The Annals of Probability
2017, Vol. 45, No. 6B, 4167–4221
https://doi.org/10.1214/16-AOP1162
© Institute of Mathematical Statistics, 2017

A CENTRAL LIMIT THEOREM FOR THE KPZ EQUATION

BY MARTIN HAIRER[1] AND HAO SHEN

University of Warwick

We consider the KPZ equation in one space dimension driven by a stationary centred space–time random field, which is sufficiently integrable and mixing, but not necessarily Gaussian. We show that, in the weakly asymmetric regime, the solution to this equation considered at a suitable large scale and in a suitable reference frame converges to the Hopf–Cole solution to the KPZ equation driven by space–time Gaussian white noise. While the limiting process depends only on the integrated variance of the driving field, the diverging constants appearing in the definition of the reference frame also depend on higher order moments.

CONTENTS

1. Introduction 4167
Structure of the article 4172
2. Framework 4173
2.1. Assumption on noise 4173
2.2. Regularity structures setup 4174
3. Renormalisation 4179
3.1. Joint cumulants 4180
3.2. Values of the renormalisation constants 4181
4. A priori bounds on the renormalised models 4185
4.1. First object and some generalities 4188
4.2. Bounds on the other objects 4196
Bounds on 4196
Bounds on 4198
Bounds on 4201
Bounds on 4202
5. General bounds 4205
6. Identification of the limit 4215
Acknowledgements 4220
References 4220

1. Introduction. Over the past fifteen years or so, there has been a great deal of interest in reaching a mathematical understanding of the "KPZ fixed point", a conjectured self-similar stochastic process describing the large-scale behaviour

Received July 2015; revised August 2016.
[1]Supported by the Philip Leverhulme Trust and the European Research Council.
MSC2010 subject classifications. Primary 60H15; secondary 35K55, 60H30.
Key words and phrases. KPZ equation, central limit theorem, Wiener chaos, cumulants.

of one-dimensional interface growth models [28]. While substantial progress has been achieved for microscopic models presenting some form of "integrable structure" (see [3, 6, 7, 10, 32, 34] and references therein for some of the major milestones), the underlying mechanisms leading to the universality of the KPZ fixed point in general are still very poorly understood.

A related question is that of the universality of the "crossover regime" in weakly asymmetric interface fluctuation models. Recall that the large-scale behaviour of perfectly symmetric models is expected to be governed by the Edwards–Wilkinson model [16], that is, the additive stochastic heat equation. Weakly asymmetric models are thus expected to exhibit a crossover regime where they transition from a behaviour described by the Edwards–Wilkinson model to one described by the KPZ fixed point. It is conjectured that this behaviour is also universal and is described by the KPZ equation [28], a nonlinear stochastic PDE formally given by

$$\partial_t h = \partial_x^2 h + \lambda(\partial_x h)^2 + \xi. \tag{1.1}$$

Here, ξ denotes space–time Gaussian white noise, $\lambda \in \mathbf{R}$, and the spatial variable x takes values in the circle S^1. One then recovers the Edwards–Wilkinson model for $\lambda = 0$, while the KPZ fixed point corresponds to $\lambda = \infty$ (modulo rescaling).

Mathematically, rigorous evidence for this was first given in [8], where the authors show that the height function of the weakly asymmetric simple exclusion process converges to (1.1) (interpreted in a suitable sense). More recently, this result was extended to several other discrete models in [2, 12, 15, 21]. One difficulty arising is that it is not clear *a priori* what (1.1) even means: solutions are not differentiable, so that $(\partial_x h)^2$ does not exist in the classical sense. The trick used in [8] to circumvent this is that the authors consider the so-called "Hopf–Cole solution" which, *by definition*, is given by $h = \lambda^{-1}\log(Z)$, where Z solves $dZ = \partial_x^2 Z\,dt + \lambda Z\,dW$ in the Itô sense; see [14].

A more direct interpretation of (1.1) was recently given in [22] using the theory of rough paths, where it was also shown that this interpretation generalises the Hopf–Cole solution. (In the sense that the solution theory relies on more input data than just a realisation of the driving noise ξ, but there is a canonical choice for this additional data for which the solution then coincides with the Hopf–Cole solution.) A somewhat different approach leading to the same result is given by the theory of regularity structures developed in [23] (see also [24] for an introduction to the theory). Its application to the KPZ equation is explained, for example, in [18], but was already implicit in [23, 25].

In [26], the authors exploit this theory to show that a rather large class of continuous weakly asymmetric interface fluctuation models is indeed described by the KPZ equation in the crossover regime. These models are of the type

$$\partial_t h = \partial_x^2 h + \sqrt{\varepsilon} P(\partial_x h) + \zeta,$$

for ζ a smooth space–time Gaussian field with finite range correlations and P an even polynomial. In that article, the Gaussianity assumption on the noise is used

in an essential way to show convergence of an associated "model" (in the sense of the theory of regularity structures) to the limiting model used in [18, 22].

The purpose of the present article is to remove the Gaussianity assumption of [26]. Since we focus on dealing with the non-Gaussian character of our model, we only consider the case of quadratic nonlinearities, but we expect that a result analogous to that of [26] also holds in our context. We assume that the driving noise is given by a stationary and centred space–time random field ζ with suitable regularity, integrability and mixing conditions given below. It is important to note that we do *not* assume that ζ is white in time, so martingale-based techniques as used, for example, in [8] and more recently in [31] do not apply here. We also remark that although we only consider the KPZ equation in this article, most of the methods such as the non-Gaussian Wick calculus, non-Gaussian diagrammatic techniques, etc. developed here should also apply to other equations such as the dynamical Φ^4 equation, the generalised parabolic Anderson model (gPAM) and so on.

Given a non-Gaussian random field ζ as above, we consider the equation

$$\partial_t h = \partial_x^2 h + \sqrt{\varepsilon}\lambda(\partial_x h)^2 + \zeta.$$

Rescaling it by setting $\tilde{h}_\varepsilon(t,x) = \varepsilon^{\frac{1}{2}} h(\varepsilon^{-2}t, \varepsilon^{-1}x)$ yields

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \lambda(\partial_x \tilde{h}_\varepsilon)^2 + \tilde{\zeta}_\varepsilon, \tag{1.2}$$

where $\tilde{\zeta}_\varepsilon = \varepsilon^{-3/2}\zeta(\varepsilon^{-2}t, \varepsilon^{-1}x)$. Our main result, Theorem 1.3 below, then states that there exists a suitable choice of inertial reference frame such that, in this frame, $\tilde{h}_\varepsilon$ converges as $\varepsilon \to 0$ to the Hopf–Cole solution of the KPZ equation. One interesting fact is that the "vertical speed" C_ε of this reference frame differs by a term of order $\varepsilon^{-1/2}$ from what it would be if ζ were replaced by a centred Gaussian field with the same covariance. This difference depends to leading order on the third cumulants of ζ and to order 1 on its fourth cumulants. The limiting solution however depends only on the covariance of ζ, which is why we call our result a "central limit theorem". We will also see that unless the covariance of ζ is invariant under spatial inversion $x \mapsto -x$, the frame in which convergence takes place also has a "horizontal speed" of order one. (This is also the case when the driving noise is Gaussian, but it didn't arise in [25, 26] since symmetry under spatial inversion was assumed there.)

Our assumptions on the space–time random field ζ are twofold. First, we need translation invariance and sufficient regularity/integrability.

ASSUMPTION 1.1. The random field ζ is stationary, centred and Lipschitz continuous. Furthermore, $\mathbf{E}(|\zeta(z)|^p + |\nabla\zeta(z)|^p) < \infty$ for all $p > 0$, and ζ is normalised so that $\mathbf{E}(\zeta(0)\zeta(z))$ integrates to 1.

Our second assumption is that there is enough independence. In order to formulate it, given $K \subset \mathbf{R}^2$, we denote by $\mathcal{F}_K$ the σ-algebra generated by the point evaluations $\{\zeta(z) : z \in K\}$.

ASSUMPTION 1.2. For any two compact sets $K_1, K_2 \subset \mathbf{R}^2$ such that $\inf_{z_i \in K_i} |z_1 - z_2| \geq 1$, $\mathcal{F}_{K_1}$ and $\mathcal{F}_{K_2}$ are independent.

The strict independency condition in Assumption 1.2 is only here for simplicity and may be relaxed. We expect both our results and our technique of proof to still hold when the dependency decays sufficiently fast (e.g., exponentially or polynomially with a sufficiently high power). Also, in Assumption 1.1, the condition that all moments of ζ are finite can be relaxed: our proof only requires us to have finite pth moments for some sufficiently large but fixed p.

One typical example of a random field ζ satisfying these assumptions is as follows. Let μ be a finite positive measure on $\mathcal{C}_0^1(\mathbf{R}^2)$ which is supported on the set of functions with support in the centred ball of radius $1/2$, and such that $\int \|\varphi\|_{\mathcal{C}^1}^p \mu(d\varphi) < \infty$ for every p. Denote by $\hat{\mu}$ a realisation of the Poisson point process on $\mathcal{C}_0^1(\mathbf{R}^2) \times \mathbf{R}^2$ with intensity $\mu \otimes dz$ and set

$$\zeta(z) = \int \varphi(z - z')\hat{\mu}(d\varphi \otimes dz') - \iint \varphi(z)\, dz \mu(d\varphi).$$

Then it is immediate that ζ satisfies Assumptions 1.1 and 1.2. Furthermore, nonlinear combinations $F(\zeta_1, \dots, \zeta_k)$ of fields of this type still satisfy these assumptions provided that F grows at most polynomially at infinity.

Given this setting, we state the main result of this article as the following theorem. Note that since we consider periodic boundary conditions, we need to replace the field ζ by a suitable "periodisation", which is formulated in Assumption 2.1 below. For simplicity of notation, we still denote by $\tilde{\zeta}_\varepsilon$ the noise obtained from rescaling the "periodised" random field.

THEOREM 1.3. *Let ζ be a random field satisfying Assumptions 1.1 and 1.2. Let $h_0^{(\varepsilon)}$ be a sequence of smooth functions on S^1 that converge in $\mathcal{C}^\beta$ as $\varepsilon \to 0$ to a limit $h_0 \in \mathcal{C}^\beta$ with $\beta \in (0, 1)$. Let $\tilde{h}_\varepsilon$ be the solution to (1.2) on S^1 with initial condition $h_0^{(\varepsilon)}$ and with $\tilde{\zeta}_\varepsilon$ as in (2.2) below satisfying Assumption 2.1. Then there exist velocities v_h and $v_v^{(\varepsilon)}$ such that, for every $T > 0$, the family of random functions $\tilde{h}_\varepsilon(t, x - v_h t) - v_v^{(\varepsilon)} t$ converges in law as $\varepsilon \to 0$ to the Hopf–Cole solution of the KPZ equation (1.1) with initial data h_0 in the space $\mathcal{C}^\eta([0, T] \times S^1)$, for any $\eta \in (0, \frac{1}{2} \wedge \beta)$.*

Furthermore, one has

$$v_v^{(\varepsilon)} = \lambda \varepsilon^{-1} C_0 + 2\lambda^2 \varepsilon^{-1/2} C_1 + \lambda^3 c, \qquad v_h = 4\lambda^2 \hat{c}, \tag{1.3}$$

where C_0 and $\hat{c}$ only depend on the second moment of the random field ζ, while C_1 depends on its third moment and c depends on the second and fourth moments. If $\mathbf{E}(\zeta(0, 0)\zeta(t, x))$ is even as a function of x, then $\hat{c} = 0$.

Here, $\mathcal{C}^\alpha$ is the Hölder space of functions or distributions with regularity α as in [23]. (This coincides with the parabolic Besov space $B^\alpha_{\infty,\infty,\mathrm{loc}}$.) The subscripts h and v in our notation only indicate that v_h and $v_v^{(\varepsilon)}$ are the "horizontal" and "vertical" speeds, respectively, as mentioned after (1.2). The choice of the constants appearing in Theorem 1.3 will be given in Section 3.2; in particular $\hat{c}$ will be defined in Lemma 3.6.

REMARK 1.4. The constant C_0 is the one given by the "naïve" Wick ordering and can be explicitly expressed as

$$C_0 = \int_{\mathbf{R}^4} P'(z)P'(\bar{z})\kappa_2(z-\bar{z})\,dz\,d\bar{z}, \tag{1.4}$$

where P denotes the heat kernel, P' is its space derivative, and $\kappa_2(z-\bar{z}) = \mathbf{E}(\zeta(z)\zeta(\bar{z}))$. (For comparison, we should identify κ_2 here with $\varrho * \varrho$ where ϱ is the mollifier function in [26].)

REMARK 1.5. At first sight, the result may appear somewhat trivial since, by the classical central limit theorem, $\tilde{\zeta}_\varepsilon \to \xi$ weakly so that it may not be surprising that solutions to (1.2) converge to solutions to (1.1). The problem of course is that the solution map to (1.1) is not continuous in any of the usual topologies in which this weak convergence takes place. This is apparent in the fact that, in order to obtain the right limit, we need to perform a nontrivial change of reference frame with divergent velocities that depend on higher order cumulants of ζ.

PROOF OF THEOREM 1.3. Let $h_\varepsilon(t,x) \stackrel{\mathrm{def}}{=} \tilde{h}_\varepsilon(t, x - v_h t) - v_v^{(\varepsilon)} t$. Then it is immediate to check that h_ε satisfies the equation

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \lambda(\partial_x h_\varepsilon)^2 - v_h\,\partial_x h_\varepsilon - v_v^{(\varepsilon)} + \zeta_\varepsilon, \tag{1.5}$$

where ζ_ε is defined in (2.2) below. To prove the convergence of h_ε as stated in the theorem, we apply the theory of regularity structures. The theory is briefly reviewed in Section 2.2, where the regularity structure $\mathscr{T}$ suitable for the study of equation (1.1) is defined. The proof of Theorem 1.3 then follows in the following way.

We consider the (local) solution map to the abstract fixed point problem for functions H taking values in a suitable regularity structure [see (2.12) below]

$$H = \mathcal{P}\mathbf{1}_{t>0}\big(\lambda(\mathscr{D}H)^2 + \Xi\big) + Ph_0, \tag{1.6}$$

where $\mathcal{P}$ is a suitable operator representing space–time convolution with the heat kernel and Ξ is an element in the regularity structure representing the noise term. It is then known from [18, 22, 23, 25, 26] that there exists a natural model $\hat{Z} = (\hat{\Pi}, \hat{\Gamma})$ for $\mathscr{T}$ such that if H solves (1.6) for $\hat{Z}$, then $\mathcal{R}H$ coincides with the Hopf–Cole

solution to the KPZ equation. Furthermore, the model in question is obtained as a limit of renormalised lifts of smooth Gaussian approximations to space–time white noise.

As a next step, we define a family of renormalised models $\hat{Z}_\varepsilon$ (see Section 3) and show that they obey the following two crucial properties:

- By appropriately choosing the values of the renormalisation constants, we can ensure that $\hat{Z}_\varepsilon$ converges in distribution as $\varepsilon \to 0$ to $\hat{Z}$. This is the content of Theorem 6.5 below.
- By appropriately choosing v_h and $v_v^{(\varepsilon)}$, we can ensure that if H_ε solves (1.6) for $\hat{Z}_\varepsilon$, then $\mathcal{R}H_\varepsilon$ coincides with the solution h_ε to (1.5). See (3.3), combined with the fact that the correct choice of renormalisation constants for the first step is given by (3.10). In particular, v_h and $v_v^{(\varepsilon)}$ indeed have the forms claimed in (1.3).

As a consequence of [23], Theorem 7.8, these two properties immediately yield the existence of a random time τ such that h_ε, when restricted to the interval $[0, \tau]$, converges in law to the solution to the Hopf–Cole solution to the KPZ equation in $\mathcal{C}^\eta([0, T] \times S^1)$, for every $\eta \in (0, \frac{1}{2} \wedge \beta)$, see also the discussion at the end of Section 2.2.

The extension of the convergence result to any finite deterministic time interval follows from [23], Proposition 7.11, and the a priori knowledge that Hopf–Cole solutions to the KPZ equation are global and α-Hölder continuous for every $\alpha < 1/2$. □

Structure of the article. The article is organised as follows. We start with the basic setup of our problem in Section 2, which includes some more discussions about the assumptions on the noise, followed by a brief introduction to the theory of regularity structures as the framework for our proof. In Section 3, we then define a collection of renormalisation constants which, through the action of the renormalisation group, determines a family of renormalised models in the sense of [23]. In Section 4, we show some key technical results (Proposition 4.6, Lemma 4.7, Proposition 4.18), then obtain bounds on arbitrary moments of these renormalised models, uniformly in the small parameter ε. This is the most technical step: because our models are built from a non-Gaussian random field, their moments are not automatically bounded by their variance, unlike in the Gaussian case where equivalence of moments holds. The proof of these moment bounds depends on some general technical tools developed in Section 5. The main result of that section is Corollary 5.15, which provides conditions that can be straightforwardly checked and yield the desired moment estimates. In the last section, Section 6, we show Theorem 6.5 which identifies the limiting model as the "KPZ model" and, therefore, implies that in a suitable reference frame the limiting solution does indeed coincide with the Hopf–Cole solution to the KPZ equation driven by Gaussian white noise.

2. Framework.

2.1. *Assumption on noise.* We assume that ζ satisfies Assumption 1.1 and 1.2. Let then $\{\zeta^{(\varepsilon)}\}_{\varepsilon\in(0,1]}$ be a family of random fields with the following properties.

ASSUMPTION 2.1. The fields $\zeta^{(\varepsilon)}$ are stationary, almost surely periodic in space with period $1/\varepsilon$, and satisfy $\sup_{\varepsilon\in(0,1]}\mathbf{E}|\zeta^{(\varepsilon)}(0)|^p<\infty$ for every $p\geq 1$.

Furthermore, for any two sets K_1, $K_2\subset\mathbf{R}^2$ that are periodic in space with period $1/\varepsilon$ and such that $\inf_{z_i\in K_i}|z_1-z_2|\geq 1$, $\mathcal{F}^{(\varepsilon)}_{K_1}$ and $\mathcal{F}^{(\varepsilon)}_{K_2}$ are independent, where $\mathcal{F}^{(\varepsilon)}_{K}$ is the σ-algebra generated by $\{\zeta^{(\varepsilon)}(z):z\in K\}$.

Finally, for every $\varepsilon>0$, there is a coupling of ζ and $\zeta^{(\varepsilon)}$ such that, for every $T>0$ and every $\delta>0$,

$$\sup_{|t|\leq T\varepsilon^{-2}}\sup_{|x|\leq(1-\delta)/(2\varepsilon)}\lim_{\varepsilon\to 0}\varepsilon^{-3}\mathbf{E}\big|\zeta(t,x)-\zeta^{(\varepsilon)}(t,x)\big|^2=0. \tag{2.1}$$

REMARK 2.2. We could have replaced (2.1) by a slightly weaker condition, but the one given here has the advantage of being easy to state. It is also easy to verify in the examples we have in mind; see Example 2.3 below.

We then set

$$\begin{aligned}\tilde\zeta_\varepsilon(t,x)&=\varepsilon^{-3/2}\zeta^{(\varepsilon)}\big(t/\varepsilon^2,x/\varepsilon\big),\\ \zeta_\varepsilon(t,x)&=\varepsilon^{-3/2}\zeta^{(\varepsilon)}\big(t/\varepsilon^2,(x-v_ht)/\varepsilon\big),\end{aligned} \tag{2.2}$$

where v_h will be specified in Section 3.2 [see Lemma 3.6 and recall from (1.3) that $v_h=4\lambda^2\hat c$]. Note that both ζ_ε and $\tilde\zeta_\varepsilon$ are periodic in space with period 1.

EXAMPLE 2.3. To show that Assumption 2.1 is not unreasonable, consider the following example. Let μ be a Poisson point process on $\mathbf{R}^2\times[0,1]$ with uniform intensity measure, let $\varphi(t,x,a)$ be a smooth compactly supported function (say with support in the ball of radius 1), and set

$$\zeta(t,x)=\int_{\mathbf{R}^2\times[0,1]}\varphi(t-s,x-y,a)\mu(ds,dy,da). \tag{2.3}$$

Let $\mu^{(\varepsilon)}$ be the periodic extension to $\mathbf{R}^2\times[0,1]$ of a Poisson point process on $\mathbf{R}\times[-1/(2\varepsilon),1/(2\varepsilon)]\times[0,1]$ with uniform intensity measure and let $\zeta^{(\varepsilon)}$ be as in (2.3), with μ replaced by $\mu^{(\varepsilon)}$. Then it is immediate to verify that (2.1) is satisfied, since the natural coupling between ζ and $\zeta^{(\varepsilon)}$ is such that $\zeta=\zeta^{(\varepsilon)}$ on $\mathbf{R}\times[K-1/(2\varepsilon),1/(2\varepsilon)-K]$, for some fixed K.

2.2. *Regularity structures setup*. In order to prove the convergence result stated in Theorem 1.3, we will make use of the theory of regularity structures developed in [23]. Before delving into the technical details, we give a short summary of the main concepts of the theory, as applied to the problem at hand, in this subsection. One can find a more concise exposition of the theory in the lecture notes [24].

Recall that a *regularity structure* consists of two pieces of data. First, a graded vector space $\mathcal{T} = \bigoplus_{\alpha \in A} \mathcal{T}_\alpha$, where the index set A, called the set of homogeneities, is a subset of $\mathbf{R}$ which is locally finite and bounded from below. In our case, each $\mathcal{T}_\alpha$ is furthermore finite-dimensional. Second, a "structure group" $\mathcal{G}$ of continuous linear transformations of $\mathcal{T}$ such that, for every $\Gamma \in \mathcal{G}$, every $\alpha \in A$, and every $\tau \in \mathcal{T}_\alpha$, one has

$$\Gamma\tau - \tau \in \mathcal{T}_{<\alpha} \qquad \text{with } \mathcal{T}_{<\alpha} \stackrel{\text{def}}{=} \bigoplus_{\beta<\alpha} \mathcal{T}_\beta. \tag{2.4}$$

A canonical example is the space of polynomials in d indeterminates, with index set $A = \mathbf{N}$, $\mathcal{T}_n$ consisting of homogeneous polynomials of $\mathfrak{s}$-scaled degree n, and $\mathcal{G}$ consisting of translations. We recall here a scaling $\mathfrak{s} \in \mathbf{N}^d$ is simply a vector $(\mathfrak{s}_1, \dots, \mathfrak{s}_d)$ of positive relatively prime integers, and the $\mathfrak{s}$-scaled degree of a monomial $X^k \stackrel{\text{def}}{=} \prod_{i=1}^d X_i^{k_i}$, where $k = (k_1, \dots, k_d)$ is a multi-index, is defined as $|k| \stackrel{\text{def}}{=} \sum_{i=1}^d \mathfrak{s}_i k_i$.

In general, [24], Section 8, gives a recipe to build a regularity structure from any given family of (subcritical in an appropriate sense) semi-linear stochastic PDEs. In our case, space and time are both one dimensional so $d = 2$, and the natural scaling to choose is the parabolic scaling $\mathfrak{s} = (2, 1)$, and thus the scaling dimension of space–time is

$$|\mathfrak{s}| = 3$$

which is fixed throughout the article. This scaling defines a distance $\|x - y\|_\mathfrak{s}$ on $\mathbf{R}^2$ by

$$\|x\|_\mathfrak{s}^4 \stackrel{\text{def}}{=} |x_0|^2 + |x_1|^4,$$

where x_0 and x_1 are time and space coordinates, respectively. To simplify the notation, in the sequel we will often just write $|x| = \|x\|_\mathfrak{s}$ for a space–time point x.

