will need to control $|\bar{\Gamma}|$ in its value and at the same time also match a subset of the terms on the right-hand side of (6.5). This will immediately imply (6.5).

Before we do the enhancement, let us first check the parities of the degrees of vertices in $\bar{\Gamma}$. For every $(s, t) \in \mathcal{S}_{\max} \times \mathcal{T}$, the contribution to $\deg(s, t)$ comes from both $\Gamma^{(s)}$ and $\Gamma^*$. By (6.38) and the definition of $\Omega^*$, we have

$$\deg(s, t) = m_s^t + \mathbf{1}_{\{t \in \mathcal{O}_2(\mathbf{m}_s) \cup \mathcal{E}_1(\mathbf{m}_s)\}} \qquad (\mathrm{mod}\ 2).$$

It is straightforward to check that $\deg(s, t)$ is odd if and only if $t \in \mathcal{O}$. The same is true for $(i, t) \in \mathcal{I} \times \mathcal{T}$ (just replace $\mathbf{m}_s$ by $\mathbf{p}_i$) so that, for every $u \in (\mathcal{S}_{\max} \cup \mathcal{I}) \times \mathcal{T}$, $\deg(u)$ is odd if $u \in (\mathcal{S}_{\max} \cup \mathcal{I}) \times \mathcal{O}$ and even if $u \in (\mathcal{S}_{\max} \cup \mathcal{I}) \times \mathcal{E}$.

We now turn to the points in $\mathcal{J} \times \mathcal{T}$. The only contribution to the degrees of these points are from $\Gamma^*$. For every $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$, by the choice of the representative point $u^* \in \mathfrak{u}$ described in Section 6.7, we see that $\deg(u^*)$ is odd if and only if

---

[9]Here and below, when we say a graph is admissible for a product where some factors contain a sum of Wick powers, we mean that the graph is admissible for one of the terms in the expansion of the product.

M. HAIRER AND W. XU

$\mathfrak{u} \in \mathscr{U}_1^{\mathcal{J}}$. Then $\deg(j, t)$ is even if $(j, t) \in \mathcal{J} \times \mathcal{E}$, and in particular is 0 if it is not a representative point for any $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$.

Hence, the only points whose parity of degree are inconsistent with their types are those in $(\mathcal{J} \times \mathcal{O}) \setminus \{u(\mathfrak{u})\}_{\mathfrak{u} \in \mathscr{U}_1^{\mathcal{J}}}$. To keep track of these points, we let

$$\mathcal{V}_\mathfrak{u} = \{v \in \mathfrak{u} \cap (\mathcal{J} \times \mathcal{O}) : \deg(v) \text{ is even}\}.$$

One can check that $\mathcal{V}_\mathfrak{u} = \mathfrak{u} \cap (\mathcal{J} \times \mathcal{O})$ for $\mathfrak{u} \in \mathscr{U}_0^{\mathcal{J}}$, $\mathcal{V}_\mathfrak{u} = (\mathfrak{u} \cap (\mathcal{J} \times \mathcal{O})) \setminus \{u^*(\mathfrak{u})\}$ for $\mathfrak{u} \in \mathscr{U}_1^{\mathcal{J}}$, and is empty for $\mathfrak{u} \in \mathscr{U}_2^{\mathcal{J}}$. In particular, $|\mathcal{V}_\mathfrak{u}|$ is even for every $\mathfrak{u} \in \mathscr{U}^{\mathcal{J}}$. Hence, we perform the following operation to $\bar{\Gamma}$:

$$(6.40) \qquad 1 \le C\theta^{|\mathcal{V}_\mathfrak{u}|} \left( \begin{array}{ccc} \circ & & \circ \\ | & \cdots & | \\ \circ & & \circ \end{array} \right) \qquad \forall \mathfrak{u} \in \mathscr{U}^{\mathcal{J}},$$

where we paired the points in $\mathcal{V}_\mathfrak{u}$ for every $\mathfrak{u}$ and linked each pair with an edge of multiplicity 1. Note that this operation is performed only to points in $\mathcal{V}_\mathfrak{u}$ but not all of $\mathfrak{u}$, so the notion $\circ$ for odd points is justified. We do this operation for every $\mathfrak{u}$. Hence, all the points in the set $\{v \in \mathcal{V}_\mathfrak{u}, \mathfrak{u} \in \mathscr{U}^{\mathcal{J}}\}$ are affected in this operation, which adds exactly one degree to each of these points. After this procedure, $\deg(v)$ is odd for every $v \in \mathcal{J} \times \mathcal{O}$.

