Stoch PDE: Anal Comp (2018) 6:211–254
https://doi.org/10.1007/s40072-017-0107-4

Weak universality of dynamical Φ^4_3: non-Gaussian noise

Hao Shen[1] · **Weijun Xu**[2]

Received: 2 February 2016 / Published online: 16 October 2017

Abstract We consider a class of continuous phase coexistence models in three spatial dimensions. The fluctuations are driven by symmetric stationary random fields with sufficient integrability and mixing conditions, but not necessarily Gaussian. We show that, in the weakly nonlinear regime, if the external potential is a symmetric polynomial and a certain average of it exhibits pitchfork bifurcation, then these models all rescale to Φ^4_3 near their critical point.

Keywords Phase coexistence · Polynomial potential · Non-Gaussian

1 Introduction

The dynamical Φ^4_d model is formally given by

$$\partial_t \Phi = \Delta \Phi - \lambda \Phi^3 + \xi, \tag{1.1}$$

where ξ is the space-time white noise in d spatial dimensions (on the torus $\mathbf{T}^d$), namely the Gaussian random field with covariance formally given by

$$\mathbf{E}\xi(s,x)\xi(t,y) = \delta(s-t)\delta^{(d)}(x-y).$$

✉ Weijun Xu
weijunx@gmail.com

Hao Shen
pkushenhao@gmail.com

1 Columbia University, New York, NY, USA

2 University of Warwick, Coventry, UK

This equation was initially derived from the stochastic quantisation of the Euclidean quantum field theory [26]. It is also believed to be the scaling limit for Kac-Ising model near critical temperature (conjectured in [14]). Note that formally, one could rescale Φ to turn the coefficient in front of the cubic term to be 1, but we still retain that λ here in order to simplify the scaling later.

In dimension $d = 1$, the equation is well-posed, and it was shown in [7] that the dynamic Ising model on the real line does rescale to it. However, as soon as $d \geq 2$, the equation becomes ill-posed in the sense that the "solution" Φ is distribution valued so that the non-linearity Φ^3 does not make sense. In $d = 2$, the solution could be constructed either via Dirichlet forms [4] or via Wick ordering ([10], by turning the term Φ^3 into the Wick product $:\Phi^3:$ with respect to the Gaussian structure induced by the linearised equation); and [23] constructed the global solution on the entire plane. Recently, Mourrat and Weber [22] showed that two dimensional Ising model with Glauber dynamics and Kac interaction does converge to the Wick ordered solution constructed in [10]; see also [27] for the study of a more general ferromagnetic interaction (Blume-Capel model) which is shown to converge to either Φ^4_2 or Φ^6_2 according to the specific regimes.

The three dimensional case is much more involved as it exhibits further logarithmic divergence beyond Wick ordering. It became amenable to analysis only very recently, via the theory of regularity structures [15], controlled para-products [8,12], or Wilson's renormalisation group approach [20]. In all these cases, the solution is obtained as the limit of smooth solutions Φ_ϵ to the equations

$$\partial_t \Phi_\epsilon = \Delta \Phi_\epsilon - \lambda \Phi_\epsilon^3 + \xi_\epsilon + C_\epsilon \Phi_\epsilon, \tag{1.2}$$

where ξ_ϵ is a regularisation of the noise ξ at scale ϵ, and $C_\epsilon \to +\infty$ is the *renormalisation constant* which contains a Wick renormalisation that diverges as ϵ^{-1} and a logarithmically divergent constant. Furthermore, these constants can be chosen such that the limit of Φ_ϵ is *independent* of the actual regularisation used. The global solution in three spatial dimensions is constructed by [24]. For $d \geq 4$, one does not expect to get any nontrivial limit [1,6,11].

The Eq. (1.1) in three dimensions is believed to be the universal model for phase coexistence near criticality. In the recent work [19], Hairer and the second author of the present article studied phase coexistence models in three spatial dimensions of the type

$$\partial_t u = \Delta u - \epsilon V'_\theta(u) + \zeta, \tag{1.3}$$

where ζ is a space-time Gaussian random field with correlation length 1, and V_θ is a polynomial potential whose coefficients depends smoothly on a parameter θ. The main result of [19], built on analogous results in the context of the KPZ equation in [17] as well as the general theory of regularity structures [15], is that if V is symmetric and its certain average exhibits a "pitchfork bifurcation near the origin" (see [19, Section 1] or Assumption 1.4 below), then there exists a choice of

$$\theta = a\epsilon \log \epsilon + \mathcal{O}(\epsilon)$$

such that the rescaled processes (also called macroscopic processes)

$$u_\epsilon(t,x) := \epsilon^{-\frac{1}{2}} u(t/\epsilon^2, x/\epsilon), \tag{1.4}$$

converge to the solution of the Φ^4_3 equation, interpreted as the limit in (1.2), and the coefficient λ of the cubic term in the limit is a linear combination of all coefficients of the polynomial V_0. We call this limit the $\Phi^4_3(\lambda)$ family, where the term "family" comes from the fact that one could perturb the finite part of the infinite renormalisation constant, and end up with a one-parameter family of limits.

The main purpose of this article is to show that when ζ is a symmetric, stationary noise with strong mixing assumption but *not necessarily Gaussian*, the macroscopic process u_ϵ still converges to the Φ^4_3 family with a proper choice of θ (depending on ϵ). Also, by the recent global existence result [16,24], the convergence is actually global in time.

Our main theorem can be loosely stated as follows (for the exact statement, including the precise notion of convergence, see Theorem 5.2).

Theorem 1.1 *Let ζ and V_θ satisfy Assumptions 1.3 and 1.4 below. Let u (depending on both ϵ and θ) be the solution to (1.3). Then, there exists $a < 0$ such that for $\theta = a\epsilon|\log\epsilon| + \mathcal{O}(\epsilon)$, the rescaled processes u_ϵ in (1.4) converge in law to the $\Phi^4_3(\widehat{a}_1)$ family, where $\widehat{a}_1$ is given by (1.8) below.*

Remark 1.2 The symmetry assumption on both the potential and the (law of) noise is essential to get Φ^4_3 as the limit. In [19], it was shown that with Gaussian (symmetric) noise but non-symmetric potential, the system would in general rescale to either OU process or Φ^3_3, depending on the value of θ. We expect the same phenomena to happen with non-symmetric noise even if the potential is symmetric.

We now give precise assumptions on the noise ζ and the potential V_θ in Theorem 1.1, starting from the noise.

Assumption 1.3 We assume the random field ζ defined on $\mathbf{R} \times (\epsilon^{-1}\mathbf{T})^3$ satisfies the following properties:

1. ζ is symmetric (in the sense that $\zeta \overset{\text{law}}{=} -\zeta$), stationary, continuous, and $\mathbf{E}|\zeta(z)|^p < +\infty$ for all $p > 1$.
2. ζ is normalised so that $\varrho(z) = \mathbf{E}\zeta(0)\zeta(z)$ integrates to 1.
3. For any two compact sets $K_1, K_2 \subset \mathbf{R} \times (\epsilon^{-1}\mathbf{T})^3$ with distance at least 1 away from each other, the two sigma fields $\mathcal{F}_{K_1}$ and $\mathcal{F}_{K_2}$ are independent, where $\mathcal{F}_{K_i}$ is the σ-algebra generated by the point evaluations $\{\zeta(z) : z \in K_i\}$.

Typical examples of such random fields include smeared-out versions of Poisson point processes with uniform space-time intensity. We refer to [18] for more details and related examples (see also the discussions below Assumption 1.2 and Example 2.3 in [18]). Given such a random field ζ, we let Ψ be the stationary solution to the equation

$$\partial_t \Psi = \Delta\Psi + \zeta. \tag{1.5}$$

Since we are in three spatial dimensions, the square of the heat kernel is integrable at large scales, so the stationary solution does exist. Let μ denote the distribution of Ψ evaluated at a space-time point (the actual point does not matter since Ψ is stationary), and define the averaged potential $\langle V_\theta \rangle$ to be

$$\langle V_\theta\rangle(x) := \int_{\mathbf{R}} V_\theta(x+y)\mu(dy). \tag{1.6}$$

Our assumption on V is the following.

Assumption 1.4 Let μ denote the distribution of the stationary solution to (1.5), where ζ is a random field satisfying Assumption 1.3. Let $V : (\theta, x) \mapsto V_\theta(x)$ be a symmetric polynomial in x with coefficients depending smoothly on θ. We assume that the averaged potential $\langle V_\theta\rangle$ obtained from V_θ via (1.6) has a pitchfork bifurcation near the origin in the sense that

$$\frac{\partial^4\langle V\rangle}{\partial x^4}(0,0) > 0, \qquad \frac{\partial^2\langle V\rangle}{\partial x^2}(0,0) = 0, \qquad \frac{\partial^3\langle V\rangle}{\partial\theta\ \partial x^2}(0,0) < 0. \tag{1.7}$$

Note that it is always the derivative V'_θ rather than V_θ that appears in the dynamics, so the constant term of V will not play any role in the context. We suppose $\langle V'_\theta\rangle$ has the form

$$\langle V_\theta\rangle'(x) = \sum_{j=0}^{m} \widehat{a}_j(\theta)x^{2j+1}, \tag{1.8}$$

which only has odd powers since both V_θ and μ in (1.6) are symmetric, then the pitchfork bifurcation assumption (1.7) reads

$$\widehat{a}_1 > 0, \qquad \widehat{a}_0 = 0, \qquad \widehat{a}'_0 < 0, \tag{1.9}$$

where we have used the notation $\widehat{a}_j = \widehat{a}_j(0)$ and $\widehat{a}'_j = \widehat{a}'_j(0)$ for simplicity. Note that the coefficients $\widehat{a}_j(\theta)$ here come in front of different powers from those in [19], since in [19] the potential V_θ is not assumed to be symmetric, and has both even and odd powers.

We briefly explain at this point the pitchfork bifurcation condition in the above Assumption (in particular why it naturally appears on $\langle V\rangle$ rather than V). Note that if u solves the microscopic Eq. (1.3), then the rescaled process u_ϵ defined in (1.4) solves the equation

$$\partial_t u_\epsilon = \Delta u_\epsilon - \sum_{j=0}^{m} a_j(\theta)\epsilon^{j-1}u_\epsilon^{2j+1} + \zeta_\epsilon, \tag{1.10}$$

where $a_j(\theta)$'s are the coefficients of V'_θ, and

$$\zeta_\epsilon(t,x) = \epsilon^{-\frac{5}{2}}\zeta(t/\epsilon^2, x/\epsilon) \tag{1.11}$$

is the rescaled noise. Heuristically, as $\epsilon \to 0$, one can think of u_ϵ^2 as "of order ϵ^{-1}" (for instance the expected value of u_ϵ^2 at a certain space-time point is given by the diagonal value of the regularised Green's function which in three space dimensions diverges at rate ϵ^{-1}). So each of the terms $\epsilon^{j-1}u_\epsilon^{2j+1}$ ($j \geq 1$) has a "diverging part" of order $\epsilon^{-1}u_\epsilon$, and the combined effect of these divergences (together with the term $a_0(\theta)\epsilon^{-1}u_\epsilon$ for $j = 0$ in 1.10) is then a constant multiple of $\epsilon^{-1}u_\epsilon$. It turns out that the value of this constant is nothing but $\int_{\mathbf{R}} V_0'(y)\mu(dy)$, and by (1.6) and (1.8) it is precisely $\widehat{a}_0$. Thus $\widehat{a}_0 = 0$ in (1.9) is a necessary condition for the right hand side of (1.10) to converge. In fact, as we will see, the condition $\widehat{a}_0 = 0$ automatically takes account of the "Wick renormalisation" (for *non-Gaussian* noise ζ_ϵ) and thus the coefficients of these divergent terms with order ϵ^{-1} will precisely cancel out each other. As mentioned above (right after 1.2), there will be further logarithmic divergence, but this divergence can be renormalised by a suitable choice of the small parameter θ thanks to $\widehat{a}_0' \neq 0$. [1] Finally, $\widehat{a}_1 > 0$ is needed for the final limit to be the $\Phi^4_3(\widehat{a}_1)$ family and exists globally: $\widehat{a}_1 = 0$ would lead to a Gaussian limit and $\widehat{a}_1 < 0$ would result in a solution which blows up in finite time.