The regularity structure relevant for the analysis of (1.1) is then built in the following way. We write $\mathcal{U}$ for a collection of formal expressions that will be useful to describe the solution h, $\mathcal{U}'$ for a collection of formal expressions useful to describe its spatial distributional derivative $\partial_x h$, and $\mathcal{V}$ for a collection of formal expressions useful to describe the right-hand side of the KPZ equation (1.1). We decree that $\mathcal{U}$ and $\mathcal{U}'$ contain at least the polynomials of two indeterminates X_0, X_1, denoting the time and space directions, respectively. We then introduce three additional

symbols, Ξ, $\mathcal{I}$ and $\mathcal{I}'$, where Ξ will be interpreted as an abstract representation of the driving noise ζ (or rather ζ_ε), and $\mathcal{I}$ and $\mathcal{I}'$ will be interpreted as the operation of convolving with a truncation of the heat kernel and its spatial derivative, respectively. In view of the structure of the equation (1.1), to have a regularity structure that is rich enough to describe the fixed-point problem we should also decree that

$$\begin{aligned} \tau, \bar\tau \in \mathcal{U}' \quad &\Rightarrow \quad \tau\bar\tau \in \mathcal{V}, \\ \tau \in \mathcal{V} \quad &\Rightarrow \quad \mathcal{I}(\tau) \in \mathcal{U}, \qquad \mathcal{I}'(\tau) \in \mathcal{U}'. \end{aligned} \tag{2.5}$$

The sets $\mathcal{U}$, $\mathcal{U}'$ and $\mathcal{V}$ are then defined as the smallest collection of formal expressions such that $\Xi \in \mathcal{V}$, $X^k \in \mathcal{U}$, $X^k \in \mathcal{U}'$, and (2.5) holds. Following [23], we furthermore decree that $\tau\bar\tau = \bar\tau\tau$ and $\mathcal{I}(X^k) = \mathcal{I}'(X^k) = 0$. We then define

$$\mathcal{W} \stackrel{\text{def}}{=} \mathcal{U} \cup \mathcal{U}' \cup \mathcal{V}.$$

For each formal expression τ, the homogeneity $|\tau| \in \mathbf{R}$ is assigned in the following way. First of all, for a multi-index $k = (k_0, k_1)$, we set $|X^k| = |k| = 2k_0 + k_1$, which is consistent with the parabolic scaling $\mathfrak{s}$ chosen above. We then set

$$|\Xi| = -\frac{3}{2} - \bar\kappa,$$

where $\bar\kappa > 0$ is a fixed small number, and we define the homogeneity for every formal expression by

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

The linear space $\mathcal{T}$ is then defined as the linear span of $\mathcal{W}$, and $\mathcal{T}_\alpha$ is the subspace spanned by $\{\tau : |\tau| = \alpha\}$. By a simple power-counting argument, one finds that as long as $\bar\kappa < 1/2$, the sets $\{\tau \in \mathcal{W} : |\tau| < \gamma\}$ are finite for every $\gamma \in \mathbf{R}$, reflecting the fact that the KPZ equation (1.1) in one space dimension is subcritical.

As in [22], we use a graphical shorthand for the formal expressions in $\mathcal{W}$. We use dots to represent the expression Ξ, and lines to represent the operator $\mathcal{I}'$. Joining of formal expressions by their roots denotes their product. For example, one has $= \mathcal{I}'(\Xi)$. Writing $\Psi = \mathcal{I}'(\Xi)$ as a shorthand, one also has $= \Psi^2$ and $= \Psi\mathcal{I}'(\Psi^2)$, etc. With this notation, the formal expressions in $\mathcal{W}$ with negative homogeneities other than Ξ are given by

$$\begin{aligned} &|\text{}| = -1 - 2\bar\kappa, \qquad |\text{}| = -\frac{1}{2} - 3\bar\kappa, \qquad |\text{}| = -\frac{1}{2} - \bar\kappa, \\ &|\text{}| = |\text{ }| = -4\bar\kappa, \qquad |\text{ }| = |\text{}| = -2\bar\kappa, \end{aligned} \tag{2.6}$$

provided that $\bar\kappa > 0$ is sufficiently small. We will denote by $\mathcal{W}_-$ the above formal expressions with negative homogeneities. In fact, we will never need to consider the full space $\mathcal{T}$, but it will be sufficient to consider the subspace generated by all elements of homogeneity less than some large enough number σ. In practice, it actually turns out to be sufficient to choose any $\sigma > \frac{3}{2} + \bar\kappa$.

The last ingredient of the regularity structure, namely the structure group $\mathcal{G}$, can also be described explicitly. However, the precise description of $\mathcal{G}$ does not matter in the present paper, so we refrain from giving it. The interested reader is referred to [23], Section 8, as well as [26], Section 3, for the more general case in which the nonlinearity of the equation also contains powers of higher order. The regularity structure $(\mathcal{T}, \mathcal{G})$ defined here is the same as in [18], Chapter 15.

Given the regularity structure, a crucial concept is that of a *model* for it, which associates to each formal expression in the regularity structure a "concrete" function or distribution on $\mathbf{R}^d$ (which will be the space–time $\mathbf{R}^2$ in our case). It consists of a pair (Π, Γ) of functions

$$\Pi\colon \mathbf{R}^d \to \mathcal{L}(\mathcal{T}, \mathcal{S}'), \qquad \Gamma\colon \mathbf{R}^d \times \mathbf{R}^d \to \mathcal{G},$$

$$x \mapsto \Pi_x \qquad (x, y) \mapsto \Gamma_{xy},$$

where $\mathcal{L}(\mathcal{T}, \mathcal{S}')$ is the space of linear maps from $\mathcal{T}$ to the space of distributions (over space–time $\mathbf{R}^2$ in our case), such that Γ_{xx} is the identity and

$$\Pi_y = \Pi_x \circ \Gamma_{xy}, \qquad \Gamma_{xy}\Gamma_{yz} = \Gamma_{xz},$$

for every $x, y, z \in \mathbf{R}^d$. Furthermore, for every $\gamma > 0$, we require that

$$\bigl|(\Pi_x \tau)(\varphi_x^\lambda)\bigr| \lesssim \lambda^{|\tau|}, \qquad \|\Gamma_{xy}\tau\|_m \lesssim |x - y|^{|\tau| - m}, \tag{2.7}$$

where "$\lesssim$" stands for "less or equal" up to a constant which is uniform over x, y in any compact set of $\mathbf{R}^d$, and over all $\lambda \in (0, 1]$, all smooth test functions $\varphi \in \mathcal{B}$, all $m < |\tau|$, and all homogeneous elements $\tau \in \mathcal{T}_{<\gamma}$ with $\|\tau\| \le 1$. Here, $\|\cdot\|_m$ denotes the norm of the component in $\mathcal{T}_m$ (which could be any norm since $\mathcal{T}_m$ is finite dimensional),

$$\mathcal{B} = \bigl\{\varphi : \|\varphi\|_{\mathcal{C}^2} \le 1, \text{ and } \varphi \text{ is supported in the unit ball}\bigr\}, \tag{2.8}$$

and $\varphi_x^\lambda(z)$ denotes the following recentred and rescaled[2] version of φ :

$$\varphi_x^\lambda(z) \stackrel{\text{def}}{=} \lambda^{-3}\varphi\biggl(\frac{z_0 - x_0}{\lambda^2}, \frac{z_1 - x_1}{\lambda}\biggr),$$

where the scaling exponents are chosen according to our scaling $\mathfrak{s} = (2, 1)$. The reason why $\mathcal{C}^2$ appears in (2.8) is that this is the smallest integer r such that $r + \min_{\tau \in \mathcal{W}} |\tau| > 0$.

Besides these requirements, we also impose that our models are *admissible*. To define this notion, as in [23], we first fix a kernel $K : \mathbf{R}^2 \to \mathbf{R}$ such that $\operatorname{supp} K \subset \{|(t, x)| \le 1, t > 0\}$, $K(t, -x) = K(t, x)$ and K coincides with the heat kernel in $|(t, x)| < 1/2$. We also require that for every polynomial Q on $\mathbf{R}^2$ of

[2]We note that the notation λ which stands for the rescaling parameter here has also been used in the equation (1.1) as the coefficient of the nonlinearity. This should not cause any confusion because it will be clear what λ means from the context.

parabolic degree 2 [i.e., $Q(t,x) = a_1 t + a_2 x^2 + a_3 x + a_4$ for $a_i \in \mathbf{R}$] one has $\int_{\mathbf{R}^2} K(t,x) Q(t,x)\,dx\,dt = 0$.

Given this truncated heat kernel K, the set $\mathscr{M}$ of *admissible models* consists of those models such that, for every multi-index k,

$$\big(\Pi_z X^k\big)(\bar z) = (\bar z - z)^k, \tag{2.9}$$

and such that

$$\begin{aligned}(\Pi_z \mathcal{I}\tau)(\bar z) &= \int_{\mathbf{R}^2} K(\bar z - x)(\Pi_z \tau)(dx) \\ &\quad - \sum_{|k|<|\mathcal{I}\tau|} \frac{(\bar z - z)^k}{k!} \int_{\mathbf{R}^2} D^k K(z-x)(\Pi_z\tau)(dx).\end{aligned} \tag{2.10}$$

We refer to [23], Section 5, for a discussion on the meaning of these expressions in general. As an example to illustrate (2.10), for $\mathcal{I}\tau = \,\raisebox{-0.2em}{<}$ with $|\,\raisebox{-0.2em}{<}\,| = \frac{1}{2} - \bar\kappa$, one has

$$(\Pi_z \,\raisebox{-0.2em}{<}\,)(\bar z) = \int_{\mathbf{R}^2} \big(K(\bar z - x) - K(z-x)\big)(\Pi_z \,\raisebox{0.2em}{!}\,)(dx).$$

When analysing the objects $\,\raisebox{-0.2em}{<}\bullet$ and $\,\raisebox{-0.2em}{<}\bullet\bullet$ in Section 4 we will see how the increment of K appearing in this expression is exploited.

REMARK 2.4. It actually turns out that if Π satisfies the *first* analytical bound in (2.7), and the structure group $\mathcal{G}$ acts on $\mathcal{T}$ in a way that is compatible with the definition of the admissible model (which is true in our case), then the second analytical bound in (2.7) is *automatically* satisfied. This is a consequence of [23], Theorem 5.14.

Given a continuous space–time function ζ_ε, there is a canonical way of building an admissible model $(\Pi^{(\varepsilon)}, \Gamma^{(\varepsilon)})$, as in [23], Section 8. First, we set $\Pi_z^{(\varepsilon)} \Xi = \zeta_\varepsilon$ independently of z, and we define it on X^k as in (2.9). Then we define $\Pi_z^{(\varepsilon)}$ recursively by (2.10), as well as the identity

$$\big(\Pi_z^{(\varepsilon)} \tau\bar\tau\big)(\bar z) = \big(\Pi_z^{(\varepsilon)}\tau\big)(\bar z)\big(\Pi_z^{(\varepsilon)}\bar\tau\big)(\bar z). \tag{2.11}$$

Note that this is only guaranteed to make sense if ζ_ε is a sufficiently regular function. It was shown in [23], Proposition 8.27, that this does indeed define an admissible model for every continuous function ζ_ε, and we will call this admissible model the canonical lift of ζ_ε. However, we emphasize that, unlike in the case of rough paths [17], not every admissible model is obtained in this way, or even as a limit of such models. In particular, while the renormalisation procedure used in Section 3 preserves the admissibility of models, it does not preserve the property of being a limit of canonical lifts of regular functions.

It was then shown in [23] how to associate an *abstract fixed point problem* to the equation (1.1) in a certain space $\mathcal{D}^{\gamma,\eta}$ of $\mathcal{T}_{<\gamma}$-valued functions. This space is the analogue in this context of a Hölder space of order γ, with the exponent η allowing for a possible singular behavior at $t=0$. In our particular case, the abstract fixed point problem can be formulated in $\mathcal{D}^{\gamma,\beta}$ for $\gamma > \frac{3}{2}+\bar{\kappa}$, and for $\beta > 0$ such that the initial data $h_0 \in \mathcal{C}^\beta$. One of the main results in [23] is the *reconstruction theorem* which states that for every $U \in \mathcal{D}^{\gamma,\beta}$ with $\gamma > 0$, there exists a unique distribution $\mathcal{R}U$ such that, near every point z, $\mathcal{R}U$ "looks like" $\Pi_z U(z)$ up to an error of order γ. The operator $U \mapsto \mathcal{R}U$ is called the reconstruction operator associated to the model (Π, Γ).

The idea to formulate this abstract fixed point is to define multiplication, differentiation, and integration against the heat kernel on elements in $\mathcal{D}^{\gamma,\eta}$, so one can write

$$H = \mathcal{P}\mathbf{1}_{t>0}\big(\lambda(\mathscr{D}H)^2 + \Xi\big) + Ph_0, \tag{2.12}$$

where Ph_0 is the heat kernel acting on the initial data h_0 and can be interpreted as an element of $\mathcal{D}^{\gamma,\beta}$ by [23], Lemma 7.5, $\mathscr{D}$ is given by $\mathscr{D}\mathcal{I}(\tau) = \mathcal{I}'(\tau)$, and the product is simply given by pointwise multiplication in $\mathcal{T}$. The symbol $\mathbf{1}_{t>0}$ represents the (scalar) indicator function of the set $\{(t,x) : t > 0\}$ and the linear operator $\mathcal{P}$ represents space–time convolution by the heat kernel in the space $\mathcal{D}^{\gamma,\beta}$, in the sense that

$$\mathcal{R}\mathcal{P}H = P * \mathcal{R}H,$$

for every $H \in \mathcal{D}^{\gamma,\eta}$ for $\gamma > 0$ and $\eta > -2$. The explicit expression for the operator $\mathcal{P}$ does not matter for our purpose but can be found in [23], Section 5, let us just mention that $\mathcal{P}H$ and $\mathcal{I}H$ only differ by elements taking values in the linear span of the X^k and that it satisfies a Schauder estimate.

Theorem [23], Theorem 7.8, shows that, for every admissible model, there exists a unique $T > 0$ such that the above fixed-point problem has a unique solution in $\mathcal{D}^{\gamma,\beta}([0,T] \times S^1)$. If the model is given by the canonical lift $(\Pi^{(\varepsilon)}, \Gamma^{(\varepsilon)})$ of a continuous function ζ_ε, then $\mathcal{R}H$ coincides with the classical solution to (1.1).

Unfortunately, these canonical lifts do not converge to a limit in $\mathscr{M}$ as $\varepsilon \to 0$. However, we will show that one can build a natural finite-dimensional family of continuous transformations $\hat{M}_\varepsilon$ of $\mathscr{M}$ such that the "renormalised models" $(\hat{\Pi}^{(\varepsilon)}, \hat{\Gamma}^{(\varepsilon)}) \stackrel{\text{def}}{=} \hat{M}_\varepsilon(\Pi^{(\varepsilon)}, \Gamma^{(\varepsilon)})$ do converge to an admissible limiting model. These transformations are parametrised by elements of a *renormalisation group* $\mathfrak{R}$ associated to our regularity structure $(\mathcal{T}, \mathcal{G})$. Since the precise definition of $\mathfrak{R}$ requires a bit more understanding of the algebraic properties of our regularity structure and is not quite relevant to the present article, we simply refer to [23], Section 8.

Once we obtain the convergence of renormalised models, the function $h = \mathcal{R}H$ with $\mathcal{R}$ the reconstruction map associated with the limiting model is then the limiting solution stated in Theorem 1.3. The Hölder regularity of the solution h is given

by the minimum of the regularity of Ph_0 and the lowest homogeneity of elements of $\mathcal{U}$ beside the Taylor polynomials which is $\mathcal{I}(\Xi)$ with homogeneity $1/2 - \bar{\kappa}$ where $\bar{\kappa} > 0$ is a sufficiently small fixed number, so [23], Proposition 3.28, yields $h \in \mathcal{C}^\eta([0,T] \times S^1)$, with $\eta \in (\frac{1}{4}, \frac{1}{2} \wedge \beta)$.

3. Renormalisation. We now give an explicit description of the renormalisation maps $\hat{M}_\varepsilon$ described above. These are parametrised by linear maps $M_\varepsilon : \mathcal{T} \to \mathcal{T}$ belonging to the "renormalisation group" $\mathfrak{R}$ which was introduced in this context in [23], Section 8.3. As a matter of fact, we only need to consider a certain 5-parameter subgroup of $\mathfrak{R}$. This subgroup consists of elements $M \in \mathfrak{R}$ of the form $M = \exp(-\sum_{i=1}^5 \ell_i L_i)$ where the ℓ_i are real-valued constants and the generators $L_i : \mathcal{T} \to \mathcal{T}$ are determined by the following contraction rules:

$$L_1 : \Xi \mapsto \mathbf{1}, \qquad L_2 : \Xi\bullet \mapsto \mathbf{1}, \qquad L_3 : \Xi\bullet\bullet \mapsto \mathbf{1}, \tag{3.1}$$
$$L_4 : \Xi\bullet\bullet \mapsto \mathbf{1}, \qquad L_5 : \Xi\Xi \mapsto \mathbf{1}.$$

For L_i with $i \neq 2$, these rules are extended to the whole regularity structure by setting $L_i\tau = 0$ if τ is not the element given above. For L_2, however, the above contraction rule should be understood in the sense that for an arbitrary formal expression τ, $L_2\tau$ equals the sum of all expressions obtained by performing a substitution of the type $\Xi\bullet \mapsto \mathbf{1}$. For instance, one has

$$L_2 \Xi\bullet\bullet = 2\,\mathbf{I}, \qquad L_2 \Xi\bullet\bullet = 2\Xi\bullet + \Upsilon.$$

Given an admissible model (Π, Γ) and an element M of the above type for some choice of constants ℓ_i, we can then build a new model $(\hat{\Pi}, \hat{\Gamma}) = \hat{M}(\Pi, \Gamma)$ by postulating that the identity

$$\hat{\Pi}_x \tau = \Pi_x M \tau, \tag{3.2}$$

holds for every $\tau \in \mathcal{T}$. This is sufficient to determine $(\hat{\Pi}, \hat{\Gamma})$ since in our case $\hat{\Gamma}$ is uniquely determined by $\hat{\Pi}$ and the knowledge that the new model is again admissible; see [23].

REMARK 3.1. The fact that $M \in \mathfrak{R}$ can be checked essentially in the same way as in [26], with the simplification that since we are only interested in a quadratic nonlinearity, the symbol $\mathcal{E}$ never appears. The only minor difference is the appearance of the generators L_2 and L_3 here. In fact, using the notation of [23], Section 8.3, one can verify that in our case one has the identities $\hat{\Delta}^M \mathcal{J}_k(\tau) = \mathcal{J}_k(M\tau) \otimes \mathbf{1}$ and $\Delta^M \tau = M\tau \otimes \mathbf{1}$. The second identity, combined with [23], equation (8.34), implies (3.2), while the combination of both identities and the fact that M is upper triangular imply that the "upper triangular condition" in [23], Definition 8.41, holds.

Let now (Π, Γ) be the model obtained by the canonical lift of an arbitrary smooth function ζ, and let $(\hat{\Pi}, \hat{\Gamma}) = \hat{M}(\Pi, \Gamma)$ be as above, again with $M = \exp(-\sum_{i=1}^{5} \ell_i L_i)$. It is then straightforward to verify that if $\mathcal{R}^M$ is the reconstruction operator associated to $(\hat{\Pi}, \hat{\Gamma})$ and H solves the corresponding abstract fixed point problem (2.12), then $h = \mathcal{R}^M H$ satisfies the renormalised equation

$$\begin{aligned} \partial_t h = \partial_x^2 h + \lambda(\partial_x h)^2 - 4\lambda^2 \ell_2\, \partial_x h + \zeta \\ - \big(\lambda \ell_1 + 2\lambda^2 \ell_3 + 4\lambda^3 \ell_4 + \lambda^3 \ell_5\big) + 4\lambda^3 \ell_2^2. \end{aligned} \tag{3.3}$$

This can be shown in the same way as [18], Proposition 15.12, except that in our case the term $2\lambda^2$ appearing in the expression of $(\partial H)^2 + \Xi$ produces a constant $2\lambda^2 \ell_3$ when being acted upon by L_3 in the definition of M. Note that the last term $4\lambda^3 \ell_2^2$ comes from the action of $\frac{1}{2}(\ell_2 L_2)^2$ in the exponential that defines M on the term $4\lambda^3$.

3.1. *Joint cumulants*. Before giving the definition of the actual values for the ℓ_i relevant to our analysis, we review the definition and basic properties of joint cumulants. Given a collection of random variables $\mathcal{X} = \{X_\alpha\}_{\alpha \in A}$ for some index set A, and a subset $B \subset A$, we write $X_B \subset \mathcal{X}$ and X^B as shorthand for

$$X_B = \{X_\alpha : \alpha \in B\}, \qquad X^B = \prod_{\alpha \in B} X_\alpha.$$

Given a finite set B, we furthermore write $\mathcal{P}(B)$ for the collection of all partitions of B, i.e. all sets $\pi \subset \mathscr{P}(B)$ (the power set of B) such that $\cup \pi = B$ and such that any two distinct elements of π are disjoint.

DEFINITION 3.2. Let $\mathcal{X}$ be a collection of random variables as above with finite moments of all orders. For any finite set $B \subset A$, we define the cumulant $\mathbf{E}_c(X_B)$ inductively over $|B|$ by

$$\mathbf{E}(X^B) = \sum_{\pi \in \mathcal{P}(B)} \prod_{\bar{B} \in \pi} \mathbf{E}_c(X_{\bar{B}}). \tag{3.4}$$

The expression (3.4) does indeed determine the cumulants uniquely by induction over $|B|$. This is because the right-hand side only involves $\mathbf{E}_c(X_B)$, which is what we want to define, as well as $\mathbf{E}_c(X_{\bar{B}})$ for some $\bar{B}$ with $|\bar{B}| < |B|$, which is already defined by the inductive hypothesis. If all the random variables are centred and jointly Gaussian, then it follows from Wick's theorem that $\mathbf{E}_c(X_B)$ always vanishes unless $|B| = 2$. Henceforth, we will use the notation κ_n for the nth joint cumulant function of the field $\zeta^{(\varepsilon)}$:

$$\kappa_n(z_1, \dots, z_n) \stackrel{\text{def}}{=} \mathbf{E}_c\big(\{\zeta^{(\varepsilon)}(z_1), \dots, \zeta^{(\varepsilon)}(z_n)\}\big). \tag{3.5}$$

Note that $\kappa_1 = 0$ since $\zeta^{(\varepsilon)}$ is assumed to be centred and κ_2 is its covariance function.

REMARK 3.3. A reader who is expertised in cumulants may know that one can also define cumulants for a collection of random variables in which some variables may appear more than once. In order to keep notation simple, in Definition 3.2 we only considered subsets B in the above definition, so that each random variable X_α is only allowed to appear at most once in the collection X_B. This is sufficient for the purpose of this article. As a matter of fact, one can easily reduce oneself to this case by considering the augmented collection $\bar{\mathcal{X}} = \{\bar{X}_{\bar{\alpha}}\}_{\bar{\alpha}\in\bar{A}}$ with $\bar{A} = A \times \mathbf{N}$ and random variables $\bar{X}_{\bar{\alpha}}$ such that $\bar{X}_{(\alpha,k)} = X_\alpha$ almost surely.

REMARK 3.4. There is a slight abuse of notation here since κ_n does of course depend on ε in general, but this dependence is very weak.

We refer to [33] and [20], Section 13.5, for the properties of joint cumulants; see also the recent article [30]. A property that will be useful in our problem is that the cumulant is zero when some of the variables are independent from the others. We formulate this in terms of our random field $\zeta^{(\varepsilon)}$. Given a collection $\{z_i\}_{i=1}^p$ of space–time points, we define $z_i \sim z_j$ if $|z_i - z_j| \le 1$, and extend this into an equivalence relation on $\{z_i\}_{i=1}^p$.

LEMMA 3.5. *The cumulants have the property that if* $\kappa_p(z_1, \dots, z_p) \neq 0$, *then* $z_1, \dots, z_p$ *all belong to the same equivalence class.* □

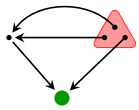
3.2. *Values of the renormalisation constants.* We now have all the ingredients in place to determine the relevant values of the renormalisation constants ℓ_i. We denote by $\kappa_p^{(\varepsilon)}$ the pth cumulant function of ζ_ε, which is the properly rescaled cumulant function of $\zeta^{(\varepsilon)}$ [rescaling according to (2.2) with a shift by v_h]. By Assumption 2.1, and the rescaling (2.2), one has

$$\bigl|\kappa_p^{(\varepsilon)}\bigr| \lesssim \varepsilon^{-p|\mathfrak{s}|/2}.$$

By Lemma 3.5, one also has $\kappa_p^{(\varepsilon)} = 0$ unless all of its arguments are located within a parabolic ball of radius $p\varepsilon$ [in fact, given a constant v, one can show that for $\varepsilon > 0$ sufficiently small, if $|(t,x)| > 2\varepsilon$, then $|(t, x - vt)| > \varepsilon$].