Thus, at this stage, all points have parities consistent with their types: $\deg(u)$ is odd if $u \in [K] \times \mathcal{O}$ and is even if $u \in [K] \times \mathcal{E}$. Since we will not make use of the choice of representative points any more, we use $u$ to denote generic points rather than just the representative ones chosen before.

We now turn to the clusters in $\mathscr{C}(\mathcal{S}_{\max}^c \times \mathcal{T}) = \mathscr{C}((\mathcal{I} \cup \mathcal{J}) \times \mathcal{T})$ (not restricting to points in $\mathcal{J} \times \mathcal{T}$ any more). For every cluster $\mathcal{W}$ in this collection, if $|\mathcal{W}| = 1$, then we do not change anything to that cluster. If $|\mathcal{W}| \ge 2$, then for $\bar{N} > \max_{\mathbf{p} \in \mathcal{M}} |\mathbf{p}|$, we cyclically connect *all* the points in $\mathcal{W}$ with edges of multiplicity $\bar{N}$ in addition to the edges and multiplicities they already have. The effect of the operation can be graphically described by

$$(6.41) \qquad 1 \le C\theta^{2\bar{N}|\mathcal{W}|} \quad \vcenter{\hbox{}} \quad, \qquad |\mathcal{W}| \ge 2,$$

where all the points on the right-hand side are in the same cluster $\mathcal{W}$, and "$+\bar{N}$" refers to that the integer $\bar{N}$ is added in addition to the multiplicity that might already exist. The exact order of the points in the cycle is arbitrary, and the notion $\circ$

refers to a generic point in $\mathcal{W}$ whose parity does not matter. We perform this operation for every $\mathcal{W} \in \mathscr{C}(\mathcal{S}_{\max}^c \times \mathcal{T})$ with $|\mathcal{W}| \geq 2$. This operation adds $2\bar{N}$ degrees to every point in $(\mathcal{I} \cup \mathcal{J}) \times \mathcal{T}$ that is not a singleton, so the parities of the degrees do not change. The graph obtained at the end of this procedure is our desired enhanced graph, which we denote by $\overline{\mathrm{Enh}}(\bar{\Gamma})$.

We now show that $\overline{\mathrm{Enh}}(\bar{\Gamma})$ does satisfy a number of properties that will allow us to deduce Theorem 6.2 from this construction. In order to precisely describe those properties, for $(k, t) \in [K] \times \mathcal{T}$, we define the external degrees of $(s, t)$ and $k$ by

$$\mathrm{deg}_{\mathrm{ex}}(k, t) = \sum_{u \notin \{k\} \times \mathcal{T}} \mathbb{E}((k, t), u), \qquad \mathbf{deg}_{\mathrm{ex}}(k) = (\mathrm{deg}_{\mathrm{ex}}(k, t))_{t \in \mathcal{T}},$$

with $|\mathbf{deg}(k)|$ being the sum of its components. We now summarise the properties of $\overline{\mathrm{Enh}}(\bar{\Gamma})$ in the following proposition.

PROPOSITION 6.24. *Let $\overline{\mathrm{Enh}}(\bar{\Gamma})$ be the graph obtained from $\bar{\Gamma}$ as above. Then we have the bound*

(6.42) $$|\bar{\Gamma}| \leq C\theta^{(2\bar{N}+1)K|\mathcal{T}|}|\overline{\mathrm{Enh}}(\bar{\Gamma})|.$$

*Furthermore, $\overline{\mathrm{Enh}}(\bar{\Gamma})$ has the following properties*:

1. *One has $\mathrm{deg}(u) \geq 1$ for all $u \in [K] \times \mathcal{T}$. Furthermore, $\mathrm{deg}(u)$ is odd if $u \in [K] \times \mathcal{O}$ and even if $u \in [K] \times \mathcal{E}$.*
2. *For every $s \in \mathcal{S}_{\max}$, there exists $\mathbf{k}_s \in \mathbf{N}_{\mathcal{E}_0(\mathbf{m}_s)}^{\mathcal{T}}$ such that $\mathrm{deg}_{\mathrm{ex}}(s) = 2\mathbf{k}_s + \mathbf{m}_s$.*
3. *For every $k \in \mathcal{S}_{\max}^c$, there exist $k' \in \mathcal{S}_{\max}^c \setminus \{k\}$ and $(t, t') \in \mathcal{T} \times \mathcal{T}$ such that $\mathbb{E}((k, t), (k', t')) \geq \bar{N}$.*

REMARK 6.25. The definition of $\mathcal{S}_{\max}$ excludes the possibility that $|\mathcal{S}_{\max}| = K - 1$. Hence, if $\mathcal{S}_{\max}^c$ is not empty, then it has at least two points, so the third property above at least makes sense. If $\mathcal{S}_{\max} = \varnothing$, then Property 2 is automatically true. On the other hand, if $\mathcal{S}_{\max} = [K]$, then Property 3 is automatically true.