Having stated the assumptions and the main result of the paper, we would like to explain the *main challenge* we overcome in this article. The proof of our universality result relies on the theory of regularity structures [15] (see Sect. 3 for a brief review of the basic concepts of this theory), and an essential ingredient in the implementation of this theory is to prove convergence of certain stochastic terms (called renormalised random models, see Sect. 3.1 for the notion of "models" and Sect. 3.3 for "renormalised" random models). The main challenge is to deal with two difficulties simultaneously: on one hand one should handle stochastic terms which are *arbitrary* polynomials of the noise, and on the other hand with the non-Gaussian noise one has to control the moments of *arbitrary* order of these stochastic terms; and the two difficulties intertwine with each other and thus make the problem nontrivial.

Indeed, if the random models are built from a Gaussian noise, then by hypercontractivity or equivalence of moments (see for instance [15, Lemma 10.5]), it suffices to bound their second moments to prove convergence. However, for non-Gaussian noise, one has to bound the p-th moments of the random models "by hand" for all p, which involves analysing complicated expressions which are generalised convolutions of singular kernels. [17, Appendix A] developed *graphic machinery* for estimating generalised convolutions, and reduced the estimates of these complicated convolutions to verifications of certain conditions on "labelled graphs" (see for instance [17, Section 6] for many examples of such graphs). The p-th moment of a random object is represented by a sum of certain labelled graphs, each obtained by "gluing together" the nodes of p identical "trees" in a certain way, and each of the identical trees represents that random object. When p is arbitrarily large which makes the labelled graph to be arbitrarily large, it is in general very hard to keep track of all the conditions that need to be checked; so in [18, Section 4, 5] and [9, Section 3], the verification procedure was further reduced to checking some conditions for every "sub-tree" of a single tree (instead of the conditions for the large graph consisting of p trees with p arbitrarily large) that represents the random object (see Assumption 4.4 and Theorem 4.6 below). This significantly simplifies the verification; see for example [18] for application to the KPZ equation.

However, in the current problem, the Eq. (1.10) involves arbitrarily large powers of u_ϵ in contrast to [18] where only the second power appears in the case of KPZ

[1] In general one only needs to assume $\widehat{a}_0' \neq 0$. The reason we actually assume $\widehat{a}_0' < 0$ is just a matter of convention: in this case, $\theta > 0$ corresponds to a double well potential and $\theta < 0$ corresponds to a single well potential - which is the usual convention of pitchfork bifurcation.

equation. As a result, the trees have arbitrarily rich ramification so that even for each single tree, it will be a highly nontrivial task to check the aforementioned conditions in Assumption 4.4 for *every* sub-tree of it. In addition, trees built from the nonlinearity with high powers of u_ϵ will in general fail the conditions (1)–(3) in Assumption 4.4. To fix the last problem, we employ the positive "homogeneities" represented by the multiplication of ϵ in (1.10), and incorporate them to render trees which do satisfy the desired conditions (see the beginning of Sect. 4.4). Furthermore, to analyse trees with arbitrarily rich ramification and efficiently check the conditions for them (after incorporating the multiplication of ϵ), we make a key observation in Lemma 4.11 below, which shows that one only needs to check the conditions for only a few sub-trees, and the verification for all other sub-trees will automatically follow from that. We expect these treatments and simplifications will apply to other situations, for example proving universality of the KPZ equation for polynomial microscopic growth models with non-Gaussian noise.

The rest of the article is organised as follows. In Sect. 2, we briefly recall the facts of non-Gaussian random variables and their Wick powers. Section 3 is devoted to the construction of the regularity structures, the abstract fixed point equation as well as the renormalisation group. Most of the set-up in these sections can be found in [15,18,19], so we keep the presentation to a minimum and refer to those articles for more details. In Sect. 4, we prove the bounds for the renormalised models constructed in Sect. 3.3 with proper renormalisation constants. These bounds are necessary to prove the convergence of renormalised models. Finally, in Sect. 5, we collect the results in all previous sections to prove the convergence of our models to the desired Φ^4_3 limit at large scales.

2 Non-Gaussian Wick powers and averaged potential

In this section, we review some basic properties of Wick polynomials for non-Gaussian random variables. We keep the presentation to the minimum here and only give definitions and statements that will be used later. More details can be found in [18, Section 3].

2.1 Joint cumulants

We start with the definition 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 short-hands for

$$X_B = \{X_\alpha : \alpha \in B\}, \qquad X^B = \prod_{\alpha \in B} X_\alpha. \tag{2.1}$$

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 $\bigcup \pi = B$ and such that any two distinct elements of π are disjoint.

Definition 2.1 Given a collection $\mathcal{X}$ of random variables as above and any finite set $B \subset A$, we define the cumulant $\mathbf{E}_c(X_B)$ inductively over $|B|$ by

$$\mathbf{E}(X^B) = \sum_{\pi \in \mathcal{P}(B)} \prod_{\bar{B} \in \pi} \mathbf{E}_c(X_{\bar{B}}). \tag{2.2}$$

Remark 2.2 It is straightforward to check that the above definition does determine the cumulants uniquely, see the discussion below [18, Definition 3.2]. We refer to [25] and [13, Section 13.5] for the properties of joint cumulants; see also the recent article [21].

From now on, we will use the notation $\mathfrak{C}_n$ for the nth joint cumulant function of the field ζ:

$$\mathfrak{C}_n(z_1, \ldots, z_n) = \mathbf{E}_c\big(\{\zeta(z_1), \ldots, \zeta(z_n)\}\big). \tag{2.3}$$

We similarly write $\mathfrak{C}_n^{(\epsilon)}$ for the n-th cumulant but with $\zeta(z_j)$'s replaced by $\zeta_\epsilon(z_j)$'s. Note that $\mathfrak{C}_{2n+1} = 0$ since ζ is assumed to be symmetric. Also, $\mathfrak{C}_2$ is its covariance function. An important property is that there exists $C > 0$ such that

$$\mathfrak{C}_p^{(\epsilon)}(z_1, \ldots, z_p) = 0 \tag{2.4}$$

whenever $|z_i - z_j| > C\epsilon$ for some i, j.

2.2 Wick polynomials

The notion of Wick products for random variables (not necessarily Gaussian) will play an essential role later. We give a definition below.

Definition 2.3 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 by setting $:X_\emptyset: = 1$ and postulating that

$$X^A = \sum_{B \subset A} :X_B: \sum_{\pi \in \mathcal{P}(A \setminus B)} \prod_{\bar{B} \in \pi} \mathbf{E}_c\big(X_{\bar{B}}\big) \tag{2.5}$$

recursively.

Remark 2.4 Again, (2.5) is sufficient to define $:X_A:$ by induction over the size of A. 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 taking expectations on both sides of (2.5) yields exactly the identity (2.2).

Note that by (2.2), we can alternatively write

$$X^A = \sum_{B \subset A} :X_B: \mathbf{E}(X^{A \setminus B}). \tag{2.5$'$}$$

There is also a formula to express the Wick product in terms of the usual products; however, (2.5) is actually the identity frequently being used in the paper: in fact we will frequently rewrite a product of generally non-Gaussian noises as a sum of terms, each term containing a Wick product as RHS of (2.5)—called the *non-Gaussian homogeneous chaos of order* $|B|$.

It is also well-known (see for instance [5, Eq. (2.5, 2.6)]) that there is an alternative characterisation of Wick products via generating functions by

$$:X_1 \cdots X_n: \; = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \frac{\exp\left(\sum_{i=1}^n t_i X_i\right)}{\mathbf{E}\exp\left(\sum_{i=1}^n t_i X_i\right)}\Big|_{t_1=\cdots=t_n=0}. \tag{2.6}$$

In the case $X_1 = \cdots = X_n$ all with distribution μ, (2.6) can be written as

$$:X^n: \; = \frac{\partial^n}{\partial t^n}\left(\frac{e^{tX}}{\mathbf{E}_\mu e^{tY}}\right)\Big|_{t=0} \tag{2.7}$$

where Y is distributed according to μ. Note that there is no n in the exponential generating function above, as the derivative is taken with respect to the same t rather than $t_1, \ldots, t_n$ separately. The form of (2.7) also suggests that we can actually define the *n-th Wick power* with respect to a measure μ as

$$W_{n,\mu}(x) = \frac{\partial^n}{\partial t^n} \frac{\exp\left(tx\right)}{\mathbf{E}_\mu \exp\left(tY\right)}\Big|_{t=0}. \tag{2.8}$$

Now, $W_{n,\mu}(\cdot)$ is a polynomial in the variable $x \in \mathbf{R}$, rather than the random variable X. One can immediately check using the definition (2.8) that $\frac{d}{dx} W_{n,\mu}(x) = n W_{n-1,\mu}$, and this means that the Wick powers form an *Appell sequence* (see [2, Section 2.2] or the original paper by Appell [3]). As a comparison to the Wick powers, we recall that the usual monomials can be generated by

$$x^n = \frac{\partial^n}{\partial t^n} e^{tx}\Big|_{t=0}. \tag{2.9}$$

We then have the following important lemma.

Lemma 2.5 *Let* μ *be a probability measure on* $\mathbf{R}$ *with finite moments of all order, and* V *be a polynomial. Let*

$$\langle V \rangle(x) = \int_{\mathbf{R}} V(x+y)\mu(dy) = \mathbf{E}_\mu V(x+Y)$$

be the average of V *against the measure* μ. *Then,* $\langle V \rangle(x) = x^n$ *if and only if* $V(x) = W_{n,\mu}(x)$.

Proof We first prove the "if" part. Let $V(x) = W_{n,\mu}(x)$, then by definitions of $W_{n,\mu}$ and $\langle\cdot\rangle$, we have

$$\langle V\rangle(x) = \mathbf{E}_\mu \frac{\partial^n}{\partial t^n}\Big(\frac{e^{tx}e^{tY}}{\mathbf{E}_\mu e^{tZ}}\Big)\Big|_{t=0} = \frac{\partial^n}{\partial t^n}\Big(\frac{e^{tx}\mathbf{E}_\mu e^{tY}}{\mathbf{E}_\mu e^{tZ}}\Big)\Big|_{t=0} = \frac{\partial^n}{\partial t^n}e^{tx}\Big|_{t=0} = x^n,$$

where for the second equality we have exchanged the differentiation with the expectation with respect to Y. For the "only if" part, suppose U is another polynomial with $\langle U\rangle = x^n$, then $U - V$ is a polynomial such that

$$\int_{\mathbf{R}} (U - V)(x + y)\mu(dy) = 0$$

for all $x \in \mathbf{R}$. It then follows easily that all coefficients of $U - V$ must be 0. Indeed, suppose that $\int_{\mathbf{R}} \sum_{i=0}^n b_i(x + y)^i \mu(dy) = 0$ where $b_n \neq 0$, then the leading order in x reads $b_n x^n \int_{\mathbf{R}} \mu(dy)$ which must vanish and thus $b_n = 0$, contradicting with the assumption $b_n \neq 0$. Therefore we have $U = V = W_{n,\mu}$. □

Lemma 2.5 allows us to rewrite (1.10) by replacing the powers by Wick powers and at the same time replacing the coefficients of V_θ' by the coefficients of $\langle V_\theta\rangle'$ (see 1.6 and 1.8). Thus the rescaled process u_ϵ satisfies the equation

$$\partial_t u_\epsilon = \Delta u_\epsilon - \sum_{j=0}^m \widehat{a}_j(\theta)\epsilon^{j-1} W_{2j+1,\mu_\epsilon}(u_\epsilon) + \zeta_\epsilon. \tag{2.10}$$

3 Regularity structures and abstract equation

In this section, we build the appropriate regularity structures in order to solve the Eq. (1.10) for the rescaled process u_ϵ, which we rewrite here

$$\partial_t u_\epsilon = \Delta u_\epsilon - \epsilon^{-\frac{3}{2}} V_\theta'(\epsilon^{\frac{1}{2}} u_\epsilon) + \zeta_\epsilon \tag{3.1}$$

in an abstract space of *modelled distributions*. The set-up is essentially identical to that in [19]. For the sake of simplicity of the presentation, we only recall the necessary set-ups here, and refer the readers to [19, Sections 2, 3] for details.