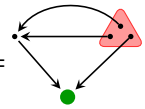
To define the renormalisation constants, we introduce some graphical notation which represent our integrations. In our graphs, a dot • represents an integration variable, and the special vertex ● represents the origin 0. Each arrow ⟶ represents the kernel $K'(y - x)$ with x and y being the starting and end points of the arrow, respectively. A red polygon with p dots inside, for instance, with $p = 4$, represents the cumulant function $\kappa_p^{(\varepsilon)}(z_1, \dots, z_p)$, with z_i given by the p

integration variables represented by the p dots. As an example,

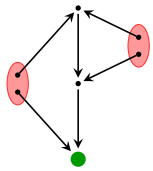
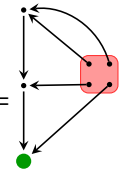


$$= \int_{\mathbf{R}^8} K'(-y)K'(-x_3)K'(y-x_1)K'(y-x_2) \times \kappa_3^{(\varepsilon)}(x_1,x_2,x_3)\,dx_1\,dx_2\,dx_3\,dy.$$

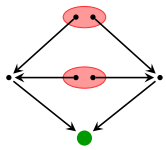
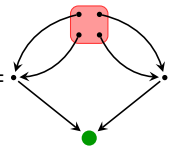
We then define a collection of renormalisation constants as follows. Here, we use the notation $Q^{(\varepsilon)} \approx \varepsilon^{-\alpha} + \sigma$ (where $\sigma \in \{0,1\}$) as a shorthand for the statement "there exist constants Q_1, Q_2 such that $\lim_{\varepsilon\to 0} |Q^{(\varepsilon)} - Q_1\varepsilon^{-\alpha} - Q_2\sigma| = 0$" [and call σ the $O(1)$ correction in the case $\sigma = 1$], and $Q^{(\varepsilon)} \approx c\log\varepsilon$ as a shorthand for "there exists a constant Q such that $\lim_{\varepsilon\to 0} |Q^{(\varepsilon)} - c\log\varepsilon - Q| = 0$". Define

$$C_0^{(\varepsilon)} =$$ $$\approx \varepsilon^{-1} + 1, \qquad C_1^{(\varepsilon)} =$$  $$\approx \varepsilon^{-\frac{1}{2}},$$

where the $O(1)$ correction in $C_0^{(\varepsilon)}$ is zero if κ_2 is even in the space variable (see Lemma 3.6 below), and $C_2^{(\varepsilon)} = 2C_{2,1}^{(\varepsilon)} + C_{2,2}^{(\varepsilon)}$ where

$$C_{2,1}^{(\varepsilon)} =$$  $$- \frac{(\hat{c}^{(\varepsilon)})^2}{2} \approx -\frac{\log\varepsilon}{32\pi\sqrt{3}}, \qquad C_{2,2}^{(\varepsilon)} =$$  $$\approx 1,$$

and $C_3^{(\varepsilon)} = 2C_{3,1}^{(\varepsilon)} + C_{3,2}^{(\varepsilon)}$ where

$$C_{3,1}^{(\varepsilon)} =$$  $$\approx \frac{\log\varepsilon}{8\pi\sqrt{3}}, \qquad C_{3,2}^{(\varepsilon)} =$$  $$\approx 1.$$

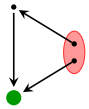
When v_h is zero, our constants $C_0^{(\varepsilon)}$, $C_{2,1}^{(\varepsilon)}$, $C_{3,1}^{(\varepsilon)}$ are the same as the renormalisation constants $C_0^{(\varepsilon)}$, $C_2^{(\varepsilon)}$, $C_3^{(\varepsilon)}$ in [26] where their divergence rates were obtained (just identify $\kappa_2^{(\varepsilon)}$ with the mollifier convolving with itself in [26]). However, since $v_h \neq 0$ in general,

$$\kappa_2^{(\varepsilon)}(t,x) \neq \hat{\kappa}^{(\varepsilon)}(t,x) \stackrel{\text{def}}{=} \varepsilon^{-3}\kappa_2\big(t/\varepsilon^2, x/\varepsilon\big), \tag{3.6}$$

so there may be lower order corrections to the asymptotic behaviors of the renormalisation constants obtained in [26]. We have the following lemma.

LEMMA 3.6. *For every sufficiently small $\varepsilon > 0$, there exists a choice of the constant $\hat{c}^{(\varepsilon)}$ such that*

$$\hat{c}^{(\varepsilon)} = \tag{3.7}$$



and such that as $\varepsilon \to 0$, $\hat{c}^{(\varepsilon)}$ converges to a finite limit $\hat{c}$. Furthermore, as $\varepsilon \to 0$ the right-hand side of (3.7) with $\hat{\kappa}^{(\varepsilon)}$ [which is defined in (3.6)] in place of $\kappa_2^{(\varepsilon)}$ also converges to $\hat{c}$. Finally, with this family of constants $\hat{c}^{(\varepsilon)}$, one has $C_0^{(\varepsilon)} \approx \varepsilon^{-1} + 1$ and the $O(1)$ correction is zero if κ_2 is even in the space variable.

We remark that (3.7) is indeed an *equation* for $\hat{c}^{(\varepsilon)}$ since on the right-hand side $\kappa_2^{(\varepsilon)}$ depends on $v_h = 4\lambda^2\hat{c}^{(\varepsilon)}$ [see (1.3) and (2.2)]. The proof of this lemma relies crucially on the following result.

LEMMA 3.7. *The function $\Theta\colon \mathbf{R}^2 \to \mathbf{R}$ given by*

$$\Theta(z) \stackrel{\text{def}}{=} \int K'(x)K'(x-y)K'(y-z)\,dx\,dy$$

is bounded.

PROOF. Since K' is an odd function in the space variable, the function $K_2(x) \stackrel{\text{def}}{=} \int K'(x-y)K'(y)\,dy$ is even in the space variable. We decompose the integral $\Theta(z) = \int K'(x)K_2(x-z)\,dx$ as integrals over three domains (similarly with the proof of [23], Lemma 10.14). The first domain is $\Omega_1 \stackrel{\text{def}}{=} \{|x| < 2|z|/3\}$. On this domain, one has $|K_2(x-z)| \lesssim |x-z|^{-1} \lesssim |z|^{-1}$, and the integration of $|K'(x)|$ over this domain is bounded by $|z|$, therefore, the integration over Ω_1 is bounded. The second domain is $\Omega_2 \stackrel{\text{def}}{=} \{|x-z| < 2|z|/3\}$, and the proof is analogous.

On the last domain $\Omega_3 = \mathbf{R}^2 \setminus (\Omega_1 \cup \Omega_2)$, if we replace $K_2(x-z)$ by $K_2(x)$, we have $|\int_{\Omega_3} K'(x)K_2(x)\,dx| \lesssim 1$. Indeed, since the integrand is odd in the space variable and Ω_1 is a domain that is symmetric under the space reflection, one has $\int_{\mathbf{R}^2\setminus\Omega_1} K'(x)K_2(x)\,dx = 0$; furthermore, one has $\int_{\Omega_2} |K'(x)K_2(x)|\,dx \lesssim 1$, so the claim follows. The difference $|K_2(x-z) - K_2(x)|$ is bounded by $|z||x|^{-2}$ using the gradient theorem, so the error caused by the replacement is bounded by $|z|\int_{\Omega_3} |K'(x)||x|^{-2}\,dx \lesssim |z||z|^{-1} = 1$. Collecting all the bounds above completes the proof for the boundedness of $\Theta(z)$. □

PROOF OF LEMMA 3.6. We can rewrite (3.7) as

$$\hat{c}^{(\varepsilon)} = \int \Theta(t,x)\hat{\kappa}^{(\varepsilon)}\big(t, x - 4\lambda^2\hat{c}^{(\varepsilon)}t\big)\,dt\,dx \stackrel{\text{def}}{=} F_\varepsilon\big(\hat{c}^{(\varepsilon)}\big). \tag{3.8}$$

It follows immediately from Lemma 3.7 and the properties of $\hat{\kappa}^{(\varepsilon)}$ that the function F_ε is bounded, uniformly in ε and in $\hat{c}^{(\varepsilon)}$.

By Assumptions 1.1 and 1.2, for sufficiently small $\varepsilon > 0$, $\partial_x \hat{\kappa}^{(\varepsilon)}$ is bounded by $C\varepsilon^{-4}\mathbf{1}_\varepsilon$ for some constant C, where $\mathbf{1}_\varepsilon$ is the characteristic function for the parabolic ball of radius ε. Combining this with the fact that $|t| \lesssim \varepsilon^2$ for (t, x) in the support of $\hat{\kappa}^{(\varepsilon)}$, we conclude that F'_ε is bounded by $C\varepsilon$ for some C, again uniformly in $\varepsilon < 1$ and in $\hat{c}^{(\varepsilon)}$. It follows that, provided that ε is sufficiently small, so that $|F'_\varepsilon| < 1/2$, (3.8) has indeed a unique solution [because the function $x - F_\varepsilon(x)$ is monotonically increasing in x with derivative bounded below by a strictly positive number], and that this solution converges to

$$F_\star = \lim_{\varepsilon \to 0} F_\varepsilon(0),$$

as $\varepsilon \to 0$. In fact, the existence of this limit $F_\star$ is straightforward to show. The right-hand side of (3.8) can be rewritten as $F_\varepsilon(0)$ plus an error which is bounded by $|\int_0^{\hat{c}^{(\varepsilon)}} F'_\varepsilon(x)\,dx| \lesssim \varepsilon|\hat{c}^{(\varepsilon)}|$. Since F_ε is bounded uniformly in ε, so is $\hat{c}^{(\varepsilon)}$; therefore this error vanishes, and taking limits on both sides of (3.8) shows that $\hat{c}^{(\varepsilon)}$ converges to $F_\star$.

Regarding the statement for $C_0^{(\varepsilon)}$, by the scaling argument in [26], if we define $\hat{C}_0^{(\varepsilon)}$ to be the same as $C_0^{(\varepsilon)}$ except that $\kappa_2^{(\varepsilon)}$ is replaced by $\hat{\kappa}^{(\varepsilon)}$, then $\hat{C}_0^{(\varepsilon)} \approx \varepsilon^{-1}$. Note that

$$C_0^{(\varepsilon)} - \hat{C}_0^{(\varepsilon)} = \int_0^{4\lambda^2 \hat{c}^{(\varepsilon)} t} \int_{\mathbf{R}^4} K'(x) K'(x - z)\, \partial_x \hat{\kappa}^{(\varepsilon)}(t, x - \theta)\, d\theta\, dx\, dz \tag{3.9}$$

and that $|\partial_x \hat{\kappa}^{(\varepsilon)}| \lesssim \varepsilon^{-4}\mathbf{1}_\varepsilon$, $|t| \lesssim \varepsilon^2$ on the support of $\hat{\kappa}^{(\varepsilon)}$, and $\hat{c}^{(\varepsilon)}$ is bounded, $C_0^{(\varepsilon)} - \hat{C}_0^{(\varepsilon)}$ converges to a finite limit. Therefore, there exists a constant Q such that

$$\lim_{\varepsilon \to 0} |C_0^{(\varepsilon)} - C_0 \varepsilon^{-1} - Q| = 0.$$

If κ_2 is even in the space variable, then since K' is odd in the space variable one has $F_\varepsilon(0) = 0$, so $\lim_{\varepsilon \to 0} \hat{c}^{(\varepsilon)} = 0$ and the order 1 correction Q vanishes, thus completing the proof. □

REMARK 3.8. In fact, the right-hand side of (3.9) divided by λ^2 converges to a limit independent of λ, which can be shown in an analogous way as the fact that $\lim_{\varepsilon \to 0} F_\varepsilon(\theta)$ does not depend on θ. Since $C_0^{(\varepsilon)}$ appears in the renormalised equation (3.3) in the form $\lambda C_0^{(\varepsilon)}$ [see (3.10) below], it turns out that the $O(1)$ correction to $C_0^{(\varepsilon)}$ contributes to the term $\lambda^3 c$ in (1.3).

Regarding the other renormalisation constants, when $v_h = 0$, it was also shown in [26], Theorem 6.5, that there exists a constant $\bar{c} \in \mathbf{R}$ depending on both κ_2 and the choice of the cutoff kernel K such that $\lim_{\varepsilon \to 0}(4C_{2,1}^{(\varepsilon)} + C_{3,1}^{(\varepsilon)}) = \bar{c}$. The prefactors in front of the logarithmically divergent constants were obtained in [22].

Following similar arguments as in [25] and [26], one has $C_1^{(\varepsilon)} \approx \varepsilon^{-1/2}$, with the limiting prefactor C_1 equal to an integral which can be represented by the same graph as the one for $C_1^{(\varepsilon)}$, except that each arrow is understood as P' and the red triangle is interpreted as κ_3 (i.e., setting $\varepsilon = 1$). It is easy to see that this integral is finite since κ_3 is continuous and P' is integrable, in particular C_1 does not depend on the choice of K. In a similar way $C_{2,2}^{(\varepsilon)}$ and $C_{3,2}^{(\varepsilon)}$ converge to finite constants $C_{2,1}$, $C_{3,2} \in \mathbf{R}$ depending on κ_4, but independent of the choice of the kernel K. These are all consequences of the scaling property of the heat kernel and the fact that K converges to heat kernel in suitable spaces under parabolic scaling. Following analogous arguments as in the proof of Lemma 3.6, one can show that the next order corrections to these constants due to the nontrivial shift v_h are $O(\varepsilon^{1/2})$ and thus all vanish in the limit.

The specification of these renormalisation constants (in particular with $\hat{c}^{(\varepsilon)}$ given by Lemma 3.6) determines an element $M_\varepsilon \in \mathfrak{R}$ as in (3.1) by setting

$$\ell_1 = C_0^{(\varepsilon)}, \qquad \ell_2 = \hat{c}^{(\varepsilon)}, \qquad \ell_3 = C_1^{(\varepsilon)}, \qquad \ell_4 = C_2^{(\varepsilon)}, \qquad \ell_5 = C_3^{(\varepsilon)}. \tag{3.10}$$

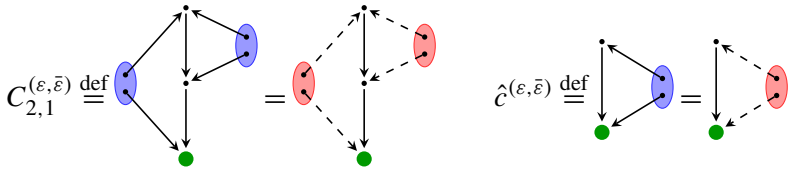
Given $\varepsilon > 0$ and a realisation of ζ_ε, this yields a renormalised model $(\hat{\Pi}^{(\varepsilon)}, \hat{\Gamma}^{(\varepsilon)})$ by acting on the canonical model $(\Pi^{(\varepsilon)}, \Gamma^{(\varepsilon)})$ lifted from ζ_ε by M_ε.

REMARK 3.9. One may wonder whether the constants $C_{2,2}^{(\varepsilon)}$ and $C_{3,2}^{(\varepsilon)}$ which involve the fourth cumulants *have to* be subtracted from the objects and . Note that although their homogeneities are zero, they converge to finite constants rather than diverge logarithmically. To obtain some limit, one does of course not have to subtract them, but we do subtract them here in order to obtain a limit which is independent (in law) of the choice of process ζ.

4. A priori bounds on the renormalised models. In the previous section, we have defined the renormalised models $\hat{Z}_\varepsilon = (\hat{\Pi}^{(\varepsilon)}, \hat{\Gamma}^{(\varepsilon)})$. The goal of this section is to obtain uniform bounds on the moments of this family of renormalised models. Actually, we will obtain a bit more, namely we obtain bounds also on a smoothened version of this model which will be useful in Section 6 when we show convergence of the models $\hat{Z}_\varepsilon$ to a limiting Gaussian model. More precisely, we define a second family of renormalised random models $\hat{Z}_{\varepsilon,\bar{\varepsilon}} = (\hat{\Pi}^{(\varepsilon,\bar{\varepsilon})}, \hat{\Gamma}^{(\varepsilon,\bar{\varepsilon})})$ in the same way as $\hat{Z}_\varepsilon$, except that the field ζ_ε used to build $\hat{Z}_\varepsilon$ is replaced by $\zeta_{\varepsilon,\bar{\varepsilon}} \stackrel{\text{def}}{=} \zeta_\varepsilon * \varrho_{\bar{\varepsilon}}$, where ϱ is a compactly supported smooth function on $\mathbf{R}^2$ integrating to 1 that is even in the space variable, and $\varrho_{\bar{\varepsilon}} = \bar{\varepsilon}^{-3}\varrho(\bar{\varepsilon}^{-2}t, \bar{\varepsilon}^{-1}x)$.

The renormalisation constants used to build the models $\hat{Z}_{\varepsilon,\bar{\varepsilon}}$ are still defined as in Section 3, except that they are now dependent on cumulants of $\zeta_{\varepsilon,\bar{\varepsilon}}$, so we denote them by $\hat{c}^{(\varepsilon,\bar{\varepsilon})}$ and $C_i^{(\varepsilon,\bar{\varepsilon})}$ with $i = 0, \dots, 3$. Note that while $\hat{c}^{(\varepsilon)}$ was obtained implicitly by solving the equation (3.7), we now consider ζ_ε (and therefore v_h) as given and view (3.7) as the definition of $\hat{c}^{(\varepsilon)}$ with given right-hand side.

For instance, $C_{2,1}^{(\varepsilon,\bar\varepsilon)}$ and $\hat c^{(\varepsilon,\bar\varepsilon)}$ are given by



respectively, where the dashed arrows represent $K'_{\bar\varepsilon}$, with $K_{\bar\varepsilon} = K * \varrho_{\bar\varepsilon}$, and represents the covariance of $\zeta_{\varepsilon,\bar\varepsilon}$. (This covariance depends on $v_h = \hat c^{(\varepsilon)}$, but not on $\hat c^{(\varepsilon,\bar\varepsilon)}$.) It is straightforward to verify that the kernel $K_{\bar\varepsilon}$ approximates the kernel K in the sense that

$$\big|K(z) - K_{\bar\varepsilon}(z)\big| \lesssim \bar\varepsilon^{\eta}|z|^{-1-\eta}, \qquad \big|K'(z) - K'_{\bar\varepsilon}(z)\big| \lesssim \bar\varepsilon^{\eta}|z|^{-2-\eta}, \tag{4.1}$$

for some sufficiently small $\eta > 0$. Our main result is the following a priori bound on $\hat Z^{(\varepsilon)}$ and on its difference with $\hat Z^{(\varepsilon,\bar\varepsilon)}$.

THEOREM 4.1. *Let* $(\hat\Pi^{(\varepsilon)}, \hat\Gamma^{(\varepsilon)})$ *and* $(\hat\Pi^{(\varepsilon,\bar\varepsilon)}, \hat\Gamma^{(\varepsilon,\bar\varepsilon)})$ *be as above. There exist* $\kappa, \eta > 0$ *such that for every* $\tau \in \{\Xi, \mathcal{I}, \mathcal{I}\Xi, \mathcal{I}(\tau_1\tau_2), \mathcal{I}^2\Xi, \mathcal{I}(\tau_1\mathcal{I}\tau_2)\}$ *and every* $p > 0$, *one has*

$$\begin{aligned} \mathbf{E}\big|\big(\hat\Pi_x^{(\varepsilon)}\tau\big)\big(\varphi_x^{\lambda}\big)\big|^p &\lesssim \lambda^{p(|\tau|+\eta)}, \\ \mathbf{E}\big|\big(\hat\Pi_x^{(\varepsilon)}\tau - \hat\Pi_x^{(\varepsilon,\bar\varepsilon)}\tau\big)\big(\varphi_x^{\lambda}\big)\big|^p &\lesssim \bar\varepsilon^{\kappa}\lambda^{p(|\tau|+\eta)} \end{aligned} \tag{4.2}$$

uniformly in all $\varepsilon, \bar\varepsilon \in (0,1]$, *all* $\lambda \in (0,1]$, *all test functions* $\varphi \in \mathcal{B}$, *and all* $x \in \mathbf{R}^2$.

REMARK 4.2. The first bound of (4.2) actually implies tightness for the family of renormalised models $\hat Z_\varepsilon = (\hat\Pi^{(\varepsilon)}, \hat\Gamma^{(\varepsilon)})$. The proof for this Kolmogorov type tightness criterion is similar to that of [23], Theorem 10.7. Since we do not really need to know tightness here, we will not elaborate on this point.

Note that although $\mathcal{I}$ and $\mathcal{I}\Xi$ both have negative homogeneities, they are of the form $\mathcal{I}(\tau)$, so that the corresponding bounds follow from the extension theorem [23], Theorem 5.14, combined with the analogous bounds on τ. The remainder of this section is devoted to the proof of (4.2) for each $\tau \in \{\mathcal{I}, \mathcal{I}\Xi, \mathcal{I}(\tau_1\tau_2), \mathcal{I}^2\Xi, \mathcal{I}(\tau_1\mathcal{I}\tau_2)\}$. In fact, we will mainly focus on the first bound of (4.2), and the second bound will follow analogously. We are going to use extensively the notion of Wick product for a collection of (not necessarily Gaussian) random variables which is defined as follows.

DEFINITION 4.3. Given a collection $\mathcal{X} = \{X_\alpha\}_{\alpha\in\mathcal{A}}$ of random variables as before, the Wick product $:X_A:$ for $A \subset \mathcal{A}$ is defined recursively by setting $\mathbf{E}:X_\emptyset: = 1$ and postulating that

$$X^A = \sum_{B\subset A} :X_B: \sum_{\pi\in\mathcal{P}(A\setminus B)} \prod_{\bar B\in\pi} \mathbf{E}_c(X_{\bar B}). \tag{4.3}$$

We will also sometimes write $:\prod_{i\in B} X_i: \stackrel{\text{def}}{=} :X_B:$, keeping in mind that $:\prod -:$ is really *one single* notation rather than composition of two consecutive operations.

Note that by (3.4) we could also write $X^A = \sum_{B\subset A} :X_B: \mathbf{E}(X^{A\setminus B})$, but (4.3) is actually the identity frequently being used in the paper. Again, (4.3) is sufficient to define $:X_A:$ by recursion over the size of A. Indeed, the term with $B = A$ is precisely the quantity we want to define, and all other terms only involve Wick products of strictly smaller sets of random variables. By the definition, we can easily see that as soon as $A \neq \emptyset$, one always has $\mathbf{E}:X_A: = 0$. Note also that if we take expectations on both sides of (4.3), then all the terms with $B \neq \emptyset$ vanish and we obtain exactly the identity (3.4). For example, given centred random variables X_i, we have $:X_i: = X_i$, $:X_1X_2: = X_1X_2 - \mathbf{E}(X_1X_2)$,

$$:X_1X_2X_3: = X_1X_2X_3 - \sum_{i\neq j\neq k} X_i\mathbf{E}(X_jX_k) - \mathbf{E}(X_1X_2X_3).$$

We refer to [4], [5], Appendix B, or [30] for the properties of non-Gaussian Wick products. The following result (Lemma 4.5) is often called "diagram formula" in the literature. As far as we know, the original proof of this particular statement was given by [35] and [19], but similar statements were known long before that. Before stating the lemma, we define an important type of partition.

DEFINITION 4.4. Let $M \times P$ be the Cartesian product of two sets M and P, and $D \subset M \times P$. We say that $\pi \in \mathcal{P}_M(D)$ if $\pi \in \mathcal{P}(D)$ and for every $B \in \pi$, there exist $(i,k), (i',k') \in B$ such that $k \neq k'$ (in particular $|B| > 1$).

In plain words, the requirement says that for each $B \in \pi$, the elements in B can not all have the same "P-index". In most cases, we are interested in the situation $D = M \times P$.

LEMMA 4.5. *Let* $M \stackrel{\text{def}}{=} \{i : 1 \le i \le m\}$ *and* $P \stackrel{\text{def}}{=} \{k : 1 \le k \le p\}$. *Let* $X_{(i,k)}$ *where* $1 \le i \le m$, $1 \le k \le p$ *be centred random variables with bounded moments of all orders. One has*

$$\mathbf{E}\Bigl(\prod_{k=1}^{p} :\prod_{i=1}^{m} X_{(i,k)}:\Bigr) = \sum_{\pi\in\mathcal{P}_M(M\times P)} \prod_{B\in\pi} \mathbf{E}_c(X_B). \tag{4.4}$$

Note that if the Wick products were replaced by usual products (say in the case $m = 1$), one would recover the definition of cumulants (3.4). In the case of centred Gaussians, only the partitions π with $|B| = 2$ for each $B \in \pi$ is allowed since joint cumulants of higher order vanish, and the statement recovers the fact that (4.4) can be computed by summing over all the ways of pairwise contracting $|M \times P|$ variables excluding "self-contractions" [i.e. factors of the form $\mathbf{E}_c(X_{(i,k)}, X_{(i',k)})$].

As an immediate consequence of Lemma 4.5 one has the following.