PROOF OF PROPOSITION 6.24. The bound (6.42) follows from the bounds (6.40) and (6.41) and that the total degree of $\theta$ is bounded by

$$\sum_{u \in \mathscr{U}^{\mathcal{J}}} |\mathcal{V}_u| + 2\bar{N} \sum_{\mathcal{W} \in \mathscr{C}(\mathcal{S}_{\max}^c \times \mathcal{T})} |\mathcal{W}| \leq (2\bar{N} + 1)K|\mathcal{T}|.$$

We now turn to the three properties. For the first one, the positivity of the degrees can be checked as follows. For $s \in \mathcal{S}_{\max}$, since the enhancement procedure does not change anything to points in $\mathcal{S}_{\max} \times \mathcal{T}$, $\mathrm{deg}(s, t)$ consists precisely the contributions from the original graphs $\Gamma^{(s)}$ and $\Gamma^*$. By (6.38) and the definition of $\Omega^*$, we have

$$\mathrm{deg}(s, t) \geq 2k_s^t + m_s^t + \mathbf{1}_{\{t \in \mathcal{O}_2^s \cup \mathcal{E}_1^s\}} + 2 \cdot \mathbf{1}_{\{t \in \mathcal{E}_0^s\}}.$$

If $m_s^t \geq 1$, then there is nothing to prove. If $m_s^t = 0$, then we have $t \in \mathcal{E}_0^s = \mathcal{E}_0(\mathbf{m}_s)$, and hence $\deg(s, t) \geq 2$. The same is true for points in $\mathcal{I} \times \mathcal{T}$ by noting that the enhancement procedure (6.41) can only add $2\bar{N}$ degrees to those points. For $u \in \mathcal{J} \times \mathcal{T}$, there are two possibilities. If $\{u\}$ is a singleton (that is, itself alone is a cluster), then $u$ is the representative point for some $\mathsf{u} \in \mathscr{U}_1^{\mathcal{J}} \cup \mathscr{U}_2^{\mathcal{J}}$. Hence, $\deg(u) \geq 1$. If $\{u\}$ is not singleton, the $u$ belongs to some $\mathcal{W} \in \mathscr{C}(\mathcal{S}_{\max}^c \times \mathcal{T})$ with $|\mathcal{W}| \geq 2$. By the enhancement (6.41), we have $\deg(u) \geq \bar{N}$. This shows that all points in $\overline{\mathrm{Enh}}(\bar{\Gamma})$ have strictly positive degrees.

As for the consistency of the parities, we have shown this property after Procedure (6.40), and that it remains unchanged under Procedure (6.41). Thus, Property 1 is true for $\overline{\mathrm{Enh}}(\bar{\Gamma})$.

For Property 2, note that the enhancement procedures above do not affect any point in $\mathcal{S}_{\max} \times \mathcal{T}$, and the graph $\Gamma^{(s)}$ only contributes to internal degrees of $\{(s, t)\}_{t \in \mathcal{T}}$ (total multiplicities of edges linking to the collection of points in $\{(s, t)\}_{t \in \mathcal{T}}$). Thus, all contributions to $\mathbf{deg}_{\mathrm{ex}}(s)$ come from $\Gamma^*$. The property then follows immediately from the first constraint in Definition 6.17.