3.1 Regularity structures and admissible models

Recall that *a regularity structure* consists of a pair $(\mathcal{T}, \mathcal{G})$, where $\mathcal{T} = \bigoplus_{\alpha\in A} \mathcal{T}_\alpha$ is a vector space graded by a (bounded below, locally finite) set A of homogeneities, and the structure group $\mathcal{G}$ is a group of linear transformations on $\mathcal{T}$ such that for every $\Gamma \in \mathcal{G}$ and $\tau \in \mathcal{T}_\alpha$, we have $\Gamma\tau - \tau \in \mathcal{T}_{<\alpha}$.

The specific regularity structure for the study of the problem (1.10) is defined as follows. Similar as [19, Section 2.1], the basis elements in $\mathcal{T}$ are built from the symbols $\mathbf{1}$, Ξ, $\{X_i\}_{i=0}^3$ and the abstract operators $\mathcal{I}$ and $\mathcal{E}^k$, where k takes values in positive integers (unlike [19] where k takes values in positive half integers). Here the symbol $\mathcal{I}$ should be thought of as "integration" and the symbol $\mathcal{E}$ should be thought of as "multiplication by ϵ". In order to reflect the parabolic scaling of the equation, for any multi-index of non-negative integers $k = (k_0, \dots, k_3)$, we define an abstract monomial of degree $|k|$ as

$$X^k = X_0^{k_0} \cdots X_3^{k_3}, \qquad |k| = 2k_0 + \sum_{i=1}^{3} k_i, \tag{3.2}$$

and the parabolic distance between $z, z' \in \mathbf{R}^4$ by

$$|z - z'| = |z_0 - z_0'|^{\frac{1}{2}} + \sum_{i=1}^{3} |z_i - z_i'|. \tag{3.3}$$

In view of the structure of (3.1), we define two sets $\mathcal{U}$ and $\mathcal{V}$ such that $X^k \in \mathcal{U}$, $\Xi \in \mathcal{V}$, and such that for every $k = 1, \dots, m$, we decree

$$\begin{aligned} \tau_1, \dots \tau_{2k+1} \in \mathcal{U} \quad &\Rightarrow \quad \mathcal{E}^{k-1}(\tau_1 \cdots \tau_{2k+1}) \in \mathcal{V}, \\ \tau \in \mathcal{V} \quad &\Rightarrow \quad \mathcal{I}(\tau) \in \mathcal{U}. \end{aligned} \tag{3.4}$$

The idea is that $\mathcal{U}$ consists of the formal symbols in the expansion of the abstract solution of the equation, and $\mathcal{V}$ consists of symbols that will appear in the expansion of the right hand side of (3.1). We then let $\mathcal{W} = \mathcal{U} \cup \mathcal{V}$ to be the set of basis elements in $\mathcal{T}$. As for the homogeneities, we set

$$|\Xi| = -\frac{5}{2} - \kappa, \qquad |X^\ell| = |\ell|,$$

and extend it to all of $\mathcal{W}$ decreeing

$$|\tau \bar{\tau}| = |\tau| + |\bar{\tau}|, \quad |\mathcal{I}(\tau)| = |\tau| + 2, \quad |\mathcal{E}^k(\tau)| = k + |\tau|.$$

Here, κ is a small but fixed positive number, and $|\ell|$ denotes the parabolic degree of the multi-index defined in (3.2). According to (3.4), we can keep adding new basis elements into $\mathcal{U}$ and $\mathcal{V}$, but for the purpose of this article, it suffices to restrict our regularity structures to elements with homogeneities less than $\frac{3}{2}$, i.e., $\mathcal{T} = \bigoplus_{\alpha < \frac{3}{2}} \mathcal{T}_\alpha$.

The basis vectors in $\mathcal{W}$ generated this way with negative homogeneities other than Ξ are of the form (with the shorthand $\Psi = \mathcal{I}(\Xi)$):

$$\begin{aligned} &\tau = \mathcal{E}^{k-1} \Psi^{2k+1-n}, && |\tau| = \frac{1}{2}(n-3) - (2k+1-n)\kappa\ ; \\ &\tau = \mathcal{E}^{\lfloor (k-1)/2 \rfloor} \left(\Psi^k \mathcal{I}(\mathcal{E}^{\lfloor \ell/2 \rfloor - 1} \Psi^\ell) \right), && |\tau| = \left\lfloor \frac{k-1}{2} \right\rfloor + \left\lfloor \frac{\ell}{2} \right\rfloor - (k+\ell)\left(\frac{1}{2} + \kappa\right) + 1 \end{aligned} \tag{3.5}$$

provided κ is sufficiently small, where in the first case $n \in \{0, 1, 2, 3\}$, and in the second one the situation when k odd and ℓ even is excluded.

As a matter of fact, the scaling, the space $\mathcal{T}$ together with the homogeneities we defined here reflect the fact that (3.1) is *subcritical*: as in [19, Lemma 2.1], one has that $\{\tau \in \mathcal{W} : |\tau| < \gamma\}$ is a finite set for every $\gamma > 0$, provided that κ is sufficiently small.

It will be convenient to view $\mathcal{E}^k$ itself as a linear map: $\tau \to \mathcal{E}^k(\tau)$. For this we introduce the graded vector space of *extended* regularity structure $\mathcal{T}_{\text{ex}}$, which is the linear span of symbols in

$$\mathcal{W} \cup \{\tau_1, \dots, \tau_{2m+1} : \tau_j \in \mathcal{U}\}.$$

Now $\mathcal{E}^k$ can be viewed as a linear map defined on $\mathcal{T}_{\text{ex}}$.

In order to define the structure group $\mathcal{G}$, we introduce a set $\mathcal{T}_+$ which is the free commutative algebra generated by symbols in $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ where

$$\mathcal{F}_1 = \{\mathbf{1}, X\}\ , \qquad \mathcal{F}_2 = \{\mathscr{J}_\ell(\tau) :\ \tau \in \mathcal{W}\backslash\{X^k\},\ |\tau| + 2 > |\ell|\}\ ,$$
$$\mathcal{F}_3 = \{\mathscr{E}_\ell^{k-1}(\tau_1 \cdots \tau_{2k+1}) :\ \tau_j \in \mathcal{U},\ k - 1 + \sum_j |\tau_j| > |\ell| \geq \sum_j |\tau_j|\}.$$

We then define the linear map $\Delta : \mathcal{T}_{\text{ex}} \to \mathcal{T}_{\text{ex}} \otimes \mathcal{T}_+$ in the same way as in [17, Section 3.2]. For any linear functional $g : \mathcal{T}_+ \to \mathbf{R}$, one obtains a linear map $\Gamma_g : \mathcal{T}_{\text{ex}} \to \mathcal{T}_{\text{ex}}$ by setting $\Gamma_g \tau = (\mathrm{id} \otimes g)\Delta\tau$. The *structure group* $\mathcal{G}$ is then defined by the collection of all maps Γ_g for linear functionals $g : \mathcal{T}_+ \to \mathbf{R}$ with the further property $g(\tau\bar{\tau}) = g(\tau)g(\bar{\tau})$. In fact $\mathcal{G}$ preserves not only $\mathcal{T}_{\text{ex}}$, but also $\mathcal{T}$; restricting $\mathcal{G}$ on $\mathcal{T}$ then gives the regularity structure $(\mathcal{T}, \mathcal{G})$.

Remark 3.1 One may think of the elements $\mathscr{J}_\ell\tau$ and $\mathscr{E}_\ell^k\tau$ essentially the same as the ℓ-th derivative of $\mathcal{I}(\tau)$ and $\mathcal{E}^k(\tau)$. But we use different notation for these elements since they will play a different role than the corresponding elements in $\mathcal{T}$.

Given a regularity structure $(\mathcal{T}, \mathcal{G})$, *a model for it* is a pair (Π, F) of functions

$$\begin{aligned} \Pi : \mathbf{R}^{1+3} &\to \mathcal{L}(\mathcal{T}, D') \qquad & F : \mathbf{R}^{1+3} &\to \mathcal{G} \\ z &\mapsto \Pi_z & z &\mapsto F_z \end{aligned}$$

satisfying the identity

$$\Pi_z F_z^{-1} = \Pi_{\bar{z}} F_{\bar{z}}^{-1}, \qquad \forall z, \bar{z} \in \mathbf{R}^{1+3}$$

as well as the bounds

$$|(\Pi_z\tau)(\varphi_z^\lambda)| \lesssim \lambda^{|\tau|}, \qquad |F_z^{-1} \circ F_{\bar{z}}|_\sigma \lesssim |z - \bar{z}|^{|\tau| - |\sigma|} \tag{3.6}$$

uniformly over all test functions $\varphi \in \mathcal{B}$, all space-time points $z, \bar{z}$ in a compact domain $\mathfrak{K}$, and all $|\tau| \in \mathcal{W}$. Here, the set $\mathcal{B}$ denotes the space of test functions

$$\mathcal{B} = \big\{\varphi : \|\varphi\|_{\mathcal{C}^2} \leq 1, \operatorname{supp}(\varphi) \subset B(0,1)\big\},$$

where $B(0, 1)$ is the unit ball centred at origin, and

$$\varphi^{\lambda}_{t',x'}(t,x) = \lambda^{-5}\varphi\big((t-t')/\lambda^2, (x-x')/\lambda\big).$$

We will also use the notation f_z for the multiplicative element in $\mathcal{T}_+^*$ (the dual space of $\mathcal{T}_+$) such that $F_z = \Gamma_{f_z}$ (recall that by definition of $\mathcal{G}$, $F_z \in \mathcal{G}$ must be of this form).

For a model $\mathfrak{M} = (\Pi, F)$ we denote by $|\!|\!|\mathfrak{M}|\!|\!|$ the smallest proportionality constant such that both bounds in (3.6) hold. $|\!|\!|\mathfrak{M}|\!|\!|$ in general depends on the compact domain $\mathfrak{K}$, but we omit it from the notation for simplicity. Since in most situations the structure group element F is completely determined by Π, we will also sometimes write $|\!|\!|\Pi|\!|\!|$ instead of $|\!|\!|\mathfrak{M}|\!|\!|$.

In addition to these minimal requirements on the models, we also impose *admissibility* for our models, which means that the model is "compatible" with a (truncated) heat kernel in the following sense. Let $K : \mathbf{R}^{1+3} \setminus \{0\} \to \mathbf{R}$ be a function that coincides with the heat kernel on $\{|z| \leq 1\}$, smooth everywhere except the origin, and annihilates all polynomials on $\mathbf{R}^{1+3}$ up to parabolic degree 2. The existence of such a kernel can be shown, for example see [15, Section 5]. With this truncated heat kernel K, we then define the set $\mathscr{M}$ of *admissible models*, namely the models (Π, f) which satisfy

$$(\Pi_z X^k)(\bar{z}) = (\bar{z} - z)^k, \qquad f_z(X^k) = (-z)^k, \qquad \forall k \tag{3.7}$$

as well as

$$\begin{aligned} f_z(\mathscr{J}_\ell \tau) &= -\int D^\ell K(z-\bar{z})(\Pi_z\tau)(d\bar{z}), \qquad 0 \leq |\ell| < |\tau| + 2 \\ \big(\Pi_z \mathcal{I}(\tau)\big)(\cdot) &= (K * \Pi_z \tau)(\cdot) + \sum_{0 \leq |\ell| < |\tau|+2} \frac{(\cdot - z)^\ell}{\ell!} \cdot f_z(\mathscr{J}_\ell \tau), \end{aligned} \tag{3.8}$$

and we set $f_z(\mathscr{J}_\ell\tau) = 0$ if $|\ell| \geq |\tau| + 2$. Note that the above notion of admissible models do not impose any restrictions on the operator $\mathcal{E}^k$.