COROLLARY 4.6. *For every $p \geq 1$ and $m \geq 1$, one has*

$$\mathbf{E}\Biggl(\prod_{k=1}^{p} :\zeta_\varepsilon\bigl(x_k^{(1)}\bigr)\cdots\zeta_\varepsilon\bigl(x_k^{(m)}\bigr):\Biggr) = \sum_{\pi\in\mathcal{P}_M(M\times P)}\prod_{B\in\pi}\kappa_{|B|}^{(\varepsilon)}\bigl(\{x_k^{(i)} : (i,k)\in B\}\bigr), \tag{4.5}$$

where $M \stackrel{\text{def}}{=} \{i : 1 \leq i \leq m\}$ and $P \stackrel{\text{def}}{=} \{k : 1 \leq k \leq p\}$.

From now on we define a further simplified notation. Given a family of space–time points $\{x_\alpha : \alpha \in B\}$ parametrised by the index set B, we write

$$\kappa_{|B|}^{(\varepsilon)}(B) \stackrel{\text{def}}{=} \kappa_{|B|}^{(\varepsilon)}\bigl(\{x_\alpha : \alpha \in B\}\bigr).$$

Here, α can be multi-indices, for instance x_α would be $x_k^{(i)}$ in the situation of (4.5).

According to Lemma 3.5, some factors $\kappa_{|B|}^{(\varepsilon)}$ in (4.5) may actually vanish. In fact, a partition π encodes the information that the space–time points that are arguments of the function $\kappa_{|B|}^{(\varepsilon)}$ for each element $B \in \pi$ must be "close to each other" in order to have nonzero contribution to (4.5).

4.1. *First object and some generalities.* We now have all the ingredients in place for the proof of Theorem 4.1. For the moment, we only show how to obtain the first bound of (4.2); we include a short discussion on how to obtain the second bound at the end of this section. We start with the case $\tau = \Xi$, for which where we need to show that $\lambda^{|\mathfrak{s}|p/2}\mathbf{E}|\zeta_\varepsilon(\varphi_0^\lambda)|^p \lesssim 1$ uniformly in $\varepsilon, \lambda > 0$ for all $p > 1$. As a consequence of Corollary 4.6 with $m = 1$, one has

$$\mathbf{E}|\zeta_\varepsilon(\varphi_0^\lambda)|^p = \sum_{\pi\in\mathcal{P}(\{1,\dots,p\})}\prod_{B\in\pi}\Biggl(\int \kappa_{|B|}^{(\varepsilon)}(B)\prod_{i\in B}\varphi_0^\lambda(x_i)\prod_{i\in B}dx_i\Biggr).$$

We bound the integral in the bracket for each $B \in \pi$, $|B| > 1$. If $\varepsilon > \lambda$, we bound $|\kappa_{|B|}^{(\varepsilon)}| \lesssim \varepsilon^{-|B||\mathfrak{s}|/2} \lesssim \lambda^{-|B||\mathfrak{s}|/2}$, and the fact that the integral of the test functions φ_0^λ are uniformly bounded. If on the other hand $\varepsilon \leq \lambda$, then we use the fact that the integrand vanishes unless all points x_i are within a ball of radius $C\varepsilon$ for some fixed $C > 0$, and this ball is centred around a point in the ball of radius λ around the origin. The volume of this region is bounded by $C\varepsilon^{(|B|-1)|\mathfrak{s}|}\lambda^{|\mathfrak{s}|}$. Since the integrand on the other hand is bounded by $C\varepsilon^{-|B||\mathfrak{s}|/2}\lambda^{-|B||\mathfrak{s}|}$, it follows that the integral is bounded by $C\varepsilon^{(|B|/2-1)|\mathfrak{s}|}\lambda^{(1-|B|)|\mathfrak{s}|} \lesssim \lambda^{-|B||\mathfrak{s}|/2}$. Since $\sum_{B\in\pi}|B| = p$, the desired bound follows.

We now turn to the simplest nontrivial object, namely $\tau = \mathcal{V}$. Using Definition 4.3 to rewrite the product $\zeta_\varepsilon(x^{(1)})\zeta_\varepsilon(x^{(2)})$ as $:\zeta_\varepsilon(x^{(1)})\zeta_\varepsilon(x^{(2)}):$ +

$\kappa_2^{(\varepsilon)}(x^{(1)}, x^{(2)})$, together with the definition of $C_0^{(\varepsilon)}$, one has

$$(\hat{\Pi}_0^{(\varepsilon)}\mathbb{V})(\varphi_0^\lambda) = \int \varphi_0^\lambda(z) K'(z - x^{(1)}) K'(z - x^{(2)}) :\zeta_\varepsilon(x^{(1)})\zeta_\varepsilon(x^{(2)}): dx^{(1)}\, dx^{(2)}\, dz. \tag{4.6}$$

The term containing $\kappa_2^{(\varepsilon)}$ is precisely canceled by $C_0^{(\varepsilon)}$. The pth moment for any even number p of the above quantity can be written as

$$\int \prod_{k=1}^{p} (\varphi_0^\lambda(z_k) K'(z_k - x_k^{(1)}) K'(z_k - x_k^{(2)})) \mathbf{E} \prod_{k=1}^{p} (:\zeta_\varepsilon(x_k^{(1)})\zeta_\varepsilon(x_k^{(2)}):)\, dx\, dz,$$

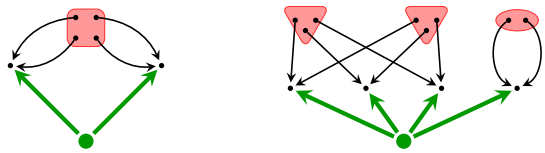
where the integration is over all $x_k^{(i)}$ and all z_k with $i \in \{1, 2\}$ and $1 \le k \le p$. Applying Corollary 4.6, one has the bound

$$\mathbf{E}|(\hat{\Pi}_0^{(\varepsilon)}\mathbb{V})(\varphi_0^\lambda)|^p \lesssim \sum_\pi \int \prod_{k=1}^{p} |\varphi_0^\lambda(z_k) K'(z_k - x_k^{(1)}) K'(z_k - x_k^{(2)})| \prod_{B\in\pi} |\kappa_{|B|}^{(\varepsilon)}(B)|\, dx\, dz, \tag{4.7}$$

where π runs over partitions in $\mathcal{P}_{\{1,2\}}(\{1, 2\} \times \{1, \dots, p\})$.

As illustrations, the left picture below shows the situation for $p = 2$ and π being the particular partition consisting of only one element which is the whole set of cardinality 4. The right picture below shows the situation for $p = 4$ with a *prohibited* partition π [namely, $\pi \in \mathcal{P}(\{1, 2\} \times \{1, \dots, p\})$ but $\pi \notin \mathcal{P}_{\{1,2\}}(\{1, 2\} \times \{1, \dots, p\})$], that is, one that does *not* appear in the sum (4.7). Graphically, the fact that the partition is prohibited is reflected in the presence of a subgraph identical to the one representing the diverging constant $C_0^{(\varepsilon)}$:

(4.8)



Here, the interpretation of these graphs is as in Section 3, but with the green arrow ⟶ representing the test function φ_0^λ.

The right-hand side of (4.7) involves an integration of functions of the following general form; we formulate it in a general way in order to deal with other objects $\tau \in \mathcal{W}_-$ as well. We are given a finite set of space–time points $B = \{x_1, \dots, x_n\}$, and a cumulant function $\kappa_{|B|}^{(\varepsilon)}(B) = \kappa_n^{(\varepsilon)}(x_1, \dots, x_n)$. By observing the right-hand side of (4.7), one realises that for every $1 \le i \le n$, there is a factor $|K'(y_i - x_i)|$ for some space–time point y_i in the integrand, which is bounded by $|y_i - x_i|^\alpha$ for some α. Here, it is allowed that some of the points $y_1, \dots, y_n$ may coincide. In the

case of (4.7) above, y_i would be one of the points z_k and $\alpha = -2$. We then have the following result for integrating such a product of cumulants and kernels (see Section 2.2 for definition of a scaling $\mathfrak{s}$ of $\mathbf{R}^d$).

LEMMA 4.7. *Given a scaling $\mathfrak{s}$ of $\mathbf{R}^d$, n points $y_1, \dots, y_n \in \mathbf{R}^d$, and real numbers $\alpha_1, \dots, \alpha_n$, assuming $-|\mathfrak{s}| < \alpha_i < 0$ for all $1 \le i \le n$, one has*

$$\int \prod_{i=1}^{n} |y_i - x_i|^{\alpha_i} \bigl|\kappa_n^{(\varepsilon)}(x_1, \dots, x_n)\bigr| \, dx_1 \cdots dx_n \lesssim \varepsilon^{(n/2-1)|\mathfrak{s}|} \int_{\mathbf{R}^d} \prod_{i=1}^{n} \bigl(|y_i - x| + \varepsilon\bigr)^{\alpha_i} \, dx, \tag{4.9}$$

uniformly in $y_1, \dots, y_n \in \mathbf{R}^d$ and $\varepsilon > 0$. Here, the integral on the left-hand side is over n variables, and the integral on the right-hand side is over only one variable.

REMARK 4.8. Lemma 4.7 is one of the main technical results in this article and will be frequently applied. The intuitive "graphical" meaning of this lemma is that for all practical purposes, we can "collapse" the p points in an expression of the type (here $p = 4$) into one single point, and this operation then generates a factor $\varepsilon^{(p/2-1)|\mathfrak{s}|}$. It is because this exponent is positive as soon as $p > 2$ that our bounds behave "as if" everything was Gaussian.

REMARK 4.9. If we consider κ_n to have homogeneity $-n|\mathfrak{s}|/2$, $|y - x|^\alpha$ to have homogeneity α, ε to have homogeneity 1, and each integration variable to have homogeneity $|\mathfrak{s}|$, then the total homogeneity is identical on both sides of (4.9). (It is equal to $n|\mathfrak{s}|/2 + \sum \alpha_i$ in both cases.)

PROOF. First of all, we define $\mathbf{1}_a(x) = 1$ if $|x| < a$ and vanishes if $|x| \ge a$. Recall then that one has the bound

$$\bigl|\kappa_n^{(\varepsilon)}(x_1, \dots, x_n)\bigr| \lesssim \varepsilon^{(n/2-1)|\mathfrak{s}|} \sum_{T} \prod_{\{x_i, x_j\} \in \mathcal{E}(T)} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_\varepsilon(x_i - x_j), \tag{4.10}$$

uniformly in $\varepsilon > 0$, where T runs over all spanning trees of the set $\{x_1, \dots, x_n\}$, and $\mathcal{E}(T)$ denotes the edge set of T. Indeed, for each T, since $|\mathcal{E}(T)| = n - 1$, the total power of ε on the right-hand side is $(n/2-1)|\mathfrak{s}| - (n-1)|\mathfrak{s}| = -n|\mathfrak{s}|/2$ which is consistent with the naive bound $|\kappa_n^{(\varepsilon)}| \lesssim \varepsilon^{-n|\mathfrak{s}|/2}$. By Lemma 3.5, $\kappa_n^{(\varepsilon)} = 0$ unless there exists a connected graph over $\{x_1, \dots, x_n\}$ such that $|x_i - x_j| < \varepsilon$ whenever $\{x_i, x_j\}$ is an edge of the graph. The bound (4.10) is then obtained by taking a spanning tree of this graph.

Let T be a fixed tree in the above sum and $e = \{x_i, x_j\}$ be an arbitrary fixed edge of T. There exists a constant $c > 1$ such that

$$\varepsilon^{-|\mathfrak{s}|} \mathbf{1}_\varepsilon(x_i - x_j) \lesssim \int_{\mathbf{R}^d} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{c\varepsilon}(x_i - x) \cdot \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{c\varepsilon}(x - x_j) \, dx \tag{4.11}$$

uniformly in $\varepsilon > 0$. This is obviously true for $\varepsilon = 1$, and the case for general $\varepsilon > 0$ follows by rescaling. We can therefore add a new "dummy variable" x representing the integration variable in (4.11), which suggests to define a new tree T' by $T' = T \cup \{x\}$ and to replace the edge e by the two edges $\{x_i, x\}$ and $\{x, x_j\}$. In other words, one sets $\mathcal{E}(T') = (\mathcal{E}(T) \setminus \{e\}) \cup \{\{x_i, x\}, \{x, x_j\}\}$.

Writing now $\mathbf{1}_\varepsilon(e) \stackrel{\text{def}}{=} \mathbf{1}_\varepsilon(x_i, x_j)$ as a shorthand if $e = \{x_i, x_j\}$, we note that $\prod_{e \in \mathcal{E}(T')} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{c\varepsilon}(e)$ is the indicator function of the event "every edge of T' has length at most $c\varepsilon$". Since the graph T' has finite diameter, this implies that every vertex of T' is at distance at most $C\varepsilon$ from the fixed vertex x, for some fixed $C > 0$, so that

$$\prod_{e \in \mathcal{E}(T)} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_\varepsilon(e) \le \int_{\mathbf{R}^d} \prod_{e \in \mathcal{E}(T')} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{c\varepsilon}(e)\, dx \le \int_{\mathbf{R}^d} \prod_{\bar{x} \in T} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{C\varepsilon}(x - \bar{x})\, dx,$$

where the first inequality is given by (4.11).

It now remains to substitute this back into (4.10) and into the left-hand side of (4.9). Integrating over all the variables $x_i \in T' \setminus \{x\}$, the claim follows from the bound

$$\int_{\mathbf{R}^d} |y_i - \bar{x}|^{\alpha_i} \varepsilon^{-|\mathfrak{s}|} \mathbf{1}_{C\varepsilon}(x - \bar{x})\, d\bar{x} \lesssim \big(|y_i - x| + \varepsilon\big)^{\alpha_i},$$

which can easily be shown by considering separately the two cases $|y_i - x| \le 2C\varepsilon$ and $|y_i - x| \ge 2C\varepsilon$. □

REMARK 4.10. In the above proof, if one chose x to be a fixed point in T, which would seem more natural in principle than the first step (4.11), the resulting bound would have one of the factors on the right-hand side of (4.9) equal to $|y_i - x|^{\alpha_i}$ (i.e. without the additional ε appearing in the statement). These occurrences of ε in our bound will however be useful in the sequel, as it will allow us to use the brutal bound $(|y_i - x| + \varepsilon)^{\alpha_i} \le \varepsilon^{\alpha_i}$.

We introduce the following notation: for any space–time point x and $\varepsilon \ge 0$, let

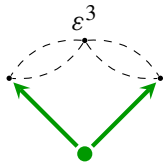
$$|x|_\varepsilon \stackrel{\text{def}}{=} |x| + \varepsilon,$$

where $|x|$ is as above the norm of x with respect to the parabolic scaling.

We can now apply Lemma 4.7 to each π and each $B \in \pi$ of the right-hand side of (4.7), so that the entire right-hand side of (4.7) is bounded by

$$\sum_\pi \int \Bigg(\prod_{k=1}^p \varphi_0^\lambda(z_k) \Bigg) \prod_{B \in \pi} \Bigg(\varepsilon^{(|B|/2-1)|\mathfrak{s}|} \prod_{(i,k) \in B} |z_k - x_B|_\varepsilon^{-2} \Bigg) \times \prod_{B \in \pi} dx_B \prod_{k=1}^p dz_k, \tag{4.12}$$

where $\pi \in \mathcal{P}_{\{1,2\}}(\{1,2\} \times \{1,\dots,p\})$ as above, and for each $B \in \pi$ we have an integration variable $x_B \in \mathbf{R}^2$. For example, in the case of the left picture in (4.8), we have $\pi = \{B\}$ with $B = \{(1,1),(1,2),(2,1),(2,2)\}$, and the corresponding term in (4.12) can be depicted graphically as



with dashed lines representing the function $x \mapsto |x|_\varepsilon^{-2}$.

REMARK 4.11. Bounds of the form (4.12) are crucial to our analysis. As already noted in Remark 4.9, although the number of integration variables appearing in (4.12) differs for each π, a simple power counting shows that the "homogeneity" of the summand is always the same, namely

$$\sum_{B\in\pi}\big((|B|/2-1)|\mathfrak{s}|-2|B|\big)+\sum_{B\in\pi}|\mathfrak{s}|=-\sum_{B\in\pi}|B|/2=-p.$$

In a certain sense, the powers of ε thus guarantee that the entire quantity is still of the correct "homogeneity".

In the rest of this subsection, we explain some generalities for bounding expressions of the type (4.12), and in the next several subsections we will follow the same routine for the other objects $\tau \in \mathcal{W}_-$. Our idea is to apply the general tools developed in Section 5, which are summarised in Proposition 4.18 below. To apply this proposition to obtain our desired bound, we will associate to $\mathcal{V}$ a "partial graph" H (in general, several such graphs for other elements $\tau \in \mathcal{W}_-$) as defined in Definition 4.12 below. Then we represent the multi-integral for each partition π in the expressions of the type (4.12), which result from application of Corollary 4.6 and Lemma 4.7, a graph $\mathcal{V}$ obtained by "Wick contracting" p copies of H, in the sense of Definition 4.13 below. The set of all allowed partitions in (4.12) is then in one-to-one correspondence with the set of possible "Wick contractions". Given an integer $p > 1$, a rescaled test function φ_0^λ, and a partial graph H, one then has a number $\Lambda^p_{\varphi_0^\lambda}(\mathcal{V})$ for every such graph $\mathcal{V}$ which equals the value of the multi-integral, and the sum of them over all Wick contractions then yields a number $\Lambda^p_{\varphi_0^\lambda}(H)$ which provides a bound for the moment (4.2). In a nutshell, the "routine" we will always follow is that

$$\mathbf{E}\big|\big(\hat{\Pi}_x^{(\varepsilon)}\tau\big)\big(\varphi_x^\lambda\big)\big|^p \overset{(4.5)\ (4.9)}{\lesssim} \Lambda^p_{\varphi_0^\lambda}(H) \overset{\text{Proposition 4.18}}{\lesssim} \lambda^{\bar{\alpha}}. \tag{4.13}$$

In general (see the next subsections), the quantity in the middle may be replaced by a sum over several graphs, and before applying (4.9) one may have to deal with

the renormalisations. Now we start to define the above mentioned terminologies precisely.

DEFINITION 4.12. A partial graph H is a connected graph with each edge e labelled by a real number m_e, which satisfies the following requirements. There exists a unique distinguished vertex $0 \in H$, as well as a unique distinguished edge $e = \{0, v_\star\}$ attached to 0 with label $m_e = 0$. (Other edges may also connect to 0 but they are not called distinguished edges.) We define $H_0 \stackrel{\text{def}}{=} H \setminus \{0\}$ and $H_\star \stackrel{\text{def}}{=} \{0, v_\star\}$. The set H_0 can be decomposed as a union of two disjoint subsets of vertices $H_0 = H_{\text{ex}} \cup H_{\text{in}}$, such that $\deg(v) = 1$ for every $v \in H_{\text{ex}}$ and $\deg(v) \geq 2$ for every $v \in H_{\text{in}} = H_0 \setminus H_{\text{ex}}$. We call a vertex in H_{ex} an "external vertex" and the only edge attached to it an "external edge". We call a vertex in H_{in} an "internal vertex", and the edges which are not distinguished or external are called "internal edges". We require that $v_\star \in H_{\text{in}}$.

When we draw a partial graph, we use a special dot ● to represent the distinguished vertex 0, a black dot • to represent a generic vertex which is not the 0, a special thick line —— to represent the distinguished edge $\{0, v_\star\}$. The labels m_e are often drawn on the edges e, such as —2+—, understood as $m_e = 2 + \delta$ for sufficiently small $\delta > 0$. As some examples of partial graphs, see for instance, (4.15), (4.21) or (4.26).

DEFINITION 4.13. Suppose that we are given a partial graph H and an integer $p > 1$, and for each $1 \leq i \leq p$ let $H^{(i)}$ be a copy of the graph H. We say that a graph $\mathcal{V}$ is obtained from H by *Wick contracting* p copies of H, if there exists an equivalence relation $\sim$ over the disjoint union of the p copies $\bigcup_{i=1}^{p} H^{(i)}$, such that each equivalence class B is one of the following three forms:

- $B = \{0^{(1)}, \ldots, 0^{(p)}\}$ where $0^{(i)}$ is the distinguished vertex of $H^{(i)}$; or
- $B = \{v\}$ where $v \in \bigcup_{i=1}^{p} H_{\text{in}}^{(i)}$; or
- $B = \{v_1, \ldots, v_{|B|}\} \subset \bigcup_{i=1}^{p} H_{\text{ex}}^{(i)}$ with $|B| > 1$, with requirement that there exist at least two elements $v, v' \in B$ such that $v \in H_{\text{ex}}^{(i)}$, $v' \in H_{\text{ex}}^{(i')}$ and $i \neq i'$;

and such that $\mathcal{V} = \bigcup_{i=1}^{p} H^{(i)} / \sim$ (namely, we identify all the vertices in each equivalence class as one vertex). The set of edges $\mathcal{E}'(\mathcal{V})$ of $\mathcal{V}$ is the union of all the sets of edges of the p copies of H, and each edge $e \in \mathcal{E}'(\mathcal{V})$ of $\mathcal{V}$ naturally inherits a label m_e from H. We still call the equivalence class consisting of the distinguished vertices 0, and call an equivalence class consisting of an internal vertex (resp. some external vertices) an in-vertex (resp., an ex-vertex), and call an edge $e \in \mathcal{E}'(\mathcal{V})$ an in-edge (resp., ex-edge) if e is an internal (resp., external) edge of H. We say that an in-vertex belongs to a certain copy $H^{(i)}$ if the only element in this equivalence class is an internal vertex of $H^{(i)}$. We write $\mathcal{V} = \mathcal{V}_{\text{ex}} \cup \mathcal{V}_{\text{in}} \cup \{0\}$ where $\mathcal{V}_{\text{ex}}$ is the set of ex-vertices and $\mathcal{V}_{\text{in}}$ is the set of in-vertices. Let $\mathcal{V}_0 = \mathcal{V} \setminus \{0\}$ and $\mathcal{V}_\star = \{0, v_\star^{(1)}, \ldots, v_\star^{(p)}\}$.

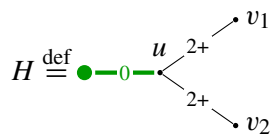
REMARK 4.14. Note that given (H, p) there could be many possible graphs $\mathcal{V}$ constructed as above. The graph $(\mathcal{V}, \mathcal{E}')$ may have multiple edges between two vertices. In the sequel, $\deg(v)$ for $v \in \mathcal{V}$ stands for the degree of v counting multiple edges.

REMARK 4.15. The graphical notation for a graph $\mathcal{V}$ in Definition 4.13 will be the same as that for a partial graph H. We remark that although we use the same solid line to represent all edges of $\mathcal{V}$, the functions associated with them will turn out to be slightly different: an in-edge will be associated with a singular function $|x|^{-m_e}$ while an ex-edge will be associated with a mollified function $|x|_\varepsilon^{-m_e}$. We will not distinguish them by complicating the graphical notation since it is always clear whether an edge is an in-edge or ex-edge.

Given a graph $\mathcal{V}$ as in Definition 4.13 and a test function φ, we can associate a number to it, as alluded above:

$$\Lambda^p_\varphi(\mathcal{V}) \stackrel{\text{def}}{=} \int \prod_{i=1}^{p} \varphi(x_{v_\star^{(i)}}) \prod_{v \in \mathcal{V}_{\text{ex}}} \varepsilon^{(\deg(v)/2-1)|\mathfrak{s}|} \times \prod_{e=\{v,\bar{v}\} \in \mathcal{E}'(\mathcal{V})} |x_v - x_{\bar{v}}|_{\varepsilon(e)}^{-m_e} \prod_{v \in \mathcal{V}_0} dx_v, \tag{4.14}$$

where the norm $|\cdot|_{\varepsilon(e)} = |\cdot|_\varepsilon$ if e is an ex-edge and $|\cdot|_{\varepsilon(e)} = |\cdot|$ if e is an in-edge. Now by observing the expressions (4.12) and (4.14) one realises that we can define a partial graph H with one internal vertex u and two external vertices v_1, v_2:

(4.15) $H \stackrel{\text{def}}{=}$ 

so that for every π in (4.12) the integral is equal to $\Lambda^p_{\varphi_0^\lambda}(\mathcal{V})$ for a graph $\mathcal{V}$ obtained from Wick contracting p copies of H. Indeed, every in-vertex (resp., ex-vertex) of $\mathcal{V}$ corresponds to an integration variable z_k (resp., x_B) in (4.12). Define

$$\Lambda^p_{\varphi_0^\lambda}(H) \stackrel{\text{def}}{=} \sum_{\mathcal{V}} \Lambda^p_{\varphi_0^\lambda}(\mathcal{V}), \tag{4.16}$$

where the sum is over all graphs $\mathcal{V}$ obtained from Wick contracting p copies of H. Then by Definition 4.4 (for the allowed partitions) and Definition 4.13 (for the allowed Wick contractions), (4.12) is equal to $\Lambda^p_{\varphi_0^\lambda}(H)$.