Property 3 is the only one which we use $\mathcal{S} = \mathcal{S}_{\max}$. By definition of $\mathcal{S}_{\max}$, for every $k \in \mathcal{S}_{\max}^c$, there exists a cluster $\mathcal{W} \in \mathscr{C}(\mathcal{S}_{\max}^c \times \mathcal{T})$ which contains at least one point from both $\{k\} \times \mathcal{T}$ and $(\mathcal{S}_{\max}^c \setminus \{k\}) \times \mathcal{T}$. For that $\mathcal{W}$, since it contains at least two points, so it has gone through the operation 6.41. In the cyclic ordering in that operation, there exists $t \in \mathcal{T}$ such that $(k, t) \in \mathcal{W}$ and is adjacent to some $(k', t')$ in the cycle for some $k' \neq k$. This implies $\mathbb{E}((k, t), (k', t')) \geq \bar{N}$. $\quad\square$

PROOF OF THEOREM 6.2. We are now ready to prove the main theorem. By the bounds (6.37) and (6.42), we know that up to some power of $\theta$, $|\mathcal{G}(\boldsymbol{\theta}, \mathbf{X})|$, and hence the left-hand side of (6.5) is controlled by $\sum |\overline{\mathrm{Enh}}(\bar{\Gamma})|$ where the sum is taken over all possible graphs $\bar{\Gamma}$ obtained from (6.39). It then suffices to show that each $\overline{\mathrm{Enh}}(\bar{\Gamma})$ is admissible for the product on the right-hand side of (6.5) for large enough $N$. This will imply $|\overline{\mathrm{Enh}}(\bar{\Gamma})|$ is controlled by that right-hand side, and hence completes the proof of the theorem.

To show the admissibility, we match the three properties of $\overline{\mathrm{Enh}}(\bar{\Gamma})$ in Proposition 6.24 with those of the admissible graphs to the right-hand side of (6.5). The collection of graphs admissible for the product

$$\prod_{k=1}^{K} \mathscr{T}_{\mathcal{M}} \left( \prod_{t \in \mathcal{O}} P_N(X_{k,t}) \prod_{t \in \mathcal{E}} Q_N(X_{k,t}) \right)$$

can be completely characterised as follows:

1. $\deg(u) \in \{1, 3, \ldots, 2N - 1\}$ if $u \in [K] \times \mathcal{O}$, and $\deg(u) \in \{2, 4, \ldots, 2N\}$ if $u \in [K] \times \mathcal{E}$.
2. For every $k \in [K]$, $\mathbf{deg}_{\mathrm{ex}}(k) \notin \mathcal{M}$.

For the first criterion, the consistency of parity and strict positivity of the degrees is given directly by Property 1 in Proposition 6.24. Also, the degrees are bounded by $2N - 1$ and $2N$ if $N$ is large enough.

We now turn to the second one. For $s \in \mathcal{S}_{\max}$, by Property 2, we have $\mathbf{deg}_{\mathrm{ex}}(s) = 2\mathbf{k}_s + \mathbf{m}_s \in \mathcal{B}(\mathbf{m}_s)$. Since $\mathbf{m}_s \in \mathcal{R} \subset \mathcal{M}^c$, the assumption on $\mathcal{M}$ explicitly states that $\mathcal{B}(\mathbf{m}_s) \subset \mathcal{M}^c$, and hence $\mathbf{deg}_{\mathrm{ex}}(s) \notin \mathcal{M}$. For $k \in \mathcal{S}_{\max}^c$, Property 3 implies that $|\mathbf{deg}_{\mathrm{ex}}(k)| \geq \bar{N} > \max_{\mathbf{p} \in \mathcal{M}} |\mathbf{p}|$. Hence, we also have $\mathbf{deg}_{\mathrm{ex}}(k) \notin \mathcal{M}$ for $k \in \mathcal{S}_{\max}^c$.

Finally, the degree of the power in $\theta$ is also bounded by large enough $N$. This completes the proof. $\quad\square$

REMARK 6.26. If we carefully keep track of the powers in $\theta$ in all of the above arguments, we see that we can take $N = CK$, where $C$ depends on $\mathcal{T}, \mathcal{M}$ and $\mathcal{R}$ only.

## APPENDIX: THE STANDARD KPZ MODEL

For the sake of completeness, we briefly describe the standard KPZ model and its approximations. These approximations correspond to the renormalised model $\widehat{\Pi}^\varepsilon$ defined in Section 3.1 in the special case when $F(u) = au^2$. We denote it by $\Pi^{\mathrm{KPZ};\varepsilon}$. Note that since those models are normalised, every $a \neq 0$ gives rise to the *same* model.