Given a smooth function ζ and $\epsilon \geq 0$, we now build *the canonical model* $\mathscr{L}_\epsilon(\zeta) = (\Pi^\epsilon, f^\epsilon)$ as follows. Define the action of $\mathscr{L}_\epsilon(\zeta)$ on X^k according to (3.7), and set set

$$(\Pi_z \Xi)(\bar{z}) = \zeta(\bar{z}).$$

We then extend the definition to all of $\mathcal{W}$ by imposing the admissibility restriction (3.8), as well as setting recursively

$$(\Pi_z^\epsilon \tau\bar{\tau})(\bar{z}) = (\Pi_z^\epsilon \tau)(\bar{z}) \cdot (\Pi_z^\epsilon \bar{\tau})(\bar{z}) \tag{3.9}$$

and

$$f_z^\epsilon(\mathscr{E}_\ell^k\tau) = -\epsilon^k(D^\ell\Pi_z\tau)(z),$$
$$(\Pi_z^\epsilon\mathcal{E}^k\tau)(\bar z) = \epsilon^k(\Pi_z^\epsilon\tau)(\bar z) + \sum_{|\ell|<k+|\tau|}\frac{(\bar z - z)^\ell}{\ell!}\cdot f_z^\epsilon(\mathscr{E}_\ell^k\tau), \tag{3.10}$$

where we only sum over those multi-indices ℓ's such that $|\ell| < k + |\tau|$. Note that the right hand sides of (3.9) and (3.10) make sense only if $\Pi_z^\epsilon\tau$ is a smooth function for every τ, and this is indeed the case if the input ζ is smooth.

Remark 3.2 It can be verified that for any smooth space-time function ζ, the canonical model $\mathscr{L}_\epsilon(\zeta)$ is indeed an admissible model (for instance the identity $\Pi_z F_z^{-1} = \Pi_{\bar z} F_{\bar z}^{-1}$ indeed holds), see [17, Section 3.6].

The above framework of regularity structures and models allows one to solve the fixed point problem for a PDE in a space of *modelled distributions*, denoted by $\mathcal{D}^\gamma$ [15, Section 3], or more precisely in the space $\mathcal{D}^{\gamma,\eta}$ in view of the initial condition of the equation [15, Section 6]. The space $\mathcal{D}^{\gamma,\eta}$ consists of mappings $U : \mathbf{R}^d \to \mathcal{T}$ endowed with norm $\|U\|_{\gamma,\eta}$, see [15, Definition 6.2], which is actually just the second and third terms on the RHS of (3.14) with $\epsilon = 1$ below. It comes with an important reconstruction map $\mathcal{R}$ (as defined in [15, Section 3]), which maps a modelled distribution to a usual distribution, and thus gives a solution to the equation in the usual sense. For the weak universality problems we are concerned with, a necessary modification of this space comes from a difficulty related with initial data. We recall this modification of spaces in the next subsection now.

3.2 ϵ-dependent spaces and the abstract fixed point equation

Since the solution to standard Φ_3^4 equation possesses regularity below $-\frac{1}{2}$, in a fixed point argument, in order to continue local solutions, we need to be able to treat initial data of the same regularity. However, if the initial data belongs to $\mathcal{C}^\eta$ with $\eta < -\frac{1}{2}$, then $u_\epsilon(t,\cdot) \sim t^{-\frac{\eta}{2}}$ for small time t, and for any fixed ϵ, the term $\epsilon^{k-1}u_\epsilon^{2k+1}$ becomes non-integrable as soon as $k \geq 2$. On the other hand, the case we are mostly interested in is the limiting equation as $\epsilon \to 0$. The idea implemented in [17] to achieve this goal is to employ cancellations between positive power of ϵ and singularities in t in the process as $\epsilon \to 0$ while retaining uniform (in ϵ) bounds. This leads to the definition of ϵ-dependent models and spaces. The definitions below mainly follow [17, Section 4.1] and [19, Section 3].

For each $\epsilon \geq 0$, let $\mathscr{M}_\epsilon$ be the collections of models in $\mathscr{M}$ with norm

$$\|\!|\Pi|\!\|_\epsilon := \|\!|\Pi|\!\| + \|\Pi\|_\epsilon + \sup_{t\in[0,1]}\|\mathcal{R}\mathcal{I}(\Xi)(t,\cdot)\|_{\mathcal{C}^\eta}, \tag{3.11}$$

where $\|\!|\cdot\|\!|$ is the standard norm on modelled distributions, $\mathcal{R}$ is the reconstruction operator associated to the underlying model, and $\|\cdot\|_\epsilon$ is defined by

$$\begin{aligned}\|\Pi\|_\epsilon := &\sup_{\tau\in\mathcal{W}}\sup_z\sup_{k,\ell}\epsilon^{|\ell|-k-|\tau|}|f_z(\mathscr{E}_\ell^k(\tau))| \\ &+\sup_{\tau\in\mathcal{U}}\sup_z\sup_\psi \lambda^{-\beta}\epsilon^{\beta-|\tau|}|(\Pi_z\tau)(\psi_z^\lambda)|,\end{aligned} \tag{3.12}$$

where $\beta=\frac{6}{5}$, $\mathcal{W}$ and $\mathcal{U}$ are defined before and after (3.4), and the supremum is taken over all $\psi\in\mathcal{B}$ that further annihilates affine functions. We also let $\mathscr{M}_0$ denote the set of admissible models such that $f_z(\mathscr{E}_\ell^k\tau)=0$ for all $\tau\in\mathcal{W}$. Note that for any positive ϵ and $\bar{\epsilon}$, $\mathscr{M}_\epsilon$ and $\mathscr{M}_{\bar{\epsilon}}$ consists of exactly the same set of models, but with norms at different scales. We compare two models $\Pi^\epsilon\in\mathscr{M}_\epsilon$ and $\Pi\in\mathscr{M}_0$ by

$$\|\!|\Pi^\epsilon;\Pi\|\!|_{\epsilon;0} := \|\!|\Pi^\epsilon;\Pi\|\!| + \|\Pi^\epsilon\|_\epsilon + \sup_{t\in[0,1]}\|\mathcal{R}^\epsilon\mathcal{I}(\Xi)(t,\cdot)-\mathcal{R}\mathcal{I}(\Xi)(t,\cdot)\|_{\mathcal{C}^\eta}, \tag{3.13}$$

and this will be used to characterise the convergence of our models to the limiting Φ_3^4 model in Theorem 5.1.

We also define the ϵ-weighted function space $\mathcal{C}_\epsilon^{\gamma,\eta}$ and modelled distribution space $\mathcal{D}_\epsilon^{\gamma,\eta}$. These are the same as [19, Definitions 3.3 and 3.5], but for the sake of completeness, we give the exact definitions again here. Let $\gamma\in(1,2)$, $\eta<0$ and $\epsilon>0$. The space $\mathcal{C}_\epsilon^{\gamma,\eta}$ consists of $\mathcal{C}^\gamma$ functions $f^{(\epsilon)}:\mathbf{R}^3\to\mathbf{R}$ with norm

$$\|f^{(\epsilon)}\|_{\gamma,\eta;\epsilon} = \|f^{(\epsilon)}\|_{\mathcal{C}^\eta} + \epsilon^{-\eta}\|f^{(\epsilon)}\|_\infty + \epsilon^{\gamma-\eta}\sup_{|x-y|<\epsilon}\frac{|Df^{(\epsilon)}(x)-Df^{(\epsilon)}(y)|}{|x-y|^{\gamma-1}}.$$

Furthermore, we set $\mathcal{C}_0^{\gamma,\eta}=\mathcal{C}^\eta$. The difference between two elements $f^{(\epsilon)}\in\mathcal{C}_\epsilon^{\gamma,\eta}$ and $f\in\mathcal{C}^\eta$ is defined by

$$\begin{aligned}\|f^{(\epsilon)};f\|_{\gamma,\eta;\epsilon} = &\|f^{(\epsilon)}-f\|_{\mathcal{C}^\eta} + \epsilon^{-\eta}\|f^{(\epsilon)}\|_\infty \\ &+\epsilon^{\gamma-\eta}\sup_{|x-y|<\epsilon}\frac{|Df^{(\epsilon)}(x)-Df^{(\epsilon)}(y)|}{|x-y|^{\gamma-1}}.\end{aligned}$$

Note that the latter two terms on the right hand side above involves only $f^{(\epsilon)}$, since these quantities may in general be ∞ for $f\in\mathcal{C}^\eta$. As in [19, Definition 3.5], for each admissible model (Π,Γ), the space $\mathcal{D}_\epsilon^{\gamma,\eta}$ consists of modelled distributions U with norm given by

$$\|U\|_{\gamma,\eta;\epsilon} = \|U\|_{\gamma,\eta} + \sup_z\sup_\ell\frac{|U(z)|_\ell}{\epsilon^{(\eta-\ell)\wedge 0}} + \sup_{(z,z')\in D_\epsilon}\sup_{\ell<\gamma}\frac{|U(z)-(F_z^{-1}\circ F_{z'})U(z')|_\ell}{|z-z'|^{\gamma-\ell}\epsilon^{\eta-\gamma}}. \tag{3.14}$$

Here, the supremum is taken over all points $z,z'\in D_\epsilon$ defined by

$$D_\epsilon = \left\{(z,z'):|z-z'|<\epsilon\wedge\sqrt{|t|\wedge|t'|}\right\},$$

where $z=(t,x)$, $z'=(t',x')$.

Let P denote the heat kernel, and K be its truncation as defined in Sect. 3.1. Then by [15, Section 5], there exists an operator $\mathcal{P}$ on $\mathcal{D}_\epsilon^{\gamma,\eta}$ such that

$$\mathcal{R}(\mathcal{P}f) = P * \mathcal{R}f.$$

Finally, we define the operator $\widehat{\mathcal{E}}^k$ (mapping $\mathcal{D}_\epsilon^{\gamma,\eta}$ to $\mathcal{D}_\epsilon^{\gamma',\eta'}$ for some other γ', η') by

$$(\widehat{\mathcal{E}}^k U)(z) = \mathcal{E}^k U(z) - \sum_\ell \frac{X^\ell}{\ell!} f_z(\mathscr{E}_\ell^k U(z)).$$

Since the only use of these spaces and the operator $\widehat{\mathcal{E}}_k$ in this article is in the statement of Theorem 3.3 below, we do not repeat their properties here, but refer to [17, Section 4] and [19, Section 3] for more details.

The following theorem is identical to [19, Theorem 3.12]. It gives the existence, uniqueness and convergence of abstract solutions to the lift of the Eq. (3.1) in the regularity structures.

Theorem 3.3 *Let $m \geq 1$, $\gamma \in (1, \frac{6}{5})$, $\eta \in (-\frac{2m+2}{4m+3}, -\frac{1}{2})$, and $\kappa > 0$ be sufficiently small. Let $\phi_0 \in \mathcal{C}_\epsilon^{\gamma,\eta}$, and consider the equation*

$$\Phi = \mathcal{P}\mathbf{1}_+\left(\Xi - \sum_{j=1}^m \lambda_j \widehat{\mathcal{E}}^{j-1}\Phi^{2j+1} - \lambda_0\Phi\right) + \widehat{P}\phi_0. \tag{3.15}$$

Then, for every sufficiently small ϵ and every model in $\mathscr{M}_\epsilon$, there exists $T > 0$ such that Equation (3.15) has a unique solution in $\mathcal{D}_\epsilon^{\gamma,\eta}$ up to time T. Moreover, T can be chosen uniformly over any fixed bounded set of initial data in $\mathcal{C}_\epsilon^{\gamma,\eta}$, any bounded set of models in $\mathscr{M}_\epsilon$, and all sufficiently small ϵ.