Proposition 4.18 below will then yield the desired bound on $\Lambda^p_{\varphi_0^\lambda}(H)$. Before stating the proposition, we need the following definitions.

DEFINITION 4.16. Given a graph $(\mathcal{V}, \mathcal{E})$ we define for any subgraph $\bar{\mathcal{V}} \subset \mathcal{V}$ the following subsets of $\mathcal{E}$:

$$\mathcal{E}_0(\bar{\mathcal{V}}) = \{e \in \mathcal{E} : e \cap \bar{\mathcal{V}} = e\}, \qquad \mathcal{E}(\bar{\mathcal{V}}) = \{e \in \mathcal{E} : e \cap \bar{\mathcal{V}} \neq \emptyset\}.$$

DEFINITION 4.17. Given $\bar{H} \subset H_0$ such that $\bar{H}_{\mathrm{ex}} \stackrel{\mathrm{def}}{=} \bar{H} \cap H_{\mathrm{ex}} \neq \emptyset$, define number $c_e(\bar{H})$ for edges e of H as follows. If $|\bar{H}_{\mathrm{ex}}| = 1$, then $c_e(\bar{H}) = 0$ for all e. If $|\bar{H}_{\mathrm{ex}}| = 2$, then assuming that $\bar{H}_{\mathrm{ex}} = \{v, \bar{v}\}$, there are two cases. The first case is that v and $\bar{v}$ are connected with the same vertex of H, and we define $c_e(\bar{H}) = 3/4$ for every $e \in \mathcal{E}(\bar{H}_{\mathrm{ex}})$; the second case is that v and $\bar{v}$ are connected with different vertices of H, and we define $c_e(\bar{H}) = 1/2$ for every $e \in \mathcal{E}(\bar{H}_{\mathrm{ex}})$. Finally, if $|\bar{H}_{\mathrm{ex}}| > 2$, then

$$c_e(\bar{H}) = \frac{3}{2} - \frac{3}{|\bar{H}_{\mathrm{ex}}| + 1}, \tag{4.17}$$

for every $e \in \mathcal{E}(\bar{H}_{\mathrm{ex}})$. Given any $\bar{H} \subset H$ such that $0 \in \bar{H}$, we define $c_e(\bar{H}) = 3/2$ for every $e \in \mathcal{E}(\bar{H}_{\mathrm{ex}})$. For any subset $\bar{H} \subset H$ we set $c_e(\bar{H}) = 0$ for every $e \notin \mathcal{E}(\bar{H}_{\mathrm{ex}})$.

PROPOSITION 4.18. *Given a partial graph H, suppose that the following conditions hold*:

(1) *For every subset $\bar{H} \subset H$ with $|\bar{H}| \geq 2$ and $|\bar{H}_{\mathrm{in}}| \geq 1$, one has*

$$\sum_{e \in \mathcal{E}_0(\bar{H})} \big(m_e - c_e(\bar{H})\big) < 3\big(|\bar{H}_{\mathrm{in}}| - \mathbf{1}_{\bar{H} \subset H_{\mathrm{in}}}\big). \tag{4.18}$$

(2) *For every nonempty $\bar{H} \subset H \setminus H_\star$, one has*

$$\sum_{e \in \mathcal{E}(\bar{H})} m_e > 3\bigg(|\bar{H} \cap H_{\mathrm{in}}| + \frac{1}{2}|\bar{H} \cap H_{\mathrm{ex}}|\bigg). \tag{4.19}$$

Then, for every $p > 1$, one has

$$\sup_{\varphi \in \mathcal{B}} \Lambda^p_{\varphi_0^\lambda}(H) \lesssim \lambda^{\bar{\alpha} p}, \tag{4.20}$$

where

$$\bar{\alpha} = 3\big(|H_{\mathrm{ex}}|/2 + |H_{\mathrm{in}} \setminus H_\star|\big) - \sum_{e \in \mathcal{E}(H)} m_e.$$

PROOF. This follows from Corollary 5.15 below, together with Remark 5.10 that the number $c_e(\bar{H})$ in Definition 4.17 yields an admissible ε-allocation rule in the sense of Definition 5.5, and the fact that $\Lambda^p_{\varphi_0^\lambda}(H) \lesssim \Lambda^p_{\varphi_0^\lambda}(\mathcal{V})$ with proportionality constant depending only on the cardinality of H and on p [see (4.16)]. □

REMARK 4.19. By Remark 5.12, if the partial graph H is such that $\deg(0) = 1$, namely the distinguished edge $e_\star = \{0, v_\star\}$ with label $m_{e_\star} = 0$ is the only edge connected to 0, then one does not need to check (4.18) for the subsets $\bar{H}$ such that $0 \in \bar{H}$.

LEMMA 4.20. *The conditions of Proposition 4.18 are satisfied for H defined in (4.15).*

PROOF. Since $\deg(0) = 1$, we only check (4.18) for $\bar{H} \subset H_0$. If $\bar{H} = \{u, v_1\}$, then since $c_e(\bar{H}) = 0$ condition (4.18) reads $2 + \delta < 3$ where $\delta > 0$ is sufficiently small. The case $\bar{H} = \{u, v_2\}$ follows in the same way by symmetry.

Still considering condition (4.18), we now look at subgraph $\bar{H} = \{u, v_1, v_2\}$. On the left-hand side of (4.18), the set $\mathcal{E}_0(\bar{H})$ consists of two edges. Since $\bar{H} \cap H_{\mathrm{ex}}$ consists of two points which are connected with the same point (i.e., u) in H, by Definition 4.17, $c_e(\bar{H}) = 3/4$ for each $e \in \mathcal{E}_0(\bar{H})$. So condition (4.18) reads $5/2 + 2\delta < 3$ for sufficiently small $\delta > 0$.

Regarding condition (4.19), if $\bar{H} = \{v_1\}$ then condition (4.19) reads $2 + \delta > \frac{3}{2}$, and the case $\bar{H} = \{v_2\}$ follows in the same way. If $\bar{H} = \{v_1, v_2\}$, then condition (4.19) reads $4 + 2\delta > 3$. So (4.19) is verified. □

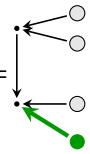
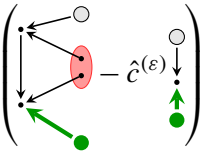
Proposition 4.18 then yields the desired bound (with $\bar{\alpha} = -1 - \kappa$ for some arbitrarily small $\kappa > 0$) on $\Lambda^p_{\varphi_0^\lambda}(H)$, and therefore Theorem 4.1 is proved for the case $\tau = \Vee$.

REMARK 4.21. While for the simplest object $\Vee$ we have directly applied Lemma 4.7 to the bound (4.7), for the more complicated objects, we will often have to deal with renormalisations *before* applying the absolute values as in (4.7).

4.2. *Bounds on the other objects.*

Bounds on $\Vee\!\!\cdot$. With the generalities discussed in the previous subsection, we now proceed to consider the object $\tau = \Vee\!\!\cdot$. Recall that we have defined some graphical notation from Section 3 to represent integrations of kernels. Besides these graphical notation, we will need another type of special vertices ○ in our graphs. Each instance of ○ stands for an integration variable x, as well as a factor $\zeta_\varepsilon(x)$. Furthermore, if more than one such vertex appear, then the corresponding product of ζ_ε is always a Wick product $:\zeta_\varepsilon(x_1) \cdots \zeta_\varepsilon(x_n):$, where the x_i are the integration variables represented by *all* of the special vertices ○ appearing in the graph. This is consistent with the graphical notation used in [25, 26]: our graphs represent integrals of product of kernels and a (non-Gaussian) Wick product of the random fields. In the particular case where ζ_ε is a Gaussian random field, this yields an element of the nth homogeneous Wiener chaos as in [25, 26], and the two notation do coincide.

With these notation, by Definition 4.3 and the definition of $C_1^{(\varepsilon)}$, combined with the fact that K annihilates constants, we have the identity:

$$(\hat{\Pi}_0^{(\varepsilon)}\Psi)(\varphi_0^\lambda) =$$  $$+ 2\Big($$  $$=: \mathcal{I}_1 + 2\mathcal{I}_2.$$

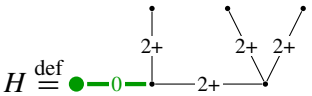
For instance,

$$\mathcal{I}_1 = \int \varphi_0^\lambda(z) K'(w - x^{(1)}) K'(w - x^{(2)}) K'(z - x^{(3)})$$

$$\times K'(z - w) : \prod_{i=1}^{3} \zeta_\varepsilon(x^{(i)}) : dx\, dw\, dz,$$

where $x = (x^{(1)}, x^{(2)}, x^{(3)})$. Note that in principle one would expect the appearance of a term involving $\kappa_3^{(\varepsilon)}$ in the expression for $\hat{\Pi}_0^{(\varepsilon)}\Psi$, but this term is precisely cancelled by $C_1^{(\varepsilon)}$.

We bound the pth moments of $\mathcal{I}_1$ and $\mathcal{I}_2$ separately. For the moment $\mathbf{E}|\mathcal{I}_1|^p$ of the first integral $\mathcal{I}_1$, we can apply Corollary 4.6 and Lemma 4.7. We are then again in the situation that there is a natural partial graph H with two internal vertices and three external vertices:

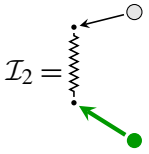
(4.21) 

so that $\mathbf{E}|\mathcal{I}_1|^p \lesssim \Lambda^p_{\varphi_0^\lambda}(H)$. As in the case $\tau = \Psi$, it is straightforward to verify the conditions of Proposition 4.18. We only check the condition (4.18) with $\bar{H}$ consisting of all the vertices of H except 0 (so $|\bar{H}| = 5$). Since $|\bar{H} \cap H_{\mathrm{ex}}| = 3$, by (4.17) in Definition 4.17 we have $c_e(\bar{H}) = 3/4$ for each of the three external edges, so (4.18) holds since the left-hand side is $8 + 4\delta - \frac{3}{4} \cdot 3$ while the right-hand side is equal to 6. We remark that the fact that the labels in the graph are actually $2 + \delta$ for some *strictly positive* δ guarantees the strict inequality in the condition (4.19). We therefore have the desired bound on the moments of $\mathcal{I}_1$ by Proposition 4.18 (with $\bar{\alpha} = -1/2 - \kappa$ for some arbitrarily small $\kappa > 0$).

Regarding the second integral $\mathcal{I}_2$, we define a function

$$Q_\varepsilon(w) \stackrel{\mathrm{def}}{=} \int K'(w - x) K'(-y) K'(-w) \kappa_2^{(\varepsilon)}(x, y)\, dx\, dy - \hat{c}^{(\varepsilon)} \delta(w). \tag{4.22}$$

We use the notation ⟿ for the function Q_ε. One then has

(4.23) $\mathcal{I}_2 =$ 

By [22], Lemma 10.16, one has

$$\left|\int K'(w-x)Q_\varepsilon(z-w)\,dw\right| \lesssim |z-x|^{-2}, \tag{4.24}$$

uniformly over $\varepsilon \in (0,1]$. Note that the renormalised distribution $\mathscr{R}Q_\varepsilon$ appearing in [22], Lemma 10.16, is precisely the same as Q_ε, because Q_ε integrates to 0 by the choice of $\hat{c}^{(\varepsilon)}$ in Lemma 3.6.

As an immediate consequence of (4.24), one has for all even p

$$\mathbf{E}|\mathcal{I}_2|^p \lesssim \int \prod_{i=1}^p \bigl(\varphi_0^\lambda(z_i)|z_i-x_i|^{-2}\bigr)\left|\mathbf{E}\Bigl(\prod_{i=1}^p \zeta_\varepsilon(x_i)\Bigr)\right| dx\,dz, \tag{4.25}$$

where the integration is over $x_1,\dots,x_p$ and $z_1,\dots,z_p$. Again, after applying Corollary 4.6 and Lemma 4.7, we can define a partial graph H with one internal vertex and one external vertex as

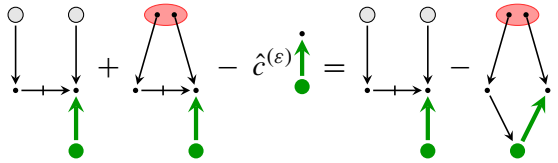
$H \stackrel{\text{def}}{=}$

(4.26)

so that (4.25) is bounded by $\Lambda^p_{\varphi_0^\lambda}(H)$. It is then very straightforward to check that the conditions of Proposition 4.18 are all satisfied. Therefore, we have the desired bound on the moments of $\mathcal{I}_2$.

This completes the proof of Theorem 4.1 for the case $\tau = \Vee$. Note that in the bound of $\mathbf{E}|\mathcal{I}_2|^p$, we actually integrated out *some of* the integration variables before applying the first step in the "routine" (4.13) and representing the rest of the multi-integral by graphs obtained by Wick contracting an H. This procedure will be useful sometimes for other objects.

Bounds on $\curlyvee$. To study the case $\tau = \curlyvee$, we introduce an additional graphical notation $\longrightarrow$, which is the same as in [25] and [26]. This barred arrow represents $K'(z-w)-K'(-w)$ where w and z are the coordinates of the starting and end point, respectively. With this notation at hand, we have

$(\hat{\Pi}_0^{(\varepsilon)}\curlyvee)(\varphi_0^\lambda) =$ 

(4.27)

where we used the definition of $\hat{c}^{(\varepsilon)}$ (i.e., Lemma 3.6) in the second equality.

The second term on the right-hand side of (4.27) is deterministic and it is bounded by $\lambda^{-\delta}$ for every $\delta > 0$ (for instance, using Lemma 4.7 and [23], Lemma 10.14). So it remains to bound the first term, which is equal to

$$\int_{\mathbf{R}^8} \varphi_0^\lambda(z)K'\bigl(w-x^{(1)}\bigr)K'\bigl(z-x^{(2)}\bigr) \times \bigl(K'(z-w)-K'(-w)\bigr):\zeta_\varepsilon\bigl(x^{(1)}\bigr)\zeta_\varepsilon\bigl(x^{(2)}\bigr):dx\,dw\,dz,$$

where the x-integration is over $x^{(1)}, x^{(2)}$. This integral can be written as a sum of two terms $\mathcal{I}_1' + \mathcal{I}_2'$, where $\mathcal{I}_i'$ denotes the integration over Ω_i for $i \in \{1, 2\}$, with

$$\Omega_1 \stackrel{\text{def}}{=} \big\{(x^{(1)}, x^{(2)}, w, z) \in \mathbf{R}^8 : |w| \geq |z|/2\big\},$$

and $\Omega_2 \stackrel{\text{def}}{=} \mathbf{R}^8 \setminus \Omega_1$. We then bound the pth moment of each of these two terms separately.

By the generalised Taylor theorem in [23], Proposition 11.1,

$$\big|K'(z-w) - K'(-w)\big| \lesssim \begin{cases} |z|^{1/2}|z-w|^{-5/2}, & \text{on } \Omega_1, \\ |z|^{1/2-\eta}|w|^{-5/2+\eta}, & \text{on } \Omega_2 \end{cases} \tag{4.28}$$

for any $\eta \in [0, 1/2]$. To see (4.28), by Taylor's theorem one has

$$\big|K'(z-w) - K'(-w)\big| \lesssim |z|^{\frac{1}{2}-\eta}\big(|z-w|^{-\frac{5}{2}+\eta} + |w|^{-\frac{5}{2}+\eta}\big)$$

for any $\eta \in [0, 1/2]$. If $|w| \geq |z|/2$, one has $|z-w| \lesssim |z| + |w| \lesssim |w|$, and thus the first bound in (4.28) by choosing $\eta = 0$. If $|w| < |z|/2$, then $|w| \lesssim |z-w|$ and one obtains the second bound in (4.28).

On the support of $\varphi_0^\lambda(z)$, one has the additional bound $|z| \leq \lambda$. One then has

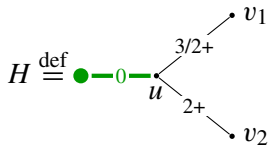
$$\begin{aligned} \mathbf{E}|\mathcal{I}_1'|^p &\lesssim \lambda^{\frac{p}{2}} \int_{\Omega_1 \times \cdots \times \Omega_1} \prod_{i=1}^p |\varphi_0^\lambda(z_i)| \big|K'\big(w_i - x_i^{(1)}\big)K'\big(z_i - x_i^{(2)}\big)\big| \\ &\quad \times |z_i - w_i|^{-5/2} \Big|\mathbf{E}\Big(\prod_{i=1}^p :\zeta_\varepsilon\big(x_i^{(1)}\big)\zeta_\varepsilon\big(x_i^{(2)}\big):\Big)\Big| \, dx\, dw\, dz \\ &\lesssim \lambda^{\frac{p}{2}} \int_{\mathbf{R}^{6p}} \prod_{i=1}^p |\varphi_0^\lambda(z_i)| |z_i - x_i^{(2)}|^{-2} |z_i - x_i^{(1)}|^{-\frac{3}{2}} \\ &\quad \times \Big|\mathbf{E}\Big(\prod_{i=1}^p :\zeta_\varepsilon\big(x_i^{(1)}\big)\zeta_\varepsilon\big(x_i^{(2)}\big):\Big)\Big| \, dx\, dz. \end{aligned} \tag{4.29}$$

Note that after applying (4.28), we bounded the integral over the Cartesian product of p copies of Ω_1 by the integration over all of $(\mathbf{R}^2)^{4p}$. In the last step, we then integrated out w_i for all $1 \leq i \leq p$.

REMARK 4.22. In the bound (4.28), just like in (4.24), we deal with renormalisations "by hand", no matter whether they are "negative renormalisation" [meaning the situations in which we have kernels with too small homogeneity such as Q in (4.22)] or "positive renormalisation" [meaning the situations in which we have to subtract the kernel at the origin as in (4.28)], rather than directly relying on the general bounds in [26]. We follow this approach because, when we apply the bounds on the cumulants (e.g., Lemma 4.7), each factor in the integrand is

bounded by its absolute value, but the renormalisations encode cancellations that would then be lost. Another reason for following this approach is that Assumption 5.2 below will be simpler to verify than the original assumption in [26].

Now we apply Corollary 4.6 and Lemma 4.7, which for $\mathbf{E}|\mathcal{I}_1'|^p$ gives the same bound as (4.12), times $\lambda^{p/2}$, except that some factors $|z_k - x_B|^{-2}$ are replaced by $|z_k - x_B|^{-3/2}$. This bound is again represented for each π by a graph obtained from Wick contracting p copies of H, where H is given by

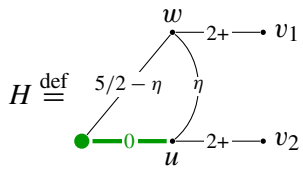


One can directly check that the conditions of Proposition 4.18 are all satisfied. In fact, the choice of exponent $-5/2$ in (4.28) was designed so that, on one hand the function $|z_i - w_i|^{-5/2}$ is still integrable when we integrate out w_i in (4.29), and on the other hand condition (4.19) is satisfied for the subgraph $\bar{H} = \{v_1\}$. This is because the left-hand side of (4.19) is then equal to $3/2 + \delta$ for some small $\delta > 0$, while the right-hand side is equal to $3/2$. We conclude that we have the desired bound on $\mathbf{E}|\mathcal{I}_1'|^p$ by Proposition 4.18 with $\bar{\alpha} = -1/2 - \kappa$ for arbitrarily small $\kappa > 0$.

Regarding the term $\mathcal{I}_2'$, we note that on the domain Ω_2 and the support of the function φ_0^λ, one has $|z|^{1/2-\eta} \lesssim \lambda^{1/2}|w - z|^{-\eta}$. Therefore, one has

$$\begin{aligned}\mathbf{E}|\mathcal{I}_2'|^p \lesssim \lambda^{\frac{p}{2}} \int_{\Omega_2 \times \cdots \times \Omega_2} \prod_{i=1}^{p} \varphi_0^\lambda(z_i) \big|K'\big(w_i - x_i^{(1)}\big) K'\big(z_i - x_i^{(2)}\big)\big| \\ \times |z_i - w_i|^{-\eta} |w_i|^{-5/2+\eta} \Big|\mathbf{E}\Big(\prod_{i=1}^{p} :\zeta_\varepsilon\big(x_i^{(1)}\big)\zeta_\varepsilon\big(x_i^{(2)}\big):\Big)\Big| \, dx\, dw\, dz.\end{aligned} \tag{4.30}$$

Again, we bound the integral by the corresponding integral over all of $(\mathbf{R}^2)^{4p}$. After applying Corollary 4.6 and Lemma 4.7, we obtain integrals which are again represented by graphs obtained from Wick contracting the partial graph:

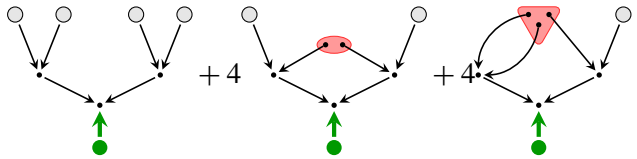


with $H_\star = \{0, u\}$. It is again straightforward to verify that the conditions of Proposition 4.18 are all satisfied, provided that we choose η and δ in a suitable way. Note that, due to existence of the edge between the vertex w and 0, we really do need

to check condition (4.18) for $\bar{H} \ni 0$. In fact, for the subgraph $\bar{H} = \{0, w, v_1\}$, the right-hand side of (4.18) is equal to 3, and since $c_{\{w,v_1\}}(\bar{H}) = 3/2$ the left-hand side is equal to $(5/2-\eta)+(2+\delta)-3/2 < 3$ for sufficiently small $\delta > 0$ such that $\delta < \eta$. The fact that the other subgraphs all satisfy condition (4.18) can be verified straightforwardly, so we omit the details.

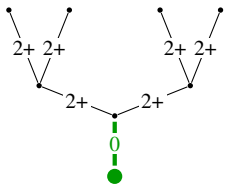
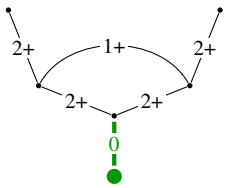
Regarding the condition (4.19), if the subgraph is given by $\bar{H} = \{w, v_1\}$, then the right-hand side of (4.19) equals $9/2$ and the left-hand side equals $(5/2-\eta)+(2+\delta)+\eta > 9/2$. For the other subgraphs, the condition follows similarly, and we conclude that the desired bound holds for $\mathbf{E}|\mathcal{I}'_2|^p$ by Proposition 4.18 with again $\bar{\alpha} = -1/2-\kappa$ for arbitrarily small $\kappa > 0$. The proof to the first bound of (4.2) is therefore completed for the case $\tau = \Xi$, noting that although one has $|\Xi| = -2\bar{\kappa}$ by (2.6), the presence of the factor $\lambda^{\frac{p}{2}}$ in front of the expressions (4.29) and (4.30) yields the required bound.

Bounds on $\mathcal{V}$. For the object $\mathcal{V}$, it is straightforward to check using (4.3) and definition of the renormalised model that

$(\hat{\Pi}_0^{(\varepsilon)}\mathcal{V})(\varphi_0^\lambda) =$  $+4$ $+4$

Note that the terms in which all the four leaves are contracted are canceled by the renormalisation constants $C_{3,1}^{(\varepsilon)}$ and $C_{3,2}^{(\varepsilon)}$. Here, the first two graphs are essentially the same as in [26] (except that our graphs represent non-Gaussian Wick products of random fields), but the last graph is new. It will be shown that the last term vanishes in the limit $\varepsilon \to 0$.