One of the main results in [21] and [25] says that $\Pi^{\mathrm{KPZ};\varepsilon}$ converges to a limit model, which we denote by $\Pi^{\mathrm{KPZ}}$. It is called the standard KPZ model. By stationarity, we can take the base point to be 0 in all cases. Recall the collection of symbols in Table 3.1. For $\square$ and $\overset{\bullet\,\bullet}{\sqcup}$, we set

$$\Pi_0^{\mathrm{KPZ};\varepsilon}\square = \Pi_0^{\mathrm{KPZ};\varepsilon}\overset{\bullet\,\bullet}{\sqcup} \equiv 0.$$

For the symbols $\circ$, $\bullet$, $\overset{\circ}{\underset{\circ}{}}$, $\overset{\bullet}{\underset{\circ}{}}$, $\overset{\bullet}{\vee}$ and $\overset{\bullet}{\underset{\circ\circ}{}}$ and using the same graphical notation as in [25], we decompose them into different homogeneous chaos so that

$$\langle \Pi_0^{\mathrm{KPZ};\varepsilon} \raisebox{-2pt}{\includegraphics{dummy}}, \varphi_0^\lambda \rangle = \quad + \quad - \quad + 2 \quad + 2 \quad - 2 \quad .$$
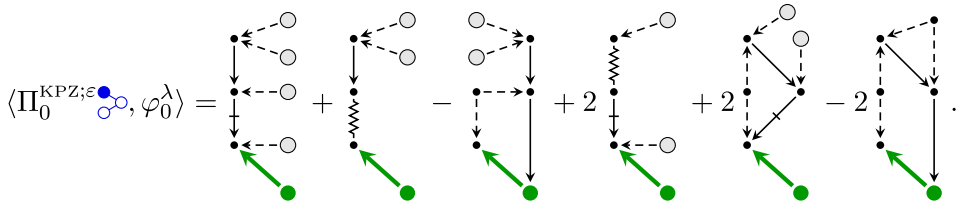
The following result shown in [14, 25] (but see also [21] for an earlier similar result in a slightly different context) is the main convergence result for the KPZ model/equation.

THEOREM A.1. *The model* $\Pi^{\mathrm{KPZ};\varepsilon}$ *converges to a limiting model* $\Pi^{\mathrm{KPZ}}$, *which is the KPZ model. For*

$$\tau \in \{ \circ, \bullet, {}^{\circ}_{\circ}, {}^{\bullet}_{\circ}, \mathbf{\vee}^{\bullet}, {}^{\bullet}_{\circ}{}^{\circ} \},$$

*the object* $\langle \Pi_0^{\mathrm{KPZ}} \tau, \varphi_0^\lambda \rangle$ *is described by the same combination of the trees in* $\langle \Pi_0^{\mathrm{KPZ};\varepsilon} \tau, \varphi_0^\lambda \rangle$ *except that each decorated arrow* $\dashrightarrow$ *is replaced by the plain kernel* $\longrightarrow$. *In addition, there exists some* $\kappa' > 0$ *such that one has the bound*

$$(A.1) \qquad \left( \mathbf{E} | \langle \Pi_0^{\mathrm{KPZ};\varepsilon} \tau - \Pi_0^{\mathrm{KPZ}} \tau, \varphi_0^\lambda \rangle |^{2n} \right)^{\frac{1}{2n}} \lesssim_n \varepsilon^{\kappa'} \lambda^{|\tau| + \kappa'}.$$

## REFERENCES

[1] AMIR, G., CORWIN, I. and QUASTEL, J. (2011). Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions. *Comm. Pure Appl. Math.* **64** 466–537. MRMR2796514

[2] BALÁZS, M., QUASTEL, J. and SEPPÄLÄINEN, T. (2011). Fluctuation exponent of the KPZ/stochastic Burgers equation. *J. Amer. Math. Soc.* **24** 683–708. MRMR2784327

[3] BERTINI, L. and GIACOMIN, G. (1997). Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** 571–607. MRMR1462228

[4] BORODIN, A. and CORWIN, I. (2014). Macdonald processes. *Probab. Theory Related Fields* **158** 225–400. MRMR3152785

[5] BRUNED, Y., HAIRER, M. and ZAMBOTTI, L. (2016). Algebraic renormalisation of regularity structures. Available at arXiv:1610.08468.

[6] CHANDRA, A. and HAIRER, M. (2016). An analytic BPHZ theorem for regularity structures. Available at arXiv:1612.08138.

[7] CORWIN, I. and SHEN, H. (2018). Open ASEP in the weakly asymmetric regime. *Comm. Pure Appl. Math.* **71** 2065–2128.