Let $\phi_0^{(\epsilon)}$ be a sequence of elements in $\mathcal{C}_\epsilon^{\gamma,\eta}$ such that $\|\phi_0^{(\epsilon)}; \phi_0\|_{\gamma,\eta;\epsilon} \to 0$ for some $\phi_0 \in \mathcal{C}^\eta$, $\Pi^\epsilon \in \mathscr{M}_\epsilon$, $\Pi \in \mathscr{M}_0$ be models such that $|\!|\!|\Pi^\epsilon; \Pi|\!|\!|_{\epsilon;0} \to 0$, and let $\lambda_j^{(\epsilon)} \to \lambda_j$ for each j. If $\Phi \in \mathcal{D}_0^{\gamma,\eta}$ solves the fixed point problem (3.15) up to time T with model Π, initial data ϕ_0 and coefficients λ_j, then for all small enough ϵ, there is a unique solution $\Phi^{(\epsilon)} \in \mathcal{D}_\epsilon^{\gamma,\eta}$ to (3.15) up to the same time T with initial data $\phi_0^{(\epsilon)}$, model Π^ϵ and coefficients $\lambda_j^{(\epsilon)}$'s, and we have

$$\lim_{\epsilon\to 0} \|\Phi^{(\epsilon)}; \Phi\|_{\gamma,\eta;\epsilon} = 0, \quad \lim_{\epsilon\to 0} \sup_{t\in[0,T]} \|(\mathcal{R}^{(\epsilon)}\Phi^{(\epsilon)})(t,\cdot) - (\mathcal{R}\Phi)(t,\cdot)\|_{\gamma,\eta;\epsilon} = 0,$$

where $\mathcal{R}^{(\epsilon)}$ and $\mathcal{R}$ are reconstructions operators to the models Π^ϵ and Π respectively.

3.3 The renormalisation group

If Φ solves the abstract Eq. (3.15) with the canonical model $\mathscr{L}_\epsilon(\zeta_\epsilon)$, then $u_\epsilon = \mathcal{R}^\epsilon\Phi$ solves the PDE

$$\partial_t u_\epsilon = \Delta u_\epsilon - \sum_{j=1}^m \lambda_j \epsilon^{j-1} u_\epsilon^{2j+1} - \lambda_0 u_\epsilon + \zeta_\epsilon.$$

The problem is that as $\epsilon \to 0$, the models $\mathscr{L}_\epsilon(\zeta_\epsilon)$ simply do not converge. To solve the problem one needs to perform suitable *renormalisations* for the equation.

The purpose of this section is to build a family $\mathfrak{R}$ of linear maps on $\mathcal{T}_{\text{ex}}$ so that when restricted to basis elements in $\mathcal{T}$, their action on admissible models yields the "renormalised models" (see 3.23 below) that will converge, and the reconstructed solution to the fixed point Eq. (3.15) satisfies some modified PDE.

For these purposes, the type of linear transformations $M \in \mathfrak{R}$ we consider will be the composition of two different maps M_0 and M^{Wick}, which acts on the original model (Π, f) in the following way:

$$(\Pi, f) \overset{M^{\text{Wick}}}{\longmapsto} (\Pi^{\text{Wick}}, f^{\text{Wick}}) \overset{M_0}{\longmapsto} (\Pi^M, f^M).$$

Here, M^{Wick} will act on the models as "Wick renormalisation", induced by a probability measure on $\mathbf{R}$, and M_0 has the interpretation as "mass renormalisation" in quantum field theory, and will be parametrised by finitely many parameters.

We start with the description of the Wick renormalisation map. For any probability measure μ on $\mathbf{R}$ with all moments finite, we define the linear map M_μ^{Wick} by setting

$$M_\mu^{\text{Wick}} \Xi = \Xi, \qquad M_\mu^{\text{Wick}} X^k = X^k$$

and

$$M_\mu^{\text{Wick}} \Psi^n = W_{n,\mu}(\Psi), \qquad \Psi = \mathcal{I}(\Xi)$$

mapping Ψ^n into the Wick polynomial induced by μ. We furthermore require that M_μ^{Wick} commutes with the abstract integration maps $\mathcal{I}$ and $\mathcal{E}^k$ as well as the multiplication by X^k, and extend it to the whole of $\mathcal{T}_{\text{ex}}$ by

$$M_\mu^{\text{Wick}}(\tau \mathcal{I}(\bar{\tau})) = M_\mu^{\text{Wick}}(\tau) \mathcal{I}(M_\mu^{\text{Wick}} \bar{\tau}). \tag{3.16}$$

For the map M_μ^{Wick} defined above, it follows from [15, Section 8] or [17, (5.7)] that there is a unique pair of linear maps

$$\Delta^{\text{Wick}} : \mathcal{T}_{\text{ex}} \to \mathcal{T}_{\text{ex}} \otimes \mathcal{T}_+, \qquad \widehat{M}^{\text{Wick}} : \mathcal{T}_+ \to \mathcal{T}_+$$

such that

$$\begin{aligned} \widehat{M}^{\text{Wick}} \mathscr{J}_\ell &= \mathcal{M}(\mathscr{J}_\ell \otimes \text{id}) \Delta^{\text{Wick}}, \\ \widehat{M}^{\text{Wick}} \mathscr{E}_\ell^k &= \mathcal{M}(\mathscr{E}_\ell^k \otimes \text{id}) \Delta^{\text{Wick}}, \\ (\text{id} \otimes \mathcal{M})(\Delta \otimes \text{id}) \Delta^{\text{Wick}} &= (M_\mu^{\text{Wick}} \otimes \widehat{M}^{\text{Wick}}) \Delta, \\ \widehat{M}^{\text{Wick}}(\tau_1 \tau_2) &= (\widehat{M}^{\text{Wick}} \tau_1)(\widehat{M}^{\text{Wick}} \tau_2), \quad \widehat{M}^{\text{Wick}} X^k = X^k, \end{aligned} \tag{3.17}$$

where $\mathcal{M} : \mathcal{T}_+ \to \mathcal{T}_+$ denotes the multiplication in Hopf algebra. Given an admissible model (Π, f), we define its *Wick renormalised model* $(\Pi^{\text{Wick}}, f^{\text{Wick}})$ by

$$\Pi_z^{\text{Wick}} \tau = (\Pi_z \otimes f_z) \Delta^{\text{Wick}} \tau, \qquad f_z^{\text{Wick}}(\sigma) = f_z(\widehat{M}^{\text{Wick}} \sigma). \tag{3.18}$$

The following proposition ensures that the new pair $(\Pi^{\text{Wick}}, f^{\text{Wick}})$ is indeed an admissible model.

Proposition 3.4 *The unique map Δ^{Wick} defined in (3.17) has the following upper-triangular property: for every $\tau \in \mathcal{T}$, one has*

$$\Delta^{\text{Wick}} \tau = \tau \otimes \mathbf{1} + \sum \tau^{(1)} \otimes \tau^{(2)}, \tag{3.19}$$

where each term in the sum satisfies $|\tau^{(1)}| > |\tau|$. As a consequence, the new pair $(\Pi^{\text{Wick}}, f^{\text{Wick}})$ is an admissible model.

Proof If Δ^{Wick} satisfies the upper-triangular property, then one can follow the same argument in [15,17] to verify that the pair $(\Pi^{\text{Wick}}, f^{\text{Wick}})$ defined in (3.18) indeed satisfies all the requirements for an admissible model. It then remains to verify (3.19) for every basis vector τ.

The case for $\tau = \Xi$ and X^k is trivial. One can also verify by hand that $\Delta^{\text{Wick}} \Psi^k = M_\mu^{\text{Wick}} \Psi^k \otimes 1$, so the property also follows for $\tau = \Psi^k$. Now, since M_μ^{Wick} commutes with $\mathcal{I}$ and $\mathcal{E}^k$ as well as satisfies the rule (3.16), it has exactly the same algebraic structures as in [17, Section 5.2] and [19, Section 2.4], so it follows from the same line of argument that the upper-triangular property is preserved under the operations

$$\tau \mapsto \mathcal{I}(\tau), \qquad \tau \mapsto \mathcal{E}^k(\tau), \qquad \tau \mapsto X^k \tau.$$

It then extends to all basis vectors, and the claim follows. □

We now turn to the description of the "mass renormalisation" map M_0. Define the generators $L_{k,\ell}$ for k, ℓ both even or both odd by

$$\begin{aligned} L_{2k,2\ell} : \quad & \mathcal{E}^{k-1}\Big(\Psi^{2k} \mathcal{I}(\mathcal{E}^{\ell-1} \Psi^{2\ell})\Big) \mapsto \mathbf{1}, \\ & \mathcal{E}^{k-1}\Big(\Psi^{2k} \mathcal{I}(\mathcal{E}^{\ell-1} \Psi^{2\ell+1})\Big) \mapsto (2\ell+1) \cdot \Psi, \\ L_{2k-1,2\ell+1} : \quad & \mathcal{E}^{k-1}\Big(\Psi^{2k-1} \mathcal{I}(\mathcal{E}^{\ell-1} \Psi^{2\ell+1})\Big) \mapsto \mathbf{1}, \\ & \mathcal{E}^{k-1}\Big(\Psi^{2k} \mathcal{I}(\mathcal{E}^{\ell-1} \Psi^{2\ell+1})\Big) \mapsto (2k) \cdot \Psi. \end{aligned} \tag{3.20}$$

Given these generators, we then consider the map M_0 of the form

$$M_0 = \exp\Big(- \sum_{k,\ell \geq 1} C_{k,\ell} L_{k,\ell} \Big). \tag{3.21}$$

M_0 is then parametrised by the set of constants $\{C_{k,\ell}\}$. Here $C_{k,\ell} = 0$ is $k+\ell$ is odd. For any admissible model $(\bar{\Pi}, \bar{f})$, we define the action of M_0 by

$$\bar{\Pi}_z^{M_0}\tau := \bar{\Pi}_z M_0 \tau, \qquad \bar{f}_z^{M_0}(\tau) := \bar{f}_z(\tau). \tag{3.22}$$

Since the only basis elements in the regularity structures that M_0 has a non-trivial effect on are of the form $\mathcal{E}^a \Psi^k \big(\mathcal{I}\big(\mathcal{E}^b \Psi^\ell\big)\big)$, it is immediate to see that M_0 commutes with the elements in the structure group in the sense that $M_0 \Gamma \tau = \Gamma M_0 \tau$ for any $\Gamma \in \mathcal{G}$ and $\tau \in \mathcal{T}$. As a consequence, one could easily deduce that the model $(\bar{\Pi}^{M_0}, \bar{f}^{M_0})$ also belongs to $\mathscr{M}$.

Combining (3.18) and (3.22), we can then define the new model (Π^M, f^M) under the action of $M = (M_\mu^{\text{Wick}}, M_0)$ by

$$\Pi_z^M \tau = (\Pi_z \otimes f_z)\Delta^{\text{Wick}}(M_0 \tau), \qquad f^M(\sigma) = f_z(\widehat{M}^{\text{Wick}}\sigma). \tag{3.23}$$

It follows from Proposition 3.4 and the arguments right after (3.22) that (Π^M, f^M) is admissible as long as (Π, f) is. With the above definitions of the renormalisation maps, we then have the following theorem.

Theorem 3.5 *Let $\phi_0 \in \mathcal{C}^1, \epsilon \geq 0$ and ζ be a smooth space-time function. Let $(\Pi, f) = \mathscr{L}_\epsilon(\zeta)$ be the canonical model, $M = (M_\mu^{\text{Wick}}, M_0)$ be the renormalisation maps defined as above, and $(\Pi^M, f^M) = M\mathscr{L}_\epsilon(\zeta)$ be the renormalised model as in (3.23). If $\Phi \in \mathcal{D}_\epsilon^{\gamma,\eta}$ is the solution to the fixed point problem (3.15) for the renormalised model $M\mathscr{L}_\epsilon(\zeta)$, and $\mathcal{R}^M$ is reconstruction map associated with the renormalised model, then the function $u = \mathcal{R}^M \Phi$ solves the classical PDE*

$$\partial_t u = \Delta u - \sum_{j=1}^m \lambda_j \epsilon^{j-1} W_{2j+1,\mu}(u) - \lambda_0 u - Cu + \zeta$$

with initial data ϕ_0, where

$$C = \sum_{k,\ell=1}^m \lambda_k \lambda_\ell \bigg((2k+1)(2k) C_{2k-1,2\ell+1} + (2k+1)(2\ell+1) C_{2k,2\ell} \bigg). \tag{3.24}$$

Proof The key of the proof is to note that $W_{n,\mu}$ defined in (2.8) is an Appell sequence, namely

$$W'_{n,\mu}(x) = n W_{n-1,\mu}(x).$$

as discussed below (2.8). A characteristic property of Appell sequences is the following identity (see for instance [2], the equation between (14) and Remark 3 therein with X distributed according to the measure μ here and $Y = 0$)

$$W_{n,\mu}(x+y) = \sum_{k=0}^n \binom{n}{k} W_{k,\mu}(x) \cdot y^{n-k} \tag{3.25}$$

as in the case of Hermite polynomial. With (3.25) at hand, the rest of the proof follows exactly the same way as [17, Section 5.3] and [19, Section 2.4] since the renormalisation map M defined here has the same structure as in those two references, and the only property they used to derive the form of the renormalised equation is the identity (3.25).