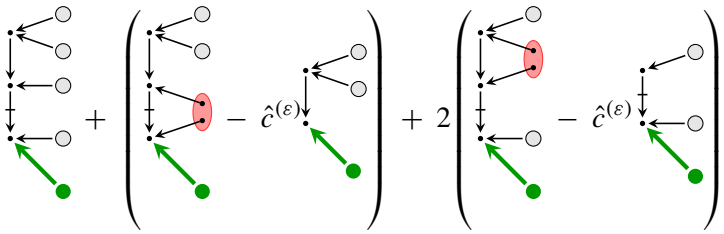
Denote by $\mathcal{I}_i^\dagger$ $(i = 1, 2, 3)$ the above terms, respectively. For $\mathcal{I}_1^\dagger$ and $\mathcal{I}_2^\dagger$, we can apply Corollary 4.6 and Lemma 4.7 to obtain the bound $\mathbf{E}|\mathcal{I}_i^\dagger|^p \lesssim \Lambda^p_{\varphi_0^\lambda}(H_i)$ where $i \in \{1, 2\}$ and H_1, H_2 are defined below. It is straightforward to check the conditions of Proposition 4.18. Note that for $\mathcal{I}_2^\dagger$, we have integrated out the second cumulant $\kappa_2^{(\varepsilon)}$, which gives the factor $|x-y|^{-1}$ (corresponding to the edge with label "1+" in H_2):

$H_1 =$  $H_2 =$  $H_3 =$

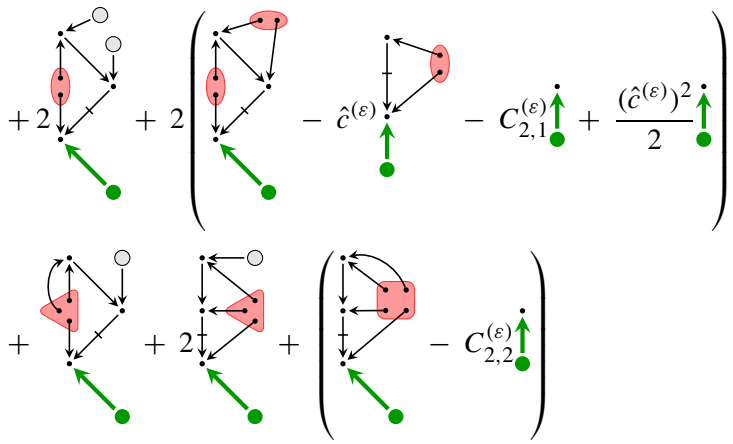
Regarding the term $\mathcal{I}_3^\dagger$, when we apply Lemma 4.7, we gain a factor $\varepsilon^{3/2}$ from the third cumulant $\kappa_3^{(\varepsilon)}$. It is clear from our graphical representation that there exist "double edges" connecting to the third cumulant function, which stand for a

function $(|x-w|+\varepsilon)^{-4}$. We can use a factor $\varepsilon^{1+\delta}$ with $0<\delta<1/2$ to improve the homogeneity of this function, namely, we have bound $\varepsilon^{1+\delta}(|x-w|+\varepsilon)^{-4} \lesssim (|x-w|+\varepsilon)^{-3+\delta}$; and we still have a factor $\varepsilon^{\frac{1}{2}-\delta}$ left. Integrating out all the variables represented by dots $\bullet$ (which are simple convolutions) except for the ones where the test functions φ_0^λ are evaluated at, we again obtain a bound $\Lambda^p_{\varphi_0^\lambda}(H_3)$ where the partial graph drawn as H_3 above (with the label $2-$ denoting $2-\delta$). It is easy to verify that the conditions of Proposition 4.18 are again satisfied for H_3. With the remaining factor $\varepsilon^{\frac{1}{2}-\delta}$, we see that as $\varepsilon\to 0$, the moments of this term converge to zero. The proof of the first bound of (4.2) is therefore completed for the case $\tau = \mathcal{V}$.

Bounds on $\mathcal{V}$. For the object $\mathcal{V}$, by Definition 4.3 for Wick products and the definition of the renormalised model $\hat\Pi^{(\varepsilon)}$, one has

$(\hat\Pi_0^{(\varepsilon)}\mathcal{V})(\varphi_0^\lambda) =$ 

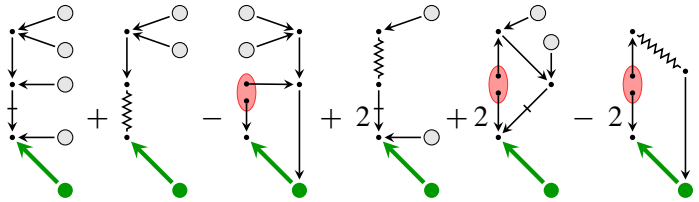
(4.31)



Here, the graphs in the first two lines are essentially the same as those in [26] (except for the renormalisation constant $\hat c^{(\varepsilon)}$), but the graphs in the last line are new due to the nontrivial higher cumulants of the random field ζ_ε. Note that the term involving the joint cumulant of ζ_ε at three points represented by the top three $\circ$ shaped vertices vanishes because the kernel K annihilates constants.

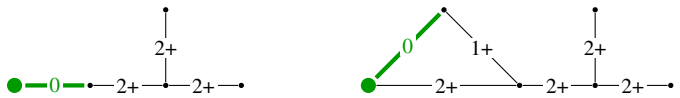
Using the notation $\wwwww$ for the kernel Q_ε defined in (4.22), and proceeding in the same way as [26] with the definition of constant $C_{2,1}^{(\varepsilon)}$, we see that the sum

of all the graphs appearing in *the first two lines* of (4.31) is equal to

(4.32) 

We denote by $\mathcal{I}_k^\star$ the kth term in (4.32) with $k = 1, \dots, 6$. Note that $\mathcal{I}_6^\star$ is deterministic and easily bounded by $\lambda^{-\eta}$ for any $\eta > 0$.

For the term $\mathcal{I}_2^\star$, we apply the bound (4.24), then it can be bounded using Corollary 4.6 and Lemma 4.7 followed by Proposition 4.18, with the partial graph shown in the left picture below. The term $\mathcal{I}_3^\star$ is bounded in the same way, with the partial graph shown in the right picture:

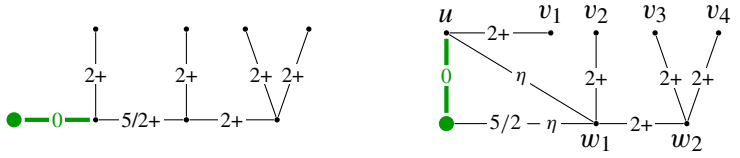


It is again straightforward to check the conditions of Proposition 4.18 for these partial graphs.

For the term $\mathcal{I}_4^\star$, we apply (4.24), then it is essentially the same as in the case $\tau = \Xi$. Regarding the term $\mathcal{I}_5^\star$, we can bound the two kernels in $\longrightarrow$ separately. In other words, we write $\mathcal{I}_5^\star$ as sum of two terms, each containing one of the two kernels. The moments of each term can then be bounded in the same way as above with partial graphs defined as follows, respectively:

The conditions of Proposition 4.18 are again verified.

We now bound moments of $\mathcal{I}_1^\star$. Using the bound (4.28) and proceeding in the same way as in the case Ξ, one can write $\mathcal{I}_1^\star$ as a sum of two terms, so that the moments of these two terms can be bounded separately. The partial graphs associated with these two terms are



These graphs still satisfy the conditions of Proposition 4.18. We only check here the conditions for some subgraphs of the second graph. For condition (4.18), if the subgraph $\bar{H} = \{0, w_1, w_2, v_2, v_3, v_4\}$, then the right-hand side of (4.18) is equal to 6, and the left-hand side is equal to $(5/2-\eta)+4(2+\delta)-3\cdot 3/2 < 6$ for sufficiently

small $\delta > 0$ such that $4\delta < \eta$. Regarding condition (4.19), if the subgraph $\bar{H} = \{w_1, w_2, v_2, v_3, v_4\}$, then the condition reads $(5/2 - \eta) + 4(2 + \delta) + \eta > 21/2$ for any $\delta > 0$. Proposition 4.18 (with $\bar{\alpha} = -1/2 - \kappa$ for some arbitrarily small $\kappa > 0$) then yields the desired bound on the pth moment, noting the overall factor $\lambda^{\frac{p}{2}}$ showing up as we proceed as in the case $\circlearrowleft$.

We proceed to consider the first graph in the last line of (4.31). Applying Lemma 4.7 to the third cumulant $\kappa_3^{(\varepsilon)}$ yields a factor $\varepsilon^{3/2}$. As before, we combine a factor $\varepsilon^{1+\eta}$ together with the "double-edge" to obtain an edge with label $3 - \eta$. Regarding the two terms in the barred edge, we deal with them separately. After integrating out some variables which are simple convolutions, the bounds boil down to straightforward verification of the conditions of Proposition 4.18 for the following two partial graphs:

Since there is the factor $\varepsilon^{1/2-\eta}$ remaining, the moments of this graph converge to zero.

With the same argument, the second graph in the last line of (4.31) also converges to zero. For the quantity in the last bracket in (4.31), after cancellation with the renormalisation constant it is equal to the same graph with the barred arrow replaced by an arrow pointing to the origin, and this graph, which represents a deterministic number, is bounded by $\varepsilon^{1-\delta}\lambda^{\delta-1}$ for sufficiently small $\delta > 0$ using [23], Lemma 10.14, and, therefore, converges to zero. We therefore obtain the desired bounds for all the terms.

PROOF OF THEOREM 4.1. Collecting all the results of this section, we obtain the first bound of (4.2), so it remains to show the second bound. Just as in the verification of the second bound in [23], Theorem 10.7, this follows in essentially the same way as the first bound. Indeed, as we consider the difference between $\hat{\Pi}_0^{(\varepsilon)}\tau$ and $\hat{\Pi}_0^{(\varepsilon,\bar{\varepsilon})}\tau$ for any $\tau \neq \Xi$, we obtain a sum of expressions of the type (4.6), but in each term some of the instances of K' are replaced by $K'_{\bar{\varepsilon}}$ and exactly one instance is replaced by $K' - K'_{\bar{\varepsilon}}$.

This is because $\zeta_{\varepsilon,\bar{\varepsilon}}$ always appears as part of an expression of the type $K' * \zeta_{\varepsilon,\bar{\varepsilon}}$, which can be rewritten as $K'_{\bar{\varepsilon}} * \zeta_{\varepsilon}$. We then use the fact that $K'_{\bar{\varepsilon}}$ satisfies the same bound as K', while $K' - K'_{\bar{\varepsilon}}$ satisfies the improved bound (4.1).

The case $\tau = \Xi$ is slightly different, since in this case $\zeta_{\varepsilon,\bar{\varepsilon}}$ appears on its own. However, the required bound there follows immediately from the fact that, on any fixed bounded domain, and for any $\alpha < 0$ and sufficiently small $\delta > 0$, and for any distribution η with $\|\eta\|_{\mathcal{C}^{\alpha+\delta}} < \infty$, one has

$$\|\eta - \varrho_{\bar{\varepsilon}} * \eta\|_{\mathcal{C}^{\alpha}} \lesssim \bar{\varepsilon}^{\delta}\|\eta\|_{\mathcal{C}^{\alpha+\delta}}.$$

The proof of Theorem 4.1 is therefore complete. □

5. General bounds. In this section, we derive some general results on bounding the type of integrals appearing in the previous section, for instance (4.12). These integrals are represented by labelled graphs which, after certain operations, fall into the scope of the general bounds in [26] (see also [25]). We view these labelled graphs as being obtained from "Wick contractions" of partial graphs, and the purpose of this section is to prove criteria on these partial graphs (i.e., Proposition 4.18) which are easy to verify and yield the desired bounds on the integrals.

We start with recalling the settings and results of the general a priori bound from [26], which should really be viewed as some version of the BPHZ theorem [9, 27, 37] in this context. The paper [26] considers labelled graphs $(\mathcal{V}, \mathcal{E})$, in which each vertex $v \in \mathcal{V}$ represents an integration variable x_v, except for a distinguished vertex $0 \in \mathcal{V}$ which represents the origin. Each edge $e = \{e_-, e_+\} \in \mathcal{E}$ is labelled by a number a_e and represents a kernel $J_e(x_{e_+} - x_{e_-})$ with homogeneity $-a_e$. There are p special vertices $v_{\star,1}, \dots, v_{\star,p}$ and p distinguished edges of the type $e = \{v_{\star,i}, 0\}$ with label $a_e = 0$, which represent factors $J_e(x_v - x_0) = \varphi_0^\lambda(x_v - 0)$.

REMARK 5.1. In [26], the edges $e = (v, \bar{v}) \in \mathcal{E}$ are oriented and decorated with labels $(a_e, r_e) \in \mathbf{R} \times \mathbf{Z}$ where r_e represents certain renormalisation procedure, and the orientation of an edge matters only if $r > 0$. In our case, since we always treat the renormalisations "by hand" before applying these general bounds, we ignore the label r_e (in other words r_e is always 0) and the orientation, so we only have a label a_e. [The orientations of the distinguished edges do not matter since we always just assume that they are associated with function $\varphi^\lambda(x_v - 0)$ rather than $\varphi^\lambda(0 - x_v)$.]

For every $a \in \mathbf{R}$, we define a (semi)norm on the space of compactly supported functions that are smooth everywhere, except at the origin:

$$\|J\|_a = \sup_{0<|x|_{\mathfrak{s}}\le 1} |x|_{\mathfrak{s}}^a \big|J(x)\big|.$$

We use the notation $\mathcal{V}_0 = \mathcal{V} \setminus \{0\}$ and $\mathcal{V}_\star = \{0, v_{\star,1}, \dots, v_{\star,p}\}$. With all of these notation at hand, a labelled graph as above, together with the corresponding collection of kernels determines a number

$$\mathcal{I}_\lambda(J) \stackrel{\text{def}}{=} \int_{(\mathbf{R}^d)^{\mathcal{V}_0}} \prod_{e\in\mathcal{E}} J_e(x_{e_+} - x_{e_-})\, dx, \tag{5.1}$$

where d is the space–time dimension, and we implicitly set $x_0 = 0$.

To bound the quantity $\mathcal{I}_\lambda(J)$, we will impose the following assumption. Recall that for a subgraph $\bar{\mathcal{V}}$, the sets $\mathcal{E}_0(\bar{\mathcal{V}})$ and $\mathcal{E}(\bar{\mathcal{V}})$ are defined in Definition 4.16.

ASSUMPTION 5.2. The labelled graph $(\mathcal{V}, \mathcal{E})$ satisfies the following properties:

1. For every subset $\bar{\mathcal{V}} \subset \mathcal{V}$ of cardinality at least 2, one has

$$\sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} a_e < |\mathfrak{s}|(|\bar{\mathcal{V}}| - 1). \tag{5.2}$$

2. For every nonempty subset $\bar{\mathcal{V}} \subset \mathcal{V} \setminus \mathcal{V}_\star$, one has the bounds

$$\sum_{e \in \mathcal{E}(\bar{\mathcal{V}})} a_e > |\mathfrak{s}||\bar{\mathcal{V}}|. \tag{5.3}$$

Note that in [26] there are four assumptions, but they reduce to the two given here since we assume that the labels r_e appearing in [26] are always 0. We then have the following result taken from [26].

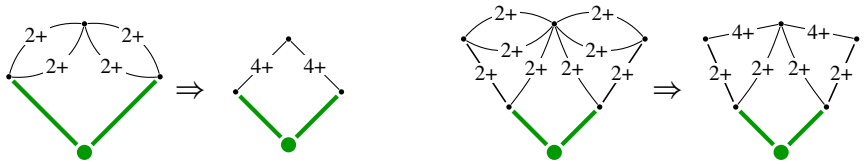
THEOREM 5.3. *Provided that Assumption 5.2 holds, there exists a constant C depending only on the number of vertices in $\mathcal{V}$ such that*

$$|\mathcal{I}_\lambda(J)| \le C\lambda^\alpha \prod_{e \in \mathcal{E}} \|J_e\|_{a_e}, \qquad \lambda \in (0, 1],$$

where $\alpha = |\mathfrak{s}||\mathcal{V} \setminus \mathcal{V}_\star| - \sum_{e \in \mathcal{E}} a_e$.

The rest of this section is devoted to verify that Assumption 5.2 indeed holds for the labelled graphs which represent our integrals, but after some necessary modifications which we explain now.

First of all, we are actually concerned with labelled graphs $(\mathcal{V}, \mathcal{E}')$ which are of the same type as the labelled graphs $(\mathcal{V}, \mathcal{E})$ defined above, except that there may exist more than one edge in $\mathcal{E}'$ between two vertices. More formally, $\mathcal{E}'$ is a multiset (i.e., it allows multiple instances of elements) of unordered pairs of vertices. [Such a graph $(\mathcal{V}, \mathcal{E}')$ is sometimes called a multigraph in the graph theory literature.] Indeed, for instance if we Wick contract two copies of the graph H defined in (4.15) [resp., in (4.21)] by identifying all the four (resp., six) external vertices as one equivalence class, then we obtain (multi)graphs on the left-hand sides of the arrows below, which clearly contain multi-edges:

(5.4) 

However, this issue can be easily resolved because given a labelled (multi)graph $(\mathcal{V}, \mathcal{E}')$, one can naturally define a labelled graph $(\mathcal{V}, \mathcal{E})$ by identifying all the multi-edges in $\mathcal{E}'$ between every pair of two vertices; the label a_e for $e = \{v, \bar{v}\} \in \mathcal{E}$ is simply defined as the sum of the labels $a_{e'}$ of all the edges $e' \in \mathcal{E}'$ between v and $\bar{v}$, see the right-hand sides of the arrows in (5.4). In the sequel, the prime in the

notation $\mathcal{E}'$ always indicates that the set $\mathcal{E}'$ allows multi-edges. For an edge $e \in \mathcal{E}'$, we will sometimes still write $e = \{v, \bar{v}\}$ or $v, \bar{v} \in e$, which simply means that e is an edge between the two vertices v and $\bar{v}$.

Another necessary modification is due to the following fact. One can verify that Assumption 5.2 fails to hold in the situations illustrated in the above two graphs. In fact, both examples contain subgraphs $\bar{\mathcal{V}}$ with $|\bar{\mathcal{V}}| = 2$ and $\sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} a_e = 4$, and the right example also contains subgraphs $\bar{\mathcal{V}}$ with $|\bar{\mathcal{V}}| = 3$ and $\sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} a_e = 8$. In both cases, (5.2) is violated, and one might worry that the corresponding integrals diverge. However, this will not cause any problem, since we will see that these graphs arise in situations where one is allowed to modify the labels of $\mathcal{E}'$ (and $\mathcal{E}$ accordingly) in such a way that Assumption 5.2 holds for the modified graph. In fact, by observing the expressions (4.12), it is clear that in the left example above, which arises in (4.15), one has an additional factor $\varepsilon^{(|B|/2-1)|\mathfrak{s}|} = \varepsilon^3$. Similarly, in the right example, which arises in (4.21), the factor is ε^6. In general for each contracted vertex v, we have a factor $\varepsilon^{3(\deg(v)/2-1)}$ arising from Lemma 4.7. These factors associated with the ex-vertices together with the fact that their neighboring ex-edges correspond to mollified functions (see Remark 4.15) can be exploited to improve the homogeneities of the edges attached to v and thus cure these "fake divergences". Here and in the sequel, $\deg(v)$ is always understood as the degree of v as a vertex in $(\mathcal{V}, \mathcal{E}')$ [rather than $(\mathcal{V}, \mathcal{E})$]; in other words $\deg(v)$ counts the multiple edges rather than regarding them as one edge.

Our main idea to verify that the Assumption 5.2 indeed holds for the objects we want to bound is based on the observation that the graphs $(\mathcal{V}, \mathcal{E}')$ are actually built by contracting simple "atomic" graphs called partial graphs (see Definition 4.12 and Definition 4.13), and we have precise knowledge on the structures of these partial graphs. We will then describe the rules to "allocate" the factors ε to the neighboring edges to cure the "fake" divergences, see Definition 5.5 below. Finally we will identify multi-edges and obtain graphs $(\mathcal{V}, \mathcal{E})$ which will be shown to satisfy Assumption 5.2.

We start with the following remark about some obvious properties of our graphs $(\mathcal{V}, \mathcal{E}')$.

REMARK 5.4. We will frequently use the following facts. Let $(\mathcal{V}, \mathcal{E}')$ be a graph constructed by Wick contracting p copies of H as in Definition 4.13. If v is an in-vertex, then $\deg(v)$ is equal to the degree of the corresponding internal vertex in H. If v is an ex-vertex, then $\deg(v)$ is equal to the cardinality of v, viewed as an equivalence class in H^p. By construction, there does not exist any edge connecting two in-vertices belonging to two different copies of H. Also, there does not exist any edge connecting two ex-vertices of $\mathcal{V}$.

DEFINITION 5.5. An *ε-allocation rule* is a way to assign, for every graph $(\mathcal{V}, \mathcal{E}')$ constructed according to the procedure in Definition 4.13 and every $v \in$

$\mathcal{V}_{\mathrm{ex}}$, a nonnegative real number $b_e^{(v)} \geq 0$ to every $e \in \{e \in \mathcal{E}' : v \in e\}$. An ε-allocation rule is called *admissible* if the following holds for every such graph $(\mathcal{V}, \mathcal{E}')$:

- For every $v \in \mathcal{V}_{\mathrm{ex}}$, one has

$$\sum_{e \ni v} b_e^{(v)} = \big(\deg(v)/2 - 1\big)|\mathfrak{s}|. \tag{5.5}$$

- For every $v \in \mathcal{V}_{\mathrm{ex}}$ and every $A \subset \{e : v \in e\}$, one has

$$\sum_{e \in A} b_e^{(v)} \geq \big(|A|/2 - 1\big)|\mathfrak{s}|. \tag{5.6}$$

- For any $v_1, v_2 \in \mathcal{V}_{\mathrm{ex}}$, $v_1 \neq v_2$, define a new graph $(\hat{\mathcal{V}}, \hat{\mathcal{E}}')$ by identifying v_1 and v_2 as one vertex w and apply the ε-allocation rule to $\hat{\mathcal{V}}$ to obtain numbers $\hat{b}_e^{(w)}$. Then the following monotonicity condition holds:

$$\hat{b}_e^{(w)} \geq b_e^{(v_1)} \quad (\forall e \ni v_1) \quad \text{and} \quad \hat{b}_e^{(w)} \geq b_e^{(v_2)} \quad (\forall e \ni v_2). \tag{5.7}$$

  (With the obvious identification of $\mathcal{E}'$ with $\hat{\mathcal{E}}'$.)

Note that by Remark 5.4 one necessarily has $\{e : v_1 \in e\} \cap \{e : v_2 \in e\} = \emptyset$. We set a convention that $b_e^{(v)} = 0$ if $v \notin e$.

REMARK 5.6. In general, the ε-allocation rule could be allowed to depend on the underlying partial graph H, as long as it is the same rule for all graphs $(\mathcal{V}, \mathcal{E}')$ built from a given H. However, in this article, we just fix one ε-allocation rule for all the graphs.

The following result will be used.

LEMMA 5.7. *Suppose that we are given an admissible ε-allocation rule. Then for any graph $\mathcal{V}$ constructed as in Definition 4.13, for any $v_1, v_2 \in \mathcal{V}_{\mathrm{ex}}$, $v_1 \neq v_2$ and any $Q_1 \subset \{e : v_1 \in e\}$, $Q_2 \subset \{e : v_2 \in e\}$, if we define a new graph $\hat{\mathcal{V}}$ by identifying v_1 and v_2 as one vertex w as in Definition 5.5 (so that $Q_1 \cup Q_2 \subset \{e : w \in e\}$), then*

$$\sum_{e \in Q_1} \hat{b}_e^{(w)} + \sum_{e \in Q_2} \hat{b}_e^{(w)} \leq \sum_{e \in Q_1} b_e^{(v_1)} + \sum_{e \in Q_2} b_e^{(v_2)} + |\mathfrak{s}|, \tag{5.8}$$

where $b_e^{(v_1)}$ and $b_e^{(v_2)}$ are numbers for $\mathcal{V}$ and $\hat{b}_e^{(w)}$ are numbers for $\hat{\mathcal{V}}$ when the ε-allocation rule is applied.