[8] CORWIN, I., SHEN, H. and TSAI, L.-C. (2018). Fluctuation exponent of the KPZ/stochastic Burgers equation. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 995–1012. MRMR3795074

[9] CORWIN, I. and TSAI, L.-C. (2017). Open ASEP in the weakly asymmetric regime. *Ann. Probab.* **45** 1771–1798. MRMR3650415

[10] DEMBO, A. and TSAI, L.-C. (2016). Weakly asymmetric non-simple exclusion process and the Kardar–Parisi–Zhang equation. *Comm. Math. Phys.* **341** 219–261. MRMR3439226

[11] DIEHL, J., GUBINELLI, M. and PERKOWSKI, N. (2017). The Kardar–Parisi–Zhang equation as scaling limit of weakly asymmetric interacting Brownian motions. *Comm. Math. Phys.* **354** 549–589. MRMR3663617

[12] EDWARDS, S. and WILKINSON, D. (1982). The surface statistics of a granular aggregate. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **381** 17–31.

[13] ERHARD, D. and HAIRER, M. (2017). Discretisation of regularity structures. Available at arXiv:1705.02836.

[14] FRIZ, P. and HAIRER, M. (2014). *A Course on Rough Paths*, *Universitext*. Springer, Cham. MRMR3289027

[15] FURLAN, M. and GUBINELLI, M. (2017). Weak universality for a class of 3D stochastic reaction-diffusion models. Available at arXiv:1708.03118.

[16] GONCALVES, P. and JARA, M. (2014). Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.* **212** 597–644. MRMR3176353

[17] GUBINELLI, M., IMKELLER, P. and PERKOWSKI, N. (2015). Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3** e6. MRMR3406823

[18] GUBINELLI, M. and PERKOWSKI, N. (2016). The Hairer–Quastel universality result at stationarity. In: *Stochastic Analysis on Large Scale Interacting Systems*. RIMS Kôkyûroku Bessatsu, B59, Res. Inst. Math. Sci. (RIMS). MRMR3675927

[19] GUBINELLI, M. and PERKOWSKI, N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269. MRMR3592748

[20] GUBINELLI, M. and PERKOWSKI, N. (2017). Energy solutions of KPZ are unique. *J. Amer. Math. Soc.* **31** 427–471. MRMR3758149

[21] HAIRER, M. (2013). Solving the KPZ equation. *Ann. Math.* **178** 559–664. MRMR3071506

[22] HAIRER, M. (2014). A theory of regularity structures. *Invent. Math.* **198** 269–504. MRMR3274562

[23] HAIRER, M. (2018). Renormalisation of parabolic stochastic PDEs. *Jpn. J. Math.* **13** 187–233. MRMR3855740

[24] HAIRER, M. and PARDOUX, É. (2015). A Wong–Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan* **67** 1551–1604. MRMR3417505

[25] HAIRER, M. and QUASTEL, J. (2018). A class of growth models rescaling to KPZ. *Forum Math. Pi* **6** e3.

[26] HAIRER, M. and SHEN, H. (2017). A central limit theorem for the KPZ equation. *Ann. Probab.* **45** 4167–4221. MRMR3737909

[27] HAIRER, M. and XU, W. (2018). Large scale behaviour of three-dimensional continuous phase coexistence models. *Comm. Pure Appl. Math.* **71** 688–746. MRMR3772400

[28] KARDAR, M., PARISI, G. and ZHANG, Y.-C. (1986). Dynamical scaling of growing interfaces. *Phys. Rev. Lett.* **56** 889–892.

[29] LABBÉ, C. (2017). Weakly asymmetric bridges and the KPZ equation. *Comm. Math. Phys.* **353** 1261–1298. MRMR3652491

[30] MATETSKI, K., QUASTEL, J. and REMENIK, D. (2017). The KPZ fixed point. Available at arXiv:1701.00018.

[31] SHEN, H. and XU, W. (2018). Weak universality of dynamical $\Phi_3^4$: Non-Gaussian noise. *Stoch. Partial Dffer. Equ. Anal. Comput.* **6** 211–254. MRMR3818405

[32] XU, W. (2018). Sharp convergence of nonlinear functionals of a class of Gaussian random fields. *Commun. Math. Stat.* **6** 509–532.

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON
LONDON, SW7 2AZ
UNITED KINGDOM
E-MAIL: m.hairer@imperial.ac.uk

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
OXFORD, OX2 6GG
UNITED KINGDOM
E-MAIL: weijunx@gmail.com