In addition, the precise expression (3.24) of the constant C can be seen as follows. We first note that this constant results from the renormalisation map M_0 defined in (3.20) and (3.21). The abstract solution Φ to (3.15) must have the form

$$\Phi = \Psi + \varphi \cdot \mathbf{1} - \mathcal{I}(P) - \varphi \cdot \mathcal{I}(P') + \varphi' \cdot X,$$

for some space-time functions φ and φ', where $\Psi = \mathcal{I}(\Xi)$ as in (3.5) above, and

$$P = \sum_{j=1}^{m} \lambda_j \mathcal{E}^{j-1} \Psi^{2j+1}, \quad P' = \sum_{j=1}^{m} (2j+1) \lambda_j \mathcal{E}^{j-1} \Psi^{2j}.$$

With this form of Φ, we see that the terms relevant to M_0 in the expansion of $\sum_j \lambda_j \mathcal{E}^{j-1} \Phi^{2j+1}$ are

$$- P'\mathcal{I}(P) - \varphi \cdot P'\mathcal{I}(P') - \varphi \cdot P''\mathcal{I}(P), \tag{3.26}$$

where in addition to P, P' above, we have

$$P'' = \sum_{j=1}^{m} (2j+1)(2j) \lambda_j \mathcal{E}^{j-1} \Psi^{2j-1}.$$

Substituting the expressions of P, P', P'' into (3.26) and employing the definition of M_0 as in (3.20) and (3.21), we then get the precise form of the constant C as in (3.24). $\square$

We have now shown that the group of transformations we build does map $\mathscr{M}$ to $\mathscr{M}$, and have derived the form of the modified PDEs under renormalised models. The rest of the article is devoted to the proof of the convergence of these renormalised models (with suitably chosen renormalisation constants) and identification of the limit of the modified equations.

4 Bounds on the renormalised models

The main goal of this section to prove the bounds for high moments of the objects $|(\widehat{\Pi}_z^\epsilon \tau)(\varphi_z^\lambda)|$, where $\widehat{\Pi}^\epsilon$ is the renormalised model as introduced in Sect. 3.3 with suitable measure μ_ϵ and constants $C_{k,\ell}$.

Let ζ be the random field satisfying Assumption 1.3, and $\zeta_\epsilon(t,x) = \epsilon^{-\frac{5}{2}} \zeta(t/\epsilon^2, x/\epsilon)$. Also, fix a space-time mollifier ρ and for any $\bar{\epsilon} > 0$, set $\rho_{\bar{\epsilon}}(t,x) = \bar{\epsilon}^{-5} \rho(t/\bar{\epsilon}^2, x/\bar{\epsilon})$. Let

$$\zeta_{\epsilon,\bar{\epsilon}} := \zeta_\epsilon * \rho_{\bar{\epsilon}}. \tag{4.1}$$

We can thus write $K * \zeta_{\epsilon,\bar{\epsilon}} = (K * \rho_{\bar{\epsilon}}) * \zeta_\epsilon$, and the kernel $K_{\bar{\epsilon}} := K * \rho_{\bar{\epsilon}}$ approximates K with the bounds

$$|K_{\bar{\epsilon}}(z) - K(z)| \lesssim \bar{\epsilon}^\delta |z|^{-3-\delta}, \qquad |DK_{\bar{\epsilon}}(z) - DK(z)| \lesssim \bar{\epsilon}^\delta |z|^{-4-\delta} \tag{4.2}$$

for all sufficiently small δ, uniformly over $\bar{\epsilon} < 1$ and $|z| < 1$. Later, we will use this bound to compare the difference between models built from ζ_ϵ and $\zeta_{\epsilon,\bar{\epsilon}}$.

Let μ_ϵ and $\mu_{\epsilon,\bar{\epsilon}}$ denote the distributions of stationary solutions Ψ_ϵ, $\Psi_{\epsilon,\bar{\epsilon}}$ to the equations

$$\partial_t \Psi_\epsilon = \Delta \Psi_\epsilon + \zeta_\epsilon, \qquad \partial_t \Psi_{\epsilon,\bar{\epsilon}} = \Delta \Psi_{\epsilon,\bar{\epsilon}} + \zeta_{\epsilon,\bar{\epsilon}}.$$

Let $M_\epsilon = (M^{\text{Wick}}_{\mu_\epsilon}, M_0^{(\epsilon)})$ and $M_{\epsilon,\bar{\epsilon}} = (M^{\text{Wick}}_{\mu_{\epsilon,\bar{\epsilon}}}, M_0^{(\epsilon,\bar{\epsilon})})$ be the renormalisation maps built in Sect. 3.3, where the constants $C^{(\epsilon)}_{k,\ell}$ and $C^{(\epsilon,\bar{\epsilon})}_{k,\ell}$'s are chosen as in Sect. 4.2 below. The main theorem of this section is then the following.

Theorem 4.1 *Let $\mathscr{L}_\epsilon(\cdot)$ be the canonical lift of the corresponding input noise to the space $\mathscr{M}_\epsilon$, and*

$$\mathfrak{M}_\epsilon = (\widehat{\Pi}^\epsilon, \widehat{f}^\epsilon) = M_\epsilon \mathscr{L}_\epsilon(\zeta_\epsilon) \quad \text{and} \quad \mathfrak{M}_{\epsilon,\bar{\epsilon}} = (\widehat{\Pi}^{\epsilon,\bar{\epsilon}}, \widehat{f}^{\epsilon,\bar{\epsilon}}) = M_{\epsilon,\bar{\epsilon}} \mathscr{L}_\epsilon(\zeta_{\epsilon,\bar{\epsilon}}) ,$$

where the constants $C^{(\epsilon)}_{k,\ell}$ and $C^{(\epsilon,\bar{\epsilon})}_{k,\ell}$ in defining the renormalisation groups are given in Sect. 4.2 below. Then, there exists $\theta > 0$ such that for every $\tau \in \mathcal{W}$ with $|\tau| < 0$ and every $p > 1$, we have

$$\mathbf{E}\left|\left(\widehat{\Pi}^\epsilon_z \tau\right)\left(\varphi^\lambda_z\right)\right|^p \lesssim \lambda^{p(|\tau|+\theta)}, \qquad \mathbf{E}\left|\left(\widehat{\Pi}^\epsilon_z \tau - \widehat{\Pi}^{\epsilon,\bar{\epsilon}}_z \tau\right)\left(\varphi^\lambda_z\right)\right|^p \lesssim \bar{\epsilon}^\theta \lambda^{p(|\tau|+\theta)}, \tag{4.3}$$

where both bounds hold uniformly over all $\epsilon, \bar{\epsilon} \in (0,1)$, all test functions $\varphi \in \mathcal{B}$ and all space-time points $z \in \mathbf{R}^4$.

We will mainly focus on the first bound in (4.3) below. Indeed, once the first bound in (4.3) is obtained for a basis vector τ, the second bound for that τ then follows in exactly the same way by considering the difference of the kernels $K_{\bar{\epsilon}} - K$ and applying (4.2), see [18] (Proof of Theorem 4.1 in the end of Section 4 therein). Also, since we are in a translation invariant setting, we will set $z = 0$ without loss of generality.

Remark 4.2 Ideally, one would like to find a limiting model $\widehat{\Pi}$ and prove a bound of the type

$$\mathbf{E}|\left(\widehat{\Pi}^\epsilon_z \tau - \widehat{\Pi}_z \tau\right)\left(\varphi^\lambda_z\right)|^p \lesssim \epsilon^\theta \lambda^{p(|\tau|+\theta)}. \tag{4.4}$$

for some positive θ. In fact, the natural candidate for such a limiting model is the one in $\mathscr{M}_0$ whose action on basis vectors without appearance of $\mathcal{E}$'s coincides with the Φ^4_3 model. In fact, we will actually prove this bound below for basis vectors that contain at least one appearance of $\mathcal{E}$, in which case we have $\widehat{\Pi}_z \tau = 0$. We expect that the bound (4.4) still holds for standard Φ^4_3 basis elements (those without appearance of $\mathcal{E}$'s), but the proof would be much more involved due to the non-Gaussian noise. Later when we identify the Φ^4_3 model as the limit, we make use of the convergence result available in the Gaussian case [19] as well as the diagonal argument used [18] to circumvent the bound (4.4) in non-Gaussian case.

4.1 Graphic notation

Roughly speaking the random variables of the type $(\Pi^\epsilon_0\tau)(\varphi^\lambda_0)$ or the ones for the renormalised models $(\widehat{\Pi}^\epsilon_0\tau)(\varphi^\lambda_0)$ in Theorem 4.1 are all integrations of convolutions or products of noises ζ_ϵ and kernels K, φ^λ and cumulant functions. For example, for $\tau = \mathcal{E}\Psi^4$, we have

$$\left(\Pi^\epsilon_0\tau\right)\left(\varphi^\lambda_0\right) = \epsilon \int (K * \zeta_\epsilon)^4\,(z)\varphi^\lambda(z)dz$$
$$= \epsilon \int \left(\prod_{i=1}^4 K(z-z_i)\zeta_\epsilon(z_i)\right)\varphi^\lambda(z)\,dz\,dz_1\cdots dz_4.$$

We will apply Definition 2.3 to rewrite a product of ζ_ϵ's (for the above example of τ, this product is $\prod_{i=1}^4 \zeta_\epsilon(z_i)$) as a sum of Wick products. As we consider the other elements τ in our regularity structure in the following, these expressions of integrations or convolutions become rather complicated, so as in [17, Section 6], [19, Section 4.2], [18, Section 3], we introduce graphic notations to represent these integrals.

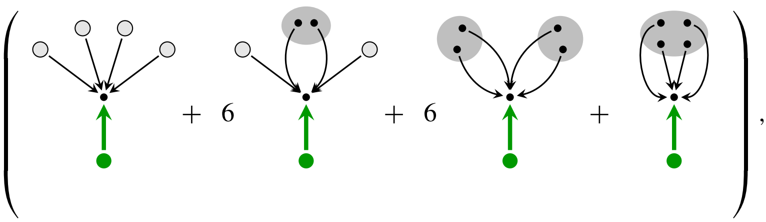
We denote by node • a space-time variable in $\mathbf{R}\times\mathbf{R}^3$ to be integrated out. The special green node ● denotes the origin 0. A bold green arrow ⟶ represents φ^λ, i.e. a generic test function rescaled by λ. Each plain arrow ⟶ represents the kernel $K(z'-z)$, and a barred arrow ⟶ represents a factor $K(z'-z)-K(-z)$, where z and z' are starting and end points of the arrow. As for cumulants, we follow essentially the same notation as in [18]: a gray ellipse with p points inside  ($p=4$ here) represents the cumulant $\mathfrak{C}^{(\epsilon)}_p(z_1,\ldots,z_p)$.

To represent the Wick products 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 $\zeta_\epsilon(x)$. Furthermore, if more than one such vertex appear, then the corresponding product of ζ_ϵ is always understood as a Wick product $:\!\zeta_\epsilon(x_1)\cdots\zeta_\epsilon(x_n)\!:$, where the x_i are the integration variables represented by *all* of the special vertices ○ appearing in the graph.