PROOF. Suppose that $Q_1 = \{e_1, \dots, e_{q_1}\}$ and $Q_2 = \{f_1, \dots, f_{q_2}\}$. Using (5.5), and the convention that $b_e^{(v)} = 0$ if $v \notin e$, the left-hand side minus the first two terms on the right-hand side of (5.8) is equal to

$$\Big(D_w - \sum_{e \notin Q_1 \cup Q_2} \hat{b}_e^{(w)}\Big) - \Big(D_{v_1} - \sum_{e \notin Q_1} b_e^{(v_1)}\Big) - \Big(D_{v_2} - \sum_{e \notin Q_2} b_e^{(v_2)}\Big), \tag{5.9}$$

where $D_u \stackrel{\text{def}}{=} (\deg(u)/2 - 1)|\mathfrak{s}|$ for any vertex $u \in \{w, v_1, v_2\}$ and the degree $\deg(u)$ is understood to be with respect to the graph $\mathcal{V}$ for $u \in \{v_1, v_2\}$ and with respect to the graph $\hat{\mathcal{V}}$ for $u = w$. Since the set $\{e : w \in e\} \setminus (Q_1 \cup Q_2)$ is equal to the disjoint union of the set $\{e : v_1 \in e\} \setminus Q_1$ and the set $\{e : v_2 \in e\} \setminus Q_2$, applying (5.7) yields

$$\sum_{e \notin Q_1} b_e^{(v_1)} + \sum_{e \notin Q_2} b_e^{(v_2)} - \sum_{e \notin Q_1 \cup Q_2} \hat{b}_e^{(w)} \le 0.$$

Therefore, (5.9) is bounded by $D_w - D_{v_1} - D_{v_2} = |\mathfrak{s}|$. □

With restrictions (5.5) and (5.7), there is not much freedom for an admissible ε-allocation rule. If $\deg(v) = 2$, then one must have $b_e^{(v)} = 0$ by (5.5). For the graphs arising in our analysis of the KPZ equation, we define an ε-allocation rule as follows:

- If $\deg(v) > 3$, or $\deg(v) = 3$ and there are 3 distinct vertices connected to v, then $b_e^{(v)} = (\deg(v)/2 - 1)|\mathfrak{s}|/\deg(v)$ for every e adjacent to v (i.e., the "even allocation" rule).
- If $\deg(v) = 3$ and there are only 2 distinct vertices connected to v then, calling $e_1, \bar{e}_1$ the two edges connecting v to the same vertex and e_2 the remaining edge, we set $b_{e_1}^{(v)} = b_{\bar{e}_1}^{(v)} = |\mathfrak{s}|/4$ and $b_{e_2}^{(v)} = 0$ (i.e. the "divergence priority" rule).

For the KPZ equation, it turns out that if $\deg(v) = 3$, then it is never the case that all three edges connect v to the same vertex.

LEMMA 5.8. *The above ε-allocation rule is admissible.*

PROOF. The equality in (5.5) always holds by definition. Regarding (5.7), let $V = \deg(v)$, $\bar{V} = \deg(\bar{v})$, and $W = \deg(w) = V + \bar{V}$ (no edge between v and $\bar{v}$ by Remark 5.4). Since $W \ge 4$ the "even allocation" rule is necessarily applied to w. If both v and $\bar{v}$ are such that the "even allocation" rules are applied, it is easy to check that monotonicity condition (5.7) reduces to

$$(V/2 - 1)|\mathfrak{s}|/V \le (W/2 - 1)|\mathfrak{s}|/W,$$

which holds since $V \le W$ and $x \mapsto \frac{x-2}{x}$ is increasing on $\mathbf{R}_+$.

Now suppose that the "even allocation" rule is applied to $\bar{v}$ while the "divergence priority" rule is applied to v (one then necessarily has $V = 3$ and $W = \bar{V} + 3$). Then

$$\frac{(\bar{V}/2 - 1)|\mathfrak{s}|}{\bar{V}} \vee \frac{|\mathfrak{s}|}{4} \le \frac{((\bar{V} + 3)/2 - 1)|\mathfrak{s}|}{\bar{V} + 3},$$

namely, the condition (5.7) holds. Here, $|\mathfrak{s}|/4$ is the largest possible value of $b_e^{(v)}$. Finally, if the "divergence priority" rules are applied to both v and $\bar{v}$, so that one necessarily has $V = \bar{V} = 3$, then one can verify (5.7) in an analogous way. □

REMARK 5.9. The above divergence priority rule is such that when $\mathcal{V}$ contains a subgraph where each dashed line represents a function $|x|_\varepsilon^{1-|\mathfrak{s}|}$ (think of it as $|K'_\varepsilon|$) and a factor $\varepsilon^{|\mathfrak{s}|/2}$ is associated with the middle vertex, the function $\varepsilon^{|\mathfrak{s}|/2}|x|_\varepsilon^{2(1-|\mathfrak{s}|)} \le |x|_\varepsilon^{2-3|\mathfrak{s}|/2}$ is integrable as long as $|\mathfrak{s}| < 4$, which turns out to be the subcriticality condition for the KPZ equation. For the dynamical Φ^4 equation, consider the subgraph with a factor $\varepsilon^{|\mathfrak{s}|}$. One should again associate all the powers of ε to the triple-edge, so that the function $\varepsilon^{|\mathfrak{s}|}|x|_\varepsilon^{3(2-|\mathfrak{s}|)} \le |x|_\varepsilon^{6-2|\mathfrak{s}|}$ is integrable as long as $|\mathfrak{s}| < 6$, which is again the subcriticality condition. (In fact, for the situation one also has to associate more powers of ε to the triple-edge than that to the double-edge when we are very close to the criticality in a certain sense.) See Remark 6.6 for discussion on the dynamical Φ^4 equation in three space dimensions.

As mentioned earlier, the ε-allocation procedure is aimed to improve the homogeneities associated to the edges, and thus cure the "fake" divergences arising from the contractions. Given a graph $(\mathcal{V}, \mathcal{E}')$ constructed from Wick contracting p copies of an partial graph H, together with an admissible ε-allocation rule, we define for every $e \in \mathcal{E}'$

$$a_e \stackrel{\text{def}}{=} m_e - b_e^{(v)} \tag{5.10}$$

if there exists ex-vertex v (which is necessarily unique by Remark 5.4) such that $v \in e$, and $a_e \stackrel{\text{def}}{=} m_e$ otherwise. In this way, given an ε-allocation rule and a Wick contraction of H, we obtain a labelled graph $(\mathcal{V}, \mathcal{E}', \{a_e\}_{e\in\mathcal{E}'})$.

Propositions 5.11 and 5.13 below are the main results in this section which state that certain conditions on H ensure that $(\mathcal{V}, \mathcal{E}', \{a_e\})$ satisfies the conditions required for applying the results of [26], Section 9.

Before being able to state Proposition 5.11, we need to introduce one more notation. Let H be a partial graph. Given an ε-allocation rule and a subgraph $\bar{H} \subset H$, we define numbers $b_e(\bar{H})$ for edges e of H as follows. Consider a graph $\mathcal{V}$ obtained from Wick contracting p copies $H^{(1)}, \dots, H^{(p)}$ of H for some $p \ge 2$, such that, considering $\bar{H}$ as a subgraph of $H^{(1)}$, *all* the elements in $\bar{H} \cap H_{\text{ex}}^{(1)}$ are identified in $\mathcal{V}$ as one single vertex v. Applying the ε-allocation rule to v, we obtain numbers $b_e^{(v)}$. We then define

$$b_e(\bar{H}) \stackrel{\text{def}}{=} \inf_{\mathcal{V}} b_e^{(v)}, \tag{5.11}$$

where the infimum is over all choices of $\mathcal{V}$ as above.

Given this definition, we furthermore define numbers $c_e(\bar{H})$ by

$$c_e(\bar{H}) = \begin{cases} b_e(\bar{H}), & \text{if } 0 \notin \bar{H} \text{ and } e \text{ touches an ex-vertex,} \\ |\mathfrak{s}|/2, & \text{if } 0 \in \bar{H} \text{ and } e \text{ touches an ex-vertex,} \\ 0, & \text{otherwise.} \end{cases}$$

This definition in engineered in such a way that, as a consequence of (5.6) and of (5.11), if $\mathcal{V}$ is obtained by Wick contracting p copies $H^{(1)}, \dots, H^{(p)}$ of H for some integer $p \geq 2$, and $\bar{H}^{(i)}$ is the pre-image of some $\bar{\mathcal{V}} \subset \mathcal{V}$ with $|\bar{\mathcal{V}}_{\mathrm{ex}}| \leq 1$ under the map $H^{(i)} \to \mathcal{V}$, then one has the bound:

$$\sum_i \sum_{e \in \mathcal{E}_0(\bar{H}^{(i)})} c_e(\bar{H}^{(i)}) \leq |\mathfrak{s}| \mathbf{1}_{0 \in \bar{\mathcal{V}}} \mathbf{1}_{\bar{\mathcal{V}}_{\mathrm{ex}} \neq \emptyset} + \sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} b_e^{(v)}, \tag{5.12}$$

where v in the unique element in $\mathcal{V}_{\mathrm{ex}}$ if $|\bar{\mathcal{V}}_{\mathrm{ex}}| = 1$ [if $|\bar{\mathcal{V}}_{\mathrm{ex}}| = 0$, (5.12) reads $0 \leq 0$.] Here, we wrote $\bar{\mathcal{V}}_{\mathrm{ex}}$ as a shorthand for $\bar{\mathcal{V}} \cap \mathcal{V}_{\mathrm{ex}}$. (We will similarly also use the notation $\bar{\mathcal{V}}_{\mathrm{in}} = \bar{\mathcal{V}} \cap \mathcal{V}_{\mathrm{in}}$.)

REMARK 5.10. The number $c_e(\bar{H})$ defined in Definition 4.17 is a special case (and thus consistent with) the general definition of $c_e(\bar{H})$ here with the ε-allocation rule specified above Lemma 5.8. Indeed, with $|\mathfrak{s}| = 3$, let $n = |\bar{H} \cap H_{\mathrm{ex}}|$. One then has $\deg(v) \geq n + 1$ by the definition of a Wick contraction. For $n > 2$, by (5.5) and the even allocation rule, one then has $b_e(\bar{H}) = \inf_{k \geq 1} 3(\frac{n+k}{2} - 1)/(n + k)$, which is equal to the right-hand side of (4.17). In the cases $n \in \{1, 2\}$, it can be also easily seen that Definition 4.17 follows from (5.11).

We then have the following criterion.

PROPOSITION 5.11. *Given a partial graph H, fix an admissible ε-allocation rule. Suppose that for every subset $\bar{H} \subset H$ with $|\bar{H}| \geq 2$ and $|\bar{H}_{\mathrm{in}}| \geq 1$ one has*

$$\sum_{e \in \mathcal{E}_0(\bar{H})} \big(m_e - c_e(\bar{H})\big) < |\mathfrak{s}| \big(|\bar{H}_{\mathrm{in}}| - \mathbf{1}_{\bar{H} \subset H_{\mathrm{in}}}\big). \tag{5.13}$$

Then, for every $p > 1$ and every graph $\mathcal{V}$ obtained from Wick contracting p copies of the graph H, the condition

$$\sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} a_e < |\mathfrak{s}| \big(|\bar{\mathcal{V}}| - 1\big) \tag{5.14}$$

holds for every subset $\bar{\mathcal{V}} \subset \mathcal{V}$ of cardinality at least 2, where a_e is defined in (5.10).

PROOF. We first remark that in the case $|\bar{H}| = 1$, (5.13) always reads $0 < 0$, so that it still holds, but without the inequality being strict. We first treat the case when $\bar{\mathcal{V}}$ has at most one ex-vertex. As a consequence of (5.12), and with the same notation as above for $\bar{H}^{(i)}$, we have

$$\begin{aligned} \sum_{e \in \mathcal{E}_0(\bar{\mathcal{V}})} a_e &\leq |\mathfrak{s}| \mathbf{1}_{0 \in \bar{\mathcal{V}}} \mathbf{1}_{\bar{\mathcal{V}} \cap \mathcal{V}_{\mathrm{ex}} \neq \emptyset} + \sum_{i \,:\, \bar{H}_{\mathrm{in}}^{(i)} \neq \emptyset} \sum_{e \in \mathcal{E}_0(\bar{H}^{(i)})} \big(m_e - c_e(\bar{H}^{(i)})\big) \\ &\leq |\mathfrak{s}| \Big(\mathbf{1}_{0 \in \bar{\mathcal{V}}} \mathbf{1}_{\bar{\mathcal{V}} \cap \mathcal{V}_{\mathrm{ex}} \neq \emptyset} + \sum_{i \,:\, \bar{H}_{\mathrm{in}}^{(i)} \neq \emptyset} \big(|\bar{H}_{\mathrm{in}}^{(i)}| - \mathbf{1}_{\bar{H}^{(i)} \subset H_{\mathrm{in}}^{(i)}}\big) \Big). \end{aligned} \tag{5.15}$$

On the other hand, we have the identity

$$|\bar{\mathcal{V}}| - 1 = \mathbf{1}_{0\in\bar{\mathcal{V}}} + |\bar{\mathcal{V}}_{\mathrm{ex}}| - 1 + \sum_i |\bar{H}_{\mathrm{in}}^{(i)}|.$$

We claim that (5.14) follows, but with an inequality that is not necessarily strict. Indeed, using the brutal bound $\mathbf{1}_{\bar{H}^{(i)}\subset H_{\mathrm{in}}^{(i)}} \geq 0$ in (5.15), it could only fail if $0 \notin \bar{\mathcal{V}}$ and $\bar{\mathcal{V}}_{\mathrm{ex}} = \emptyset$. In this case, however, one has necessarily $\bar{H}^{(i)} \subset H_{\mathrm{in}}^{(i)}$ for every i, so that the inequality is restored by the fact that at least one of the $\bar{H}^{(i)}$ must be nonempty. To see that the inequality is indeed strict, we note that the second inequality in (5.15) is actually strict, unless $|\bar{H}^{(i)}| = 1$ for every nonempty $\bar{H}_{\mathrm{in}}^{(i)}$. In this case, however, since $|\bar{\mathcal{V}}| > 1$ by assumption, it must be the case that either (1) $\bar{H}^{(i)} \subset H_{\mathrm{in}}^{(i)}$ for every i and, therefore, that at least two of the $\bar{H}^{(i)}$ are nonempty, which again restores the strict inequality; or (2) $\bar{\mathcal{V}}$ has one of the following simple forms: $\{0, v\}$ for some $v \neq 0$, $\{0, u, v\}$ for some in-vertex u and some ex-vertex v, $\{u, v\}$ for some in-vertex u and some ex-vertex v, and in all these cases it is straightforward to check the desired bound.

We now turn to the case where $\bar{\mathcal{V}}$ contains at least two ex-vertices; call them v_1 and v_2. The idea is then to compare this to the graph $\hat{\mathcal{V}}$ obtained by contracting v_1 and v_2, with $\hat{\bar{\mathcal{V}}} \subset \hat{\mathcal{V}}$ the corresponding subgraph. We also write $\hat{a}_e$ for the edge-weights of $\hat{\bar{\mathcal{V}}}$. It then follows from Lemma 5.7 that

$$\sum_{e\in\mathcal{E}_0(\bar{\mathcal{V}})} a_e \leq \sum_{e\in\mathcal{E}_0(\hat{\bar{\mathcal{V}}})} \hat{a}_e + |\mathfrak{s}|.$$

Since furthermore $|\hat{\bar{\mathcal{V}}}| = |\bar{\mathcal{V}}| - 1$, the required bound for $\bar{\mathcal{V}}$ follows from the corresponding bound for $\hat{\bar{\mathcal{V}}}$. Successively contracting external vertices, we can therefore reduce ourselves to the case where $\bar{\mathcal{V}}$ contains only one ex-vertex. If the resulting graph still has at least two vertices, the claim follows. If not, then we have $\bar{\mathcal{V}} \subset \mathcal{V}_{\mathrm{ex}}$, so that $\mathcal{E}_0(\bar{\mathcal{V}})$ is empty and the bound (5.14) is trivially satisfied. □

REMARK 5.12. If H is such that $\deg(0) = 1$ (which is often the case in practice), then (5.13) only needs to be verified for graphs $\bar{H}$ with $0 \notin \bar{H}$. Indeed, let $\bar{H}$ be such that $0 \in \bar{H}$, set $\bar{H}_0 = \bar{H} \setminus \{0\}$, and assume that $\bar{H}_0$ satisfies (5.13). If $\bar{H}_{\mathrm{ex}} = \emptyset$, then it is clear that the condition (5.13) for $\bar{H}_0$ implies that for $\bar{H}$, so we only consider the case $\bar{H}_{\mathrm{ex}} \neq \emptyset$.

Since $b_e(\bar{H})$ is defined by (5.11) as an infimum over all choices of $\mathcal{V}$, it is in particular bounded by the case where $\mathcal{V}$ is such that all vertices of $\bar{H}_{\mathrm{ex}}$ are identified to one single vertex v which contains only one additional edge not belonging to $\bar{H}$. It thus follows from (5.5) that

$$\sum_e c_e(\bar{H}_0) = \sum_e b_e(\bar{H}_0) \leq \Big(\frac{|\bar{H}_{\mathrm{ex}}| + 1}{2} - 1\Big)|\mathfrak{s}| \leq \frac{|\bar{H}_{\mathrm{ex}}||\mathfrak{s}|}{2} = \sum_e c_e(\bar{H}),$$

so that condition (5.13) indeed holds for $\bar{H}$ whenever it holds for $\bar{H}_0$.

PROPOSITION 5.13. *Given a partial graph H, fix an admissible ε-allocation rule such that the numbers $b_e^{(v)}$ always satisfy $b_e^{(v)} \le |\mathfrak{s}|/2$. Assume that for every nonempty subgraph $\bar{H} \subset H \setminus H_\star$ one has*

$$\sum_{e \in \mathcal{E}(\bar{H})} m_e > \Big(|\bar{H}_{\mathrm{in}}| + \frac{1}{2} |\bar{H}_{\mathrm{ex}}| \Big) |\mathfrak{s}|. \tag{5.16}$$

Then, for every graph $\mathcal{V}$ obtained by Wick contracting several copies of H, and for every nonempty subgraph $\bar{\mathcal{V}} \subset \mathcal{V} \setminus \mathcal{V}_\star$, one has

$$\sum_{e \in \mathcal{E}(\bar{\mathcal{V}})} a_e > |\mathfrak{s}||\bar{\mathcal{V}}|, \tag{5.17}$$

where a_e is defined in (5.10).

REMARK 5.14. It is easy to see that the KPZ allocation rule satisfies $b_e^{(v)} \le |\mathfrak{s}|/2$.

PROOF. Given $\bar{\mathcal{V}} \subset \mathcal{V} \setminus \mathcal{V}_\star$, let $m = |\bar{\mathcal{V}}_{\mathrm{in}}|$ and $q = |\bar{\mathcal{V}}_{\mathrm{ex}}|$, so that $|\bar{\mathcal{V}}| = m + q$. Denote by $H^{(i)}$ the ith copy of H in $\mathcal{V}$ as before, and denote by m_i the number of in-vertices of $\bar{\mathcal{V}}$ belonging to $H^{(i)}$, so that $m = \sum_i m_i$. We also denote by n_i the number of external vertices in $H^{(i)}$ whose images under the quotient map are ex-vertices of $\bar{\mathcal{V}}$. With this notation, (5.16) applied to $\bar{\mathcal{V}} \cap H^{(i)}$ reads

$$\sum_{e \in \mathcal{E}(\bar{\mathcal{V}}) \cap \mathcal{E}(H^{(i)})} m_e > (m_i + n_i/2)|\mathfrak{s}|,$$

and the intersection here is well defined because by construction $\mathcal{E}(\mathcal{V}) = \bigcup_i \mathcal{E}(H^{(i)})$. In fact, we actually have an even stronger bound. Indeed, there may exist an external edge $e = (u, v)$ of $H^{(i)}$ (say u is an in-vertex and v is an ex-vertex of $H^{(i)}$), such that $u \in \bar{\mathcal{V}}$ but $v \notin \bar{\mathcal{V}}$. Suppose that there are $\tilde{n}_i$ such edges in $H^{(i)}$. We apply (5.16) to the *union* of $\bar{\mathcal{V}} \cap H^{(i)}$ *and* these $\tilde{n}_i$ ex-vertices:

$$\sum_{e \in \mathcal{E}(\bar{\mathcal{V}}) \cap \mathcal{E}(H^{(i)})} m_e > (m_i + n_i/2 + \tilde{n}_i/2)|\mathfrak{s}|.$$

Summing over i, one has

$$\sum_{e \in \mathcal{E}(\bar{\mathcal{V}})} m_e > (m + n/2 + \tilde{n}/2)|\mathfrak{s}|, \tag{5.18}$$

where $n \stackrel{\mathrm{def}}{=} \sum_i n_i$ and $\tilde{n} \stackrel{\mathrm{def}}{=} \sum_i \tilde{n}_i$. Let now $d_1, \dots, d_q$ be the degrees of the q ex-vertices (understood as $\mathcal{V}$-degrees, namely the number of edges in $\mathcal{V}$, not $\bar{\mathcal{V}}$, attached to the vertex). By the procedure used to construct $\mathcal{V}$ from H and the fact

that the degree of every external vertex of H is 1, one has $\sum_{i=1}^{q} d_i = n$, so that (5.18) is equivalent to

$$\sum_{e\in\mathcal{E}(\bar{\mathcal{V}})} m_e - \sum_{j=1}^{q}(d_j/2-1)|\mathfrak{s}| - \tilde{n}|\mathfrak{s}|/2 > (m+q)|\mathfrak{s}|.$$

By (5.5), we know that for each of the q ex-vertices v we have $(d_j/2-1)|\mathfrak{s}| = \sum_{e\ni v} b_e^{(v)}$, and every $e \ni v$ here belongs to $\mathcal{E}(\bar{\mathcal{V}})$. All the other edges in $\mathcal{E}(\bar{\mathcal{V}})$ that acquire nonzero $b_e^{(v)}$ are of the form (u, v) described above, and there are $\tilde{n}$ of them, each satisfying $b_e^{(v)} \le |\mathfrak{s}|/2$. Therefore, the left-hand side of the above inequality is smaller or equal to $\sum_{e\in\mathcal{E}(\bar{\mathcal{V}})} a_e$. We then obtain the desired bound (5.17). □

COROLLARY 5.15. *Given a partial graph H with labels $\{m_e\}_{e\in\mathcal{E}(H)}$ and $p \ge 2$, let $(\mathcal{V}, \mathcal{E}')$ be a graph obtained by Wick contracting p copies of H. If the conditions of Proposition 5.11 and Proposition 5.13 are satisfied for H, then the quantity defined in (4.14) is bounded as*

$$\Lambda^{p}_{\varphi_0^\lambda}(\mathcal{V}) \lesssim \lambda^{\bar{\alpha}p},$$

where

$$\bar{\alpha} = |\mathfrak{s}|\big(|H_{\mathrm{ex}}|/2 + |H_{\mathrm{in}} \setminus H_\star|\big) - \sum_{e\in\mathcal{E}(H)} m_e.$$

PROOF. As described in Definition 4.13, the graph $(\mathcal{V}, \mathcal{E}')$ naturally inherits labels $\{m_e\}_{e\in\mathcal{E}'}$ from H, and by definition (5.10) of a_e, one has labelled graph $(\mathcal{V}, \mathcal{E}', \{a_e\})$. Let $(\mathcal{V}, \mathcal{E})$ be the graph obtained from $(\mathcal{V}, \mathcal{E}')$ by simply identifying each multi-edge as one single edge whose label is the sum of the labels a_e of the multi-edge as described above. With slight abuse of notation we still denote by a_e the labels of $(\mathcal{V}, \mathcal{E})$. Then we are precisely in the setting of Theorem 5.3, where each function $J_e(x)$ is simply $|x|^{-a_e}$ which has finite norm $\|J_e\|_{a_e}$, except $J_e(x) = \varphi_0^\lambda(x)$ for the distinguished edges e. Since the way we define a_e in (5.10) simply encodes the bound $\varepsilon^{b_e^{(v)}}|x_v - x_{\bar{v}}|_\varepsilon^{-m_e} \lesssim |x_v - x_{\bar{v}}|^{-a_e}$ for each $e = \{v, \bar{v}\}$, and we precisely use up all the factors $\varepsilon^{(\deg(v)/2-1)|\mathfrak{s}|}$ in (4.14) due to the condition (5.5), one has

$$\Lambda^{p}_{\varphi_0^\lambda}(\mathcal{V}) \lesssim \big|\mathcal{I}_\lambda(J)\big|.$$

So it is enough to show that Assumption 5.2 is satisfied for $(\mathcal{V}, \mathcal{E})$. In fact, the conclusions of Proposition 5.11 and Proposition 5.13 are obviously the same as the conditions (5.2) and (5.3) of Assumption 5.2, respectively (note that the sum of labels a_e on the left-hand side of each of these conditions is the same after we identify the multi-edges). So Assumption 5.2 is satisfied.

Finally, the power $\bar{\alpha}$ is obtained from Definition 4.13 of the procedure of constructing the graph $\mathcal{V}$, the condition (5.5) of the admissible ε-allocation rule, the relation (5.10) between a_e and m_e, and the power α defined in Theorem 5.3. □

6. Identification of the limit. It was shown in [26] (but see also [22] and [18] for very similar results) that if we replace ζ_ε by ξ_ε with $\xi_\varepsilon = \varrho_\varepsilon * \xi$ where ξ is the Gaussian space–time white noise, the renormalised models built from ξ_ε converge to a limit $\hat{Z} = (\hat{\Pi}, \hat{\Gamma})$, which we call the KPZ model.[3] The goal of this section is to show that our renormalised models $(\hat{\Pi}^{(\varepsilon)}, \hat{\Gamma}^{(\varepsilon)})$ built from ζ_ε defined above converge to the same limit. We prove this by applying a "diagonal argument" along the lines of the one used in [31] in a similar context. First of all, we have a central limit theorem for the random field ζ_ε.

PROPOSITION 6.1. *Under Assumption 2.1, for every $\alpha < -\frac{3}{2}$, ζ_ε converges in law to space–time white noise ξ in $\mathcal{C}^\alpha(\mathbf{R} \times S^1)$.*

The proof of this proposition relies on the following result.

LEMMA 6.2. *Let ζ be a random variable in $\mathcal{C}^\alpha(\mathbf{R} \times \mathbf{T}^d)$ for some $\alpha < 0$ which is stationary and such that, for any finite collection of test functions $\eta_1, \dots, \eta_N$ with $\operatorname{supp} \eta_i$ not intersecting the hyperplane $\{x = vt\}$ (viewed as a subset of $\mathbf{R} \times \mathbf{T}^d$ in the natural way) for a constant vector $v \in \mathbf{R}^d$, the joint law of $\{\zeta(\eta_i)\}$ is Gaussian with covariance $\langle \eta_i, \eta_j \rangle$. Then ζ is space–time white noise.*

PROOF. Choose an orthonormal basis $\{e_n\}$ of $L^2(\mathbf{R} \times \mathbf{T}^d)$ such that each e_n is sufficiently smooth ($\mathcal{C}^{|\alpha|}$ will do) and supported away from $\{x = vt\}$. Then, as a consequence of the assumption, $\zeta(e_n)$ is an i.i.d. sequence of normal random variables. As a consequence, we can set $\xi = \sum_n \zeta(e_n) e_n$, which converges almost surely in the sense of distributions to a space–time white noise ξ. It remains to show that $\xi = \zeta$ almost surely, as random variables in $\mathcal{C}^\alpha$, which follows if we can show that $\xi(\varphi) = \zeta(\varphi)$ almost surely, for any smooth test function φ with support in some ball of radius $1/5$ (say).

If φ is supported away from $\{x = vt\}$, this is immediate. Otherwise, write $\varphi = \lim_{n\to\infty} \varphi_n$, where each φ_n is compactly supported away from $\{x = vt\}$ and the convergence is sufficiently fast so that $\xi(\varphi_n) \to \xi(\varphi)$ almost surely. (It is easy to see that such an approximation always exists.) Let now $(T\varphi)(t, x) = \varphi(t, x - 1/2)$ and note that one also has $\zeta(T\varphi_n) \to \zeta(T\varphi)$ almost surely since the supports of $T\varphi_n$ avoid $\{x = vt\}$ and ζ coincides with ξ for such test functions. By stationarity,

[3] This is a slight misnomer, since the constant parts of $\hat{\Pi}_z \mathbb{V}$ and $\hat{\Pi}_z \mathbb{V}$ depend on the choice of cutoff function K. However, the model does not depend on ϱ and, when comparing it to the model built from ζ_ε, we will always implicitly assume that we make the same choice for K.

the collection of random variables $\{\zeta(T\varphi), \zeta(T\varphi_n) : n \geq 0\}$ is equal in law to the collection $\{\zeta(\varphi), \zeta(\varphi_n) : n \geq 0\}$, so that $\zeta(\varphi_n) \to \zeta(\varphi)$ almost surely. Since we already know that $\zeta(\varphi_n) = \xi(\varphi_n)$ almost surely, we conclude that $\zeta(\varphi) = \xi(\varphi)$ as required. □

REMARK 6.3. The assumption that ζ is stationary is crucial here, otherwise the conclusion does not hold in general. (Just add to ζ a Dirac measure located at the origin.)

PROOF OF PROPOSITION 6.1. Setting $\zeta_\varepsilon^{(0)}(t,x) = \varepsilon^{-3/2}\zeta(t/\varepsilon^2, (x - v_h t)/\varepsilon)$, it is shown in Lemma 6.4 below that, when testing $\zeta_\varepsilon^{(0)}$ against a finite number of test functions $\eta_1, \dots, \eta_N$, the law of the resulting $\mathbf{R}^n$-valued random variable converges to a Gaussian distribution with covariance $C_{ij} = \langle \eta_i, \eta_j \rangle$.

Following the same argument as in Lemma 6.4, one shows that the sequence ζ_ε is tight, so that it has some accumulation point ξ. Since each of the ζ_ε is stationary, ξ is also stationary. By Lemma 6.2, it therefore suffices to verify that it has the correct finite-dimensional distributions when tested against test functions avoiding the hyperplane $x = v_h t$. By stationarity, we can replace this by the hyperplane $x = v_h t \pm \frac{1}{2}$ (which is one single plane if we view the spatial variable as taking values on the circle).

For any $\delta > 0$, define the space–time domain

$$D_\delta = \big\{(t,x) : |t| \leq T, |x - v_h t| \leq (1-\delta)/2\big\},$$

and consider a finite collection of test functions as above, but with supports contained in D_δ. Choosing the coupling mentioned in the assumption, we then have the bound:

$$\begin{aligned}\mathbf{E}\big|(\zeta_\varepsilon - \zeta_\varepsilon^{(0)})(\eta_i)\big|^2 &\lesssim \sup_{(t,x)\in \operatorname{supp}\eta_i} \mathbf{E}\big|\zeta_\varepsilon(t,x) - \zeta_\varepsilon^{(0)}(t,x)\big|^2 \\ &\lesssim \varepsilon^{-3} \sup_{|t|\leq T\varepsilon^{-2}} \sup_{|x|\leq (1-\delta)/(2\varepsilon)} \mathbf{E}\big|\zeta(t,x) - \zeta^{(\varepsilon)}(t,x)\big|^2,\end{aligned}$$

which converges to 0 as $\varepsilon \to 0$ by assumption. □

The only missing piece in the proof of Proposition 6.1 is the following simple lemma. Note that Assumptions 1.1 and 1.2 can be also easily formulated on $\mathbf{R} \times \mathbf{R}^d$.

LEMMA 6.4. *Let ζ be a random field on $\mathbf{R} \times \mathbf{R}^d$ equipped with scaling $\mathfrak{s} = (2, 1, \dots, 1)$ satisfying Assumptions 1.1 and 1.2. Let*

$$\zeta_\varepsilon^{(0)}(z) \stackrel{\text{def}}{=} \varepsilon^{-(d+2)/2}\zeta\Big(\frac{t}{\varepsilon^2}, \frac{x - v^{(\varepsilon)}t}{\varepsilon}\Big),$$

where $v^{(\varepsilon)}$ converges to a constant vector $v \in \mathbf{R}^d$ and $z = (t, x) \in \mathbf{R} \times \mathbf{R}^d$. The sequence of random fields $\zeta_\varepsilon^{(0)}$ converges in distribution to Gaussian white noise ξ with scaling $\mathfrak{s}$ as $\varepsilon \to 0$ in the space $\mathcal{C}^\gamma$ for every $\gamma < -\frac{d+2}{2}$.

PROOF. We first show that $\mathbf{E}\|\zeta_\varepsilon^{(0)}\|^p_{\mathcal{C}^{-|\mathfrak{s}|/2}}$ (where $|\mathfrak{s}| = d + 2$) is bounded uniformly in $\varepsilon > 0$ for all $p > 1$. For this it suffices to show that $\lambda^{|\mathfrak{s}|p/2}\mathbf{E}|\zeta_\varepsilon^{(0)}(\varphi_0^\lambda)|^p \lesssim 1$ uniformly in $\varepsilon, \lambda > 0$ for all $p > 1$, which follows in exactly the same way as the proof of the case $\tau = \Xi$ in Theorem 4.1.

To identify the limit, it follows from the multidimensional version of Carleman's theorem [1, 13] that it is enough to show that for any given finite number of compactly supported smooth functions $\varphi_1, \dots, \varphi_n$, one has convergence of joint moments $\mathbf{E}(\zeta_\varepsilon^{(0)}(\varphi_1) \cdots \zeta_\varepsilon^{(0)}(\varphi_n)) \to \mathbf{E}(\xi(\varphi_1) \cdots \xi(\varphi_n))$ as $\varepsilon \to 0$. This implies convergence $\zeta_\varepsilon^{(0)} \to \xi$ in distribution, and the first part of the proof gives convergence in the desired topology. The calculation for the joint moment of $\zeta_\varepsilon^{(0)}(\varphi_i)$ is the same as above, except that φ_0^λ is replaced by φ_i. For $B \in \pi$, $|B| > 2$, the integral is bounded by a positive power of ε and thus converges to zero. The nonvanishing terms are the partitions π such that $|B| = 2$ for every $B \in \pi$. Since κ_2 is normalised in Assumption 1.1 such that the second cumulant of $\zeta_\varepsilon^{(0)}$ converges to the Dirac distribution, the joint moment converges to that of the Gaussian white noise by standard Wick theorem. □

THEOREM 6.5. *Let $\hat{Z}_\varepsilon = (\hat{\Pi}^{(\varepsilon)}, \hat{\Gamma}^{(\varepsilon)})$ be the renormalised model built from ζ_ε defined in the previous sections* [*with the choice of renormalisation constants given by* (3.10)]. *Let $\hat{Z} = (\hat{\Pi}, \hat{\Gamma})$ be the KPZ random model. Then, as $\varepsilon \to 0$, one has $\hat{Z}_\varepsilon \to \hat{Z}$ in distribution in the space $\mathscr{M}$ of admissible models for $\mathscr{T}$.*

PROOF. For any fixed bounded domain $D \subset \mathbf{R}^2$, recall that we have an associated "norm" $|||\cdot|||_D$ and pseudometric $|||\cdot, \cdot|||_D$ on the space $\mathscr{M}$ of admissible models, defined as in [23], equations (2.16)–(2.17). Since the topology on $\mathscr{M}$ is generated by these pseudometrics, it is sufficient to show that

$$\lim_{\varepsilon \to 0} \mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}|||_D = 0, \tag{6.1}$$

for every bounded domain D. Since D does not matter much in the sequel, we henceforth omit it from our notation.

In order to prove (6.1), we proceed by going through the "intermediate" model $\hat{Z}_{\varepsilon,\bar{\varepsilon}}$ previously defined in Section 4. Recall that this is built in the same way as $\hat{Z}_\varepsilon$, except that the field ζ_ε used to build the model is replaced by $\zeta_{\varepsilon,\bar{\varepsilon}} \stackrel{\text{def}}{=} \zeta_\varepsilon * \varrho_{\bar{\varepsilon}}$, where ϱ is a compactly supported smooth function on $\mathbf{R}^2$ integrating to 1 that is even in the space variable, and with $\varrho_{\bar{\varepsilon}} = \bar{\varepsilon}^{-3}\varrho(\bar{\varepsilon}^{-2}t, \bar{\varepsilon}^{-1}x)$. We furthermore define a model $\hat{Z}_{0,\bar{\varepsilon}}$ obtained again in the same way, but this time replacing ζ_ε by $\zeta_{0,\bar{\varepsilon}} \stackrel{\text{def}}{=} \xi * \varrho_{\bar{\varepsilon}}$, where ξ is a realisation of space–time white noise.

With these notation at hand, we then have

$$\mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}||| \le \mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}_{\varepsilon,\bar\varepsilon}||| + \mathbf{E}|||\hat{Z}_{\varepsilon,\bar\varepsilon}, \hat{Z}_{0,\bar\varepsilon}||| + \mathbf{E}|||\hat{Z}_{0,\bar\varepsilon}, \hat{Z}|||,$$

which is valid for any fixed $\bar\varepsilon > 0$ and for any coupling between ξ and ζ_ε, so that

$$\lim_{\varepsilon\to 0} \mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}||| \le \lim_{\bar\varepsilon\to 0}\lim_{\varepsilon\to 0}\big(\mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}_{\varepsilon,\bar\varepsilon}||| + \mathbf{E}|||\hat{Z}_{\varepsilon,\bar\varepsilon}, \hat{Z}_{0,\bar\varepsilon}||| + \mathbf{E}|||\hat{Z}_{0,\bar\varepsilon}, \hat{Z}|||\big). \tag{6.2}$$

Note the order of limits: we *first* take $\varepsilon \to 0$ and *then* only take $\bar\varepsilon \to 0$.

The convergence of $\hat{Z}_{0,\bar\varepsilon}$ to the KPZ model in the sense that

$$\lim_{\bar\varepsilon\to 0} \mathbf{E}|||\hat{Z}_{0,\bar\varepsilon}, \hat{Z}|||^p = 0,$$

for every p was already shown in [26], Theorem 6.1 (see also [18, 22] in a slightly different setting). Regarding the first term, it suffices to note that, by [23], Theorem 10.7, and the second bound in Theorem 4.1, we obtain the bound

$$\mathbf{E}|||\hat{Z}_\varepsilon, \hat{Z}_{\varepsilon,\bar\varepsilon}||| \lesssim \bar\varepsilon^\kappa,$$

uniformly over ε sufficiently small, so that this term also vanishes.

It therefore remains to bound the second term. This term involves both ζ_ε and ξ, so we need to specify a coupling between them. For this, we recall that, given any sequence X_n of random variables on a complete separable metric space $\mathcal{X}$, weak convergence to a limit X and uniform boundedness of some moment of order strictly greater than p of $d(0, X_n)$ implies convergence in the p-Wasserstein metric, which in turn implies that, for every n, there exists a coupling between X_n and X such that

$$\lim_{n\to\infty} \mathbf{E} d(X_n, X)^p = 0;$$

see, for example, [36]. Taking for $\mathcal{X}$ the space $\mathcal{C}^{-2}$ (say) on any bounded domain $\mathfrak{K}$ with the metric given by its norm, it follows from Proposition 6.1 and Theorem 4.1 (with $\tau = \Xi$) that we can find couplings between ζ_ε and ξ such that

$$\lim_{\varepsilon\to 0} \mathbf{E}\|\xi - \zeta_\varepsilon\|^p_{-2;\mathfrak{K}} = 0.$$

At this point, we use the fact that, for any fixed $\bar\varepsilon > 0$, convolution with $\varrho_{\bar\varepsilon}$ maps $\mathcal{C}^{-2}$ into $\mathcal{C}^1$ (actually $\mathcal{C}^\infty$), so that

$$\lim_{\varepsilon\to 0} \mathbf{E}\|\zeta_{0,\bar\varepsilon} - \zeta_{\varepsilon,\bar\varepsilon}\|^p_{1;\mathfrak{K}} = 0. \tag{6.3}$$

Consider now the space $\mathcal{Y}$ of all stationary random processes η on some given probability space with the following additional properties:

- The process η is almost surely periodic in space with period 1.
- One has $\mathbf{E}\|\eta\|^p_{1;\mathfrak{K}} < \infty$ for every $p \ge 1$.

We endow the space $\mathcal{Y}$ with the seminorm

$$\|\eta\|_{\mathcal{Y}}^p = \mathbf{E}\|\eta\|_{1;\mathfrak{K}}^p,$$

and we denote by $\hat{\Psi}$ the map $\zeta_\varepsilon \mapsto \hat{Z}_\varepsilon$, viewed as a map from $\mathcal{Y}$ into the space of $\mathscr{M}$-valued random variables. Here, the requirement that the argument be a stationary stochastic process is needed for the renormalisation constants to be well defined by the formulae in Section 3.2.

It is then immediate from the definitions that, for some sufficiently large p, $\hat{\Psi}$ satisfies

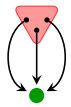
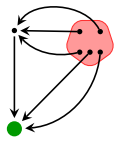
$$\mathbf{E}\|\hat{\Psi}(\zeta), \hat{\Psi}(\eta)\| \lesssim \|\zeta - \eta\|_{\mathcal{Y}}(1 + \|\zeta\|_{\mathcal{Y}} + \|\eta\|_{\mathcal{Y}})^p.$$

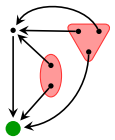
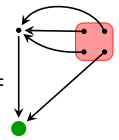
Combining this with (6.3) and the fact that $\hat{Z}_{\varepsilon,\bar{\varepsilon}} = \hat{\Psi}(\zeta_{\varepsilon,\bar{\varepsilon}})$, it immediately follows that

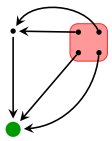
$$\lim_{\varepsilon\to 0} \mathbf{E}\|\hat{Z}_{\varepsilon,\bar{\varepsilon}}, \hat{Z}_{0,\bar{\varepsilon}}\| = 0,$$

for every fixed (sufficiently small) $\bar{\varepsilon} > 0$, so that the second term in (6.2) also vanishes, thus concluding the proof. □

REMARK 6.6. Our proof also extends to the dynamical Φ^4_3 model driven by a general noise ζ under the same assumptions as here (except that the spatial variable takes values in $\mathbf{T}^3$ instead of S^1). Besides the constant $C_1^{(\varepsilon)} \approx \varepsilon^{-1}$ and the logarithmically divergent constant $C_2^{(\varepsilon)}$ as in the Gaussian case, we then have the following new renormalisation constants appearing:

$\bar{C}_1^{(\varepsilon)} =$  $\approx \varepsilon^{-3/2}$, $\bar{C}_2^{(\varepsilon)} =$  $\approx \varepsilon^{-1/2}$,

$\bar{C}_3^{(\varepsilon)} =$  $\approx \varepsilon^{-1/2}$, $\bar{C}_4^{(\varepsilon)} =$  $\xrightarrow{\varepsilon\to 0} c_4$,

$\bar{C}_5^{(\varepsilon)} =$  $\xrightarrow{\varepsilon\to 0} c_5$,

where c_4, c_5 are some finite constants. The renormalised equation is then given by

$$\begin{aligned}\partial_t \Phi_\varepsilon = \Delta\Phi_\varepsilon - \lambda\Phi_\varepsilon^3 + (3\lambda C_1^{(\varepsilon)} - 9\lambda^2 C_2^{(\varepsilon)} - 6\lambda^2 c_4 - 9\lambda^2 c_5)\Phi_\varepsilon \\ + (\lambda\bar{C}_1^{(\varepsilon)} - 3\lambda^2\bar{C}_2^{(\varepsilon)} - 18\lambda^2\bar{C}_3^{(\varepsilon)}) + \zeta_\varepsilon.\end{aligned}$$

Using the method developed in this article, one can show that Φ_ε converges to a limit which coincides with the solution to the dynamical Φ^4_3 equation driven by Gaussian space–time white noise as constructed in [11, 23, 29].

Acknowledgements. We are indebted to Hendrik Weber and Jeremy Quastel for several interesting discussions on this topic. We also thank Henri Elad-Altman and the two anonymous referees for their very careful proofreading.

REFERENCES

[1] AKHIEZER, N. I. (1965). *The Classical Moment Problem and Some Related Questions in Analysis*. Hafner Publishing, New York. MR0184042

[2] ALBERTS, T., KHANIN, K. and QUASTEL, J. (2010). Intermediate disorder regime for directed polymers in dimension $1+1$. *Phys. Rev. Lett.* **105** 090603. Available at arXiv:1003.1885.

[3] 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. MR2796514

[4] AVRAM, F. and TAQQU, M. S. (1987). Noncentral limit theorems and Appell polynomials. *Ann. Probab.* **15** 767–775. MR0885142

[5] AVRAM, F. and TAQQU, M. S. (2006). On a Szegö type limit theorem and the asymptotic theory of random sums, integrals and quadratic forms. In *Dependence in Probability and Statistics*. *Lect. Notes Stat.* **187** 259–286. Springer, New York. MR2283259

[6] BAIK, J., DEIFT, P. and JOHANSSON, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178. MR1682248

[7] BALÁZS, M., QUASTEL, J. and SEPPÄLÄINEN, T. (2011). Fluctuation exponent of the KPZ/stochastic Burgers equation. *J. Amer. Math. Soc.* **24** 683–708. MR2784327

[8] BERTINI, L. and GIACOMIN, G. (1997). Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** 571–607.

[9] BOGOLIUBOW, N. N. and PARASIUK, O. S. (1957). Über die Multiplikation der Kausalfunktionen in der Quantentheorie der Felder. *Acta Math.* **97** 227–266. MR0095027

[10] BORODIN, A. and CORWIN, I. (2014). Macdonald processes. *Probab. Theory Related Fields* **158** 225–400.

[11] CATELLIER, R. and CHOUK, K. (2013). Paracontrolled distributions and the 3-dimensional stochastic quantization equation. Preprint. Available at arXiv:1310.6869.

[12] CORWIN, I. and TSAI, L.-C. (2015). KPZ equation limit of higher-spin exclusion processes. Preprint. Available at arXiv:1505.04158.

[13] CRAMÉR, H. and WOLD, H. (1936). Some Theorems on Distribution Functions. *J. Lond. Math. Soc. S* **1-11** 290.

[14] DA PRATO, G. and ZABCZYK, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge Univ. Press, Cambridge. MR3236753

[15] 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. MR3439226

[16] EDWARDS, S. F. and WILKINSON, D. R. (1982). The surface statistics of a granular aggregate. *Proc. Roy. Soc. London Ser. A* **381** 17–31.

[17] FRIZ, P. and VICTOIR, N. (2006). A note on the notion of geometric rough paths. *Probab. Theory Related Fields* **136** 395–416.

[18] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths*. *Universitext*. Springer, Cham. MR3289027

[19] GIRAITIS, L. and SURGAILIS, D. (1986). Multivariate Appell polynomials and the central limit theorem. In *Dependence in Probability and Statistics*. *Progr. Probab. Statist.* **11** 21–71. Birkhäuser Boston, Boston, MA.
[20] GLIMM, J. and JAFFE, A. (1987). *Quantum Physics*: *A Functional Integral Point of View*, 2nd ed. Springer, New York.
[21] GONÇALVES, P. and JARA, M. (2014). Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.* **212** 597–644. MR3176353
[22] HAIRER, M. (2013). Solving the KPZ equation. *Ann. of Math.* (2) **178** 559–664. MR3071506
[23] HAIRER, M. (2014). A theory of regularity structures. *Invent. Math.* **198** 269–504. MR3274562
[24] HAIRER, M. (2015). Introduction to regularity structures. *Braz. J. Probab. Stat.* **29** 175–210. MR3336866
[25] HAIRER, M. and PARDOUX, É. (2015). A Wong–Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan* **67** 1551–1604. Available at arXiv:1409.3138.
[26] HAIRER, M. and QUASTEL, J. (2015). A class of growth models rescaling to KPZ. Preprint. Available at arXiv:1512.07845.
[27] HEPP, K. (1969). On the equivalence of additive and analytic renormalization. *Comm. Math. Phys.* **14** 67–69. MR0250603
[28] KARDAR, M., PARISI, G. and ZHANG, Y.-C. (1986). Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** 889–892.
[29] KUPIAINEN, A. (2016). Renormalization group and stochastic PDEs. *Annales Henri Poincaré* **17** 497–535. MR3459120
[30] LUKKARINEN, J. and MARCOZZI, M. (2015). Wick polynomials and time-evolution of cumulants. Preprint. Available at arXiv:1503.05851.
[31] MOURRAT, J.-C. and WEBER, H. (2014). Convergence of the two-dimensional dynamic Ising–Kac model to φ_2^4. Preprint. Available at arXiv:1410.1179.
[32] O'CONNELL, N. and YOR, M. (2002). A representation for non-colliding random walks. *Electron. Commun. Probab.* **7** 1–12. (electronic).
[33] PECCATI, G. and TAQQU, M. S. (2011). *Wiener Chaos*: *Moments*, *Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Springe, Milan. MR2791919
[34] SASAMOTO, T. and SPOHN, H. (2010). Exact height distributions for the KPZ equation with narrow wedge initial condition. *Nuclear Phys. B* **834** 523–542.
[35] SURGAILIS, D. (1983). On Poisson multiple stochastic integrals and associated equilibrium Markov processes. In *Theory and Application of Random Fields*. *Lecture Notes in Control and Information Sciences* **49** 233–248. Springer, Berlin.
[36] VILLANI, C. (2009). *Optimal Transport Old and New*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Springer, Berlin. MR2459454
[37] ZIMMERMANN, W. (1969). Convergence of Bogoliubov's method of renormalization in momentum space. *Comm. Math. Phys.* **15** 208–234. MR0255162

MATHEMATICS INSTITUTE
UNIVERSITY OF WARWICK
COVENTRY, CV47AL
UNITED KINGDOM
E-MAIL: M.Hairer@Warwick.ac.uk

DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
2990 BROADWAY
NEW YORK, NEW YORK 10027
USA
E-MAIL: pkushenhao@gmail.com