With these notations, for the canonical model $(\Pi^\epsilon, f^\epsilon) = \mathscr{L}_\epsilon(\zeta_\epsilon)$, we can apply (2.5) and write each $(\widehat{\Pi}^\epsilon_0\tau)(\varphi^\lambda_0)$ as a sum of terms ("non-Gaussian chaos") each represented by a graph. For example, for $\tau = \mathcal{E}\Psi^4$, we have

$(\Pi^\epsilon_0\mathcal{E}\Psi^4)(\varphi^\lambda_0)$

$= \epsilon$



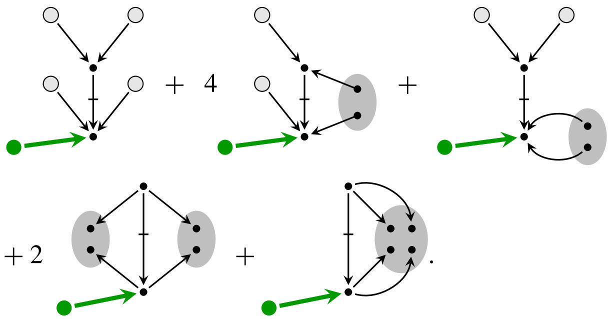
where the first and second terms above have the full expressions

$$I_1 = \int \varphi_0^\lambda(z) \prod_{j=1}^4 K(z - x_j) : \prod_{j=1}^4 \zeta_\epsilon(x_j) \colon dxdz,$$

$$I_2 = 6 \int \varphi_0^\lambda(z) \prod_{j=1}^4 K(z - x_j) \cdot \mathfrak{C}_2^{(\epsilon)}(x_1, x_2) :\zeta_\epsilon(x_3)\zeta_\epsilon(x_4)\colon dxdz,$$

and we used $x = (x_1, \dots, x_4)$. If $\widehat{\Pi}^\epsilon$ is the renormalised model, then $(\widehat{\Pi}_0^\epsilon \tau)(\varphi_0^\lambda)$ consists only the first term above. Similarly, for $\tau = \Psi^2 \mathcal{I}(\Psi^2)$, we have

$$(\Pi_0^\epsilon \Psi^2 \mathcal{I}(\Psi^2))(\varphi_0^\lambda) =$$



The reason that there is no contraction for odd number of vertices is because the noise is symmetric so all its odd cumulants vanish, and there is no contraction between two top vertices because the kernel K chosen in Sect. 3.1 annihilates constants.

In order to bound the moments of the random variables $(\widehat{\Pi}_0^\epsilon \tau)(\varphi_0^\lambda)$, we bound the moments of the term represented by *each* graph. Assumption 4.4 and Theorem 4.6 below will allow us to conclude the desired moment bounds by simply verifying certain graphical conditions. To state these conditions we introduce the following terminologies.

Let $(H, \mathcal{E})$ be a graph where H is the set of vertices and $\mathcal{E}$ (with a slight abuse of notation as this letter also stands for the operator as multiplying by ϵ) is the collection of edges. Each such graph consists of a set H_{ex} of external vertices (i.e. noise vertices that have not been integrated out) and a set H_{in} of internal vertices and the origin 0, namely $H = \{0\} \cup H_{ex} \cup H_{in}$. There is a distinguished vertex $v_\star$ connected with 0 by a line representing the test function φ^λ. Write

$$H_0 = H \setminus \{0\} \qquad \text{and} \qquad H_\star = \{0, v_\star\}.$$

For a sub-graph $\bar{H} \subset H$, we write $\bar{H}_{in} = \bar{H} \cap H_{in}$ and $\bar{H}_{ex} = \bar{H} \cap H_{ex}$.

The set of edges is decomposed as $\mathcal{E} = \mathcal{E}_2 \cup \mathcal{E}_h$ where $\mathcal{E}_2$ is a set of usual directed edges, which represent integration kernels, and $\mathcal{E}_h$ is a set of hyper-edges (i.e. edges consisting of more than two vertices), which represent cumulants of order higher than two. Each edge $e = (e_-, e_+) \in \mathcal{E}_2$ comes with a pair of numbers $(a_e, r_e) \in \mathbf{R}^+ \times \mathbf{R}$. The number $a_e > 0$ measures the singularity of the kernel J_e associated to the edge

e. More precisely, for each edge e with a_e, the kernel J_e associated to it satisfies the bound

$$|D^k J_e(z)| \lesssim |z|^{-a_e-|k|}.$$

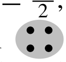
If an edge $e \in \mathcal{E}_2$ is connected to an external vertex (noise), then the kernel J_e has the bound

$$|D^k J_e(z)| \lesssim (|z| + \epsilon)^{-a_e-|k|},$$

reflecting the fact that the input of the noise is at scale ϵ. For every $\beta > 0$, we associate the kernel J_e with norm $\|J_e\|_{a_e;\beta}$ by

$$\|J_e\|_{a_e;\beta} := \sup_{|z|\leq 1, |k|\leq\beta} |z|^{a_e+|k|} |D^k J_e(z)|.$$

This quantity is always finite by assumption on the kernel J_e. The number r_e, on the other hand, gives the corresponding renormalisations needed for the edge: if $r_e > 0$, then one needs to subtract Taylor expansions up to order r_e; if $r_e < 0$, then one needs to define a renormalised version of the kernel by subtracting Taylor polynomials of the test function at the origin up to degree $|r_e|$. The orientation of an edge matters only when $r_e > 0$. Precise definitions of the positive/negative renormalisations could be found in [17, Section 8] and [19, Section 4].

For each hyper-edge $e \in \mathcal{E}_h$, we set $r_e = 0$, and associate it with the degree $a_e = \frac{5n}{2}$, where n is the number of vertices in this hyper-edge. We use the same notation  to represent a hyper-edge connecting n noise vertices, and it represents the kernel $\mathfrak{C}_\epsilon(z_1, \ldots, z_n)$. Note that this kernel has the correct scaling behaviour corresponding to the degree $\frac{5n}{2}$, and satisfies (2.4).

Remark 4.3 The degree of a hyper-edge indeed corresponds to the correct behaviour of cumulants of the rescaled field ζ_ϵ. This is guaranteed by the scale set in (1.11) as well as (2.4). For normal edges $e \in \mathcal{E}_2$, we assign their initial degree to be $a_e = 3$. But many of these (for those connected to external noise vertices) will be reduced later by multiplication of powers of ϵ's. All edges represented by the plain arrow $\longrightarrow$ have $r_e = 0$, but the pair corresponding to the barred arrow $\longmapsto$ is always $(a_e, r_e) = (3, 1)$.

For a sub-graph $\bar{H} \subset H$, we define the sets $\mathcal{E}^\uparrow(\bar{H})$, $\mathcal{E}^\downarrow(\bar{H})$, $\mathcal{E}_0(\bar{H})$ and $\mathcal{E}(\bar{H})$ in the same way as [17,19] by

$$\begin{aligned}
\mathcal{E}^\uparrow(\bar{H}) &= \{e \in \mathcal{E} : e \cap \bar{H} = e_-, r_e > 0\};\\
\mathcal{E}^\downarrow(\bar{H}) &= \{e \in \mathcal{E} : e \cap \bar{H} = e_+, r_e > 0\};\\
\mathcal{E}_0(\bar{H}) &= \{e \in \mathcal{E} : e \cap \bar{H} = e\};\\
\mathcal{E}(\bar{H}) &= \{e \in \mathcal{E} : e \cap \bar{H} \neq \emptyset\}.
\end{aligned}$$

We see that $\mathcal{E}^{\uparrow}(\bar{H})$ and $\mathcal{E}^{\downarrow}(\bar{H})$ are the set of outgoing and incoming edges of $\bar{H}$ with the further constraints $r_e > 0$ and $r_e < 0$ respectively. In particular, $\mathcal{E}^{\uparrow}(\bar{H})$ and $\mathcal{E}^{\downarrow}(\bar{H})$ are subsets of $\mathcal{E}_2$. On the other hand, a hyper-edge e can be in $\mathcal{E}_0(\bar{H})$ only if *all* the vertices of e are in $\bar{H}$. We now make a few assumptions on our graphs that ensure the high moments of the corresponding object have the correct scaling behaviour.

Assumption 4.4 [9, Assumption 3.15] The labelled graph $(H, \mathcal{E})$ satisfies the following properties.

1. For every edge $e \in \mathcal{E}_2$, one has $a_e + (r_e \wedge 0) < 5$.
2. For every subset $\bar{H} \subset H_0$ of cardinality at least 3, one has

$$\sum_{e \in \mathcal{E}_0(\bar{H})} a_e < 5\Big(|\bar{H}_{in}| + \frac{1}{2}(|\bar{H}_{ex}| - 1 - \mathbf{1}_{\bar{H}_{ex}=\emptyset})\Big). \tag{4.5}$$

3. For every subset $\bar{H} \subset H$ containing 0 and of cardinality at least 2, one has

$$\sum_{e \in \mathcal{E}_0(\bar{H})} a_e + \sum_{e \in \mathcal{E}^{\uparrow}(\bar{H})} (a_e + r_e - 1) - \sum_{e \in \mathcal{E}^{\downarrow}(\bar{H})} r_e < 5\Big(|\bar{H}_{in}| + \frac{1}{2}|\bar{H}_{ex}|\Big). \tag{4.6}$$

4. For every non-empty subset $\bar{H} \subset H_\star$, one has the bounds

$$\sum_{e \in \mathcal{E}(\bar{H}) \setminus \mathcal{E}^{\downarrow}(\bar{H})} a_e + \sum_{e \in \mathcal{E}^{\uparrow}(\bar{H})} r_e - \sum_{e \in \mathcal{E}^{\downarrow}(\bar{H})} (r_e - 1) > 5\Big(|\bar{H}_{in}| + \frac{1}{2}|\bar{H}_{ex}|\Big). \tag{4.7}$$

Note that the number 5 in the above conditions is the parabolically scaled space-time dimension.

Remark 4.5 The p-th moment of the object represented by a graph can be expressed as a sum of finitely many terms, each obtained by Wick-contracting p copies of the graph (that is, excluding self-contraction). We refer to [18] for more details on the Wick-contraction.

The following theorem from [9] gives a sufficient condition on the graph H for its p-th moment to have the correct scaling behaviour.

Theorem 4.6 *Suppose that a graph $(H, \mathcal{E})$ satisfies Assumption 4.4, and the kernels $\mathfrak{C}_\epsilon(z_1, \dots, z_n)$ represented by the hyper-edges are the cumulants of the rescaled noise satisfying Assumption 1.3. If $I^H_{p,\lambda}$ denotes the p-th moment of the (random) object represented by H with test function scaled at λ, then there exists $\beta > 0$ depending on the structure of the graph only such that*

$$I^H_{p,\lambda} \lesssim \lambda^{\alpha p} \cdot \prod_e \|J_e\|^p_{a_e;\beta}, \quad \alpha = 5|H_{in} \setminus H_\star| + \frac{5}{2}|H_{ex}| - \sum_e a_e,$$

where the proportionality constant depends on the structure of graph and cumulants of the unscaled noise $\zeta(\cdot)$ only.

4.2 Values of renormalisation constants

With the graphic notations, we now give explicit expressions to the constants $C_{k,\ell}^{(\epsilon)}$'s that appear in the renormalisation map $M_0^{(\epsilon)}$ for (k, ℓ) such that $k + \ell$ is even. Given a tuple of two integers (k, ℓ), we say that π is a *pairing* of (k, ℓ) if π is a collection of n tuples (each consisting of two integers):

$$\pi = \big\{(k_1, \ell_1), \dots, (k_n, \ell_n)\big\} \tag{4.8}$$

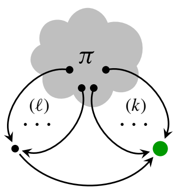
for some n such that $k_j, \ell_j \geq 1$ for all j and

$$\sum_j k_j = k, \qquad \sum_j \ell_j = \ell.$$

We write $|\pi| = n$. Let $\mathcal{P}(k, \ell)$ be the set of all the pairings of (k, ℓ). For example, we have

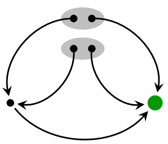
$$\mathcal{P}(3,3) = \Big\{\big\{(1,1),(1,1),(1,1)\big\}, \big\{(1,1),(2,2)\big\}, \big\{(1,2),(2,1)\big\}, \big\{(3),(3)\big\}\Big\}.$$

Given a pairing $\pi \in \mathcal{P}(k, \ell)$, we define

$$C_{k,\ell,\pi}^{(\epsilon)} = \epsilon^{\frac{k+\ell-4}{2}}$$


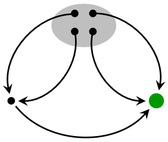
In the above picture the cloudy area represents a product of $|\pi|$ cumulants specified by π: if π is given by (4.8) then the i-th cumulant ($1 \leq i \leq n$) in the product is a cumulant function of order $k_i + \ell_i$ with k_i variables selected from the k ones on the right and ℓ_i variables selected from the ℓ ones on the left. Note that there is a slight abuse of notation here as π itself depends on k and ℓ, but we choose to keep this notation for simplicity. Intuitively, a pairing π is a way to "contract" the k vertices on the right and the ℓ ones on the left.

As for the values of these constants, it is easy to see that $C_{2,2,\pi}^{(\epsilon)}$ diverges logarithmically when $\pi = \{(1,1),(1,1)\}$ (see [15, Remark 1.17 and Section 10.5]), while all other $C_{k,\ell,\pi}^{(\epsilon)}$'s converge to a finite limit. In fact, we have that if $\pi = \{(1,1),(1,1)\}$ then

$$C_{2,2,\pi}^{(\epsilon)} =$$

$$= C_{\log} \cdot |\log \epsilon| + \mathcal{O}(1), \tag{4.9}$$

for some universal constant $C_{\log}$[2], while for π' being the single contraction of all four points together, we have

$$C^{(\epsilon)}_{2,2,\pi'} \quad = $$



$$= \quad C_{2,2,\pi'} + \mathcal{O}(\epsilon). \tag{4.10}$$

For fixed $k, \ell \geq 2$ and $\pi \in \mathcal{P}(k, \ell)$, we let $\pi!$ denote all the ways to contract (k, ℓ) vertices according to the pairing π. For example, for $(k, \ell) = (3, 3)$, we have

$$\begin{aligned} \pi! &= 6, \qquad \pi = \{(1,1),(1,1),(1,1)\};\\ \pi! &= 9, \qquad \pi = \{(1,1),(2,2)\};\\ \pi! &= 9, \qquad \pi = \{(1,2),(2,1)\};\\ \pi! &= 1, \qquad \pi = \{(3),(3)\}. \end{aligned}$$

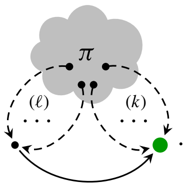
With this notation, we let

$$C^{(\epsilon)}_{k,\ell} = \sum_{\pi \in \mathcal{P}(k,\ell)} \pi! \cdot C^{(\epsilon)}_{k,\ell,\pi}, \tag{4.11}$$

where the sum is taken over all pairings for (k, ℓ). The precise values of these constants do not matter, so we do not give explicit formulae of values of $\pi!$ for general partitions π.

As for $C^{(\epsilon,\bar{\epsilon})}_{k,\ell,\pi}$'s, they are almost the same as the $C^{(\epsilon)}_{k,\ell,\pi}$'s except that each plain arrow $\longrightarrow$ connected to a noise is replaced by a dashed arrow $\dashrightarrow$, representing a factor $K * \rho_{\bar{\epsilon}}$. This is because we have $K * \zeta_{\epsilon,\bar{\epsilon}} = K * \rho_{\bar{\epsilon}} * \zeta_\epsilon$. Since we have put the mollifier $\rho_{\bar{\epsilon}}$ into the kernel K, the notation  still represents the cumulants of the field ζ_ϵ. For example, we have

$$C^{(\epsilon,\bar{\epsilon})}_{k,\ell,\pi} \quad = \quad \epsilon^{\frac{k+\ell-4}{2}}$$



The sum $C^{(\epsilon,\bar{\epsilon})}_{k,\ell}$'s are defined the same way as in (4.11).

---

[2] In fact, for $\pi = \{(1,1),(1,1)\}$, we have the expression

$$C^{(\epsilon)}_{2,2,\pi} = \int \big(\mathbf{E}(\Psi_\epsilon(z)\Psi_\epsilon(z'))\big)^2 K(z - z')dz',$$

where the integration is taken in five space-time dimensions. Since $\mathbf{E}(\Psi_\epsilon \Psi_\epsilon(z')) \sim (|z - z'| + \epsilon)^{-1}$, the integrand has a singularity of order -5, hence is logarithmically divergent. With a little more effort, one can actually show that the leading constant multiplying the $\log \epsilon$ factor is independent of the actual correlation function of the noise. But this is not important for our purpose, so we omit the details here.

4.3 First order renormalisation bounds

We are now ready to prove Theorem 4.1 for all basis vectors $\tau \in \mathcal{W}$ with $|\tau| < 0$, as listed in (3.5). We first prove the bound (4.3) for $\tau = \mathcal{E}^{k-1}\Psi^{2k+1-n}$ where $n \leq 3$. For the canonical model Π^ϵ, then we have

$$\Pi_z^\epsilon \tau = \epsilon^{k-1}(\Pi_z^\epsilon \Psi)^{2k+1-n}.$$

Let $\Psi_\epsilon = P * \zeta_\epsilon$ be the stationary solution to the linear heat equation driven by ζ_ϵ, and μ_ϵ is the distribution of Ψ_ϵ, we have

$$W_{k,\mu_\epsilon}(\Psi_\epsilon(z)) = \int P(z-x_1)\cdots P(z-x_k) :\zeta_\epsilon(x_1)\cdots\zeta_\epsilon(x_k): dx_1\cdots dx_k.$$

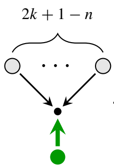
Indeed, assuming this identity holds for all $\ell < k$ we can prove that it also holds for k as follows. By (2.5'),

$$\zeta_\epsilon(x_1)\cdots\zeta_\epsilon(x_k) = \sum_{B\subset\{1,\ldots,k\}} :\{\zeta_\epsilon(x_i)\}_{i\in B}: \mathbf{E}\prod_{i\notin B}\zeta_\epsilon(x_i).$$

Multiplying both sides by $P(z-x_1)\cdots P(z-x_k)$ and integrating over $x_1\cdots x_k$, then using inductive assumption, one has

$$\begin{aligned}\Psi_\epsilon(z)^k = &\sum_{B\subsetneq\{1,\ldots,k\}} W_{|B|,\mu_\epsilon}(\Psi_\epsilon(z))\,\mathbf{E}\Big(\Psi_\epsilon(z)^{n-|B|}\Big)\\ &+\int\prod_{i=1}^k P(z-x_i) :\zeta_\epsilon(x_1)\cdots\zeta_\epsilon(x_k): \prod_{i=1}^k dx_i\end{aligned}$$

where the second term on the right hand side corresponds to the summand with $B = \{1,\ldots,k\}$, and to get the first term on the right hand side, we used the fact that P is deterministic and switched the order of expectation and integrations. By (2.5') again, we see that the second term on the RHS is precisely $W_{k,\mu_\epsilon}(\Psi_\epsilon(z))$, so the above identity for $W_{k,\mu_\epsilon}(\Psi_\epsilon(z))$ holds for all k. As a consequence, we have

$$\big(\widehat{\Pi}_0^\epsilon\tau\big)\big(\varphi_0^\lambda\big) = \big(\Pi_0^\epsilon M_{\mu_\epsilon}\tau\big)\big(\varphi_0^\lambda\big) = \quad \epsilon^{k-1}$$ 

. (4.12)

Here, each plain arrow $\longrightarrow$ has degree $a_e = 3$, which could be reduced by assigning suitable powers of ϵ's to it. We have the following proposition.

Proposition 4.7 *Both bounds in* (4.3) *hold for* $\tau = \mathcal{E}^{k-1}\Psi^{2k+1-n}$ *with* $k \geq 1$ *and* $n = 0, 1, 2, 3$.

Proof We first prove the first bound in (4.3), and briefly discuss how the second one follows from it immediately. In order to make use of the positive homogeneity of $\mathcal{E}$, we first assign powers of ϵ's to the edges of the graph in (4.12). We do it in the following way. Let $k \geq 1$ be given. Take any $2k-2$ edges connected to the noise from the $2k+1-n$ ones, and assign a $(\frac{1}{2}-\delta)$ power of ϵ to each of these edges. Thus, in the $2k+1-n$ noise edges, $3-n$ of them have degree $a_e = 3$, and all the rest $2k-2$ ones have degree $\frac{5}{2}+\delta$, and the graph comes with a multiplication of $\epsilon^{(2k-2)\delta}$.

Now, since τ has homogeneity $|\tau| = -\frac{1}{2}(3-n)-(2k+1-n)\kappa$, and the homogeneity (i.e. value of α in Theorem 4.6) of its associated graph (after the allocation of ϵ's) is

$$\frac{5}{2}|H_{\text{ex}}| - \sum_{e\in\mathcal{E}} a_e = \frac{5}{2}(2k+1-n) - 3\,(3-n) - \left(\frac{5}{2}+\delta\right)\,(2k-2)$$
$$= -\frac{1}{2}(3-n) - (2k-2)\,\delta,$$

it follows that if we take δ small enough, then in view of Theorem 4.6, it suffices to check all the conditions in Assumption 4.4.

The first condition is obvious. For the rest of the conditions, we take an arbitrary sub-graph $\bar{H}$ that contains ℓ noise vertices: p of them have degree 3, and the rest $\ell - p$ have degree $\frac{5}{2}+\delta$. By assumption, we know $p \leq 3$.

For (4.5), if $v_\star \in \bar{H}$, then the condition reads

$$3p + \left(\frac{5}{2}+\delta\right)(\ell - p) < 5\left(1+\frac{\ell}{2}\right),$$

which certainly holds since $p \leq 3$. If $v_\star \notin \bar{H}$, then the right hand side of (4.5) becomes 0, which trivially make the condition satisfied.

As for (4.6), since the graph does not contain any edge with $r_e \neq 0$, its left hand side is identical to that of (4.5), while the right hand side is strictly larger. Thus, the condition is automatically implied by (4.5) for this τ.

We finally turn to (4.7), where we only consider the case $v_\star \notin \bar{H}$. The RHS of (4.7) is $5\ell/2$, while the LHS is a sum of ℓ elements, each being at least $\frac{5}{2}+\delta$, so the condition also holds.

By Theorem 4.6, we have already shown the first bound in (4.3). As for the second one, the expression turns out to be a sum of graphs of the same type, but in each term exactly one instance of K is replaced by $K - K_{\bar{\epsilon}}$. The bound for the difference of the kernels together with the first bound in (4.3) immediately imply the second one. □

Remark 4.8 Condition (4.5) fails only when $p \geq 5$. Since we are in a regime where $p = 3$, this reflects the sub-criticality of our equation. It also indicates that the equation becomes critical when the nonlinear term is u_ϵ^5 but without any ϵ's in front of it.

Remark 4.9 There is a multiple of $\epsilon^{(2k-2)\delta}$ of the graph after the ϵ-allocation. This gives

$$\mathbf{E}\left|\left(\widehat{\Pi}_0^\epsilon \mathcal{E}^{k-1}\Psi^{2k+1-n}\right)\left(\varphi_0^\lambda\right)\right|^p \lesssim \epsilon^{(2k-2)\delta p}\lambda^{p(|\tau|+\theta)}.$$

In the case $k \geq 2$, the power of ϵ is strictly positive, which corresponds to the stronger bound in (4.4) with $\widehat{\Pi}_z \tau = 0$.

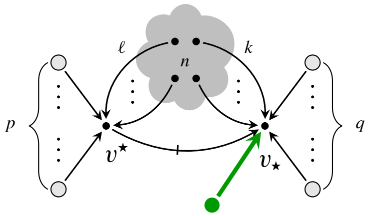
4.4 Bounds for second order objects in higher homogeneous chaos

The second order objects are homogeneous chaos decomposition of basis vectors of the form

$$\tau = \mathcal{E}^{\lfloor (k-1)/2 \rfloor} \big(\Psi^k \mathcal{I}(\mathcal{E}^{\lfloor \ell/2 \rfloor - 1} \Psi^\ell) \big). \tag{4.13}$$

Note that by the expansion of the right hand side of the abstract Eq. 3.15, the only negative-homogeneity terms that will *not* appear are those when k is odd and ℓ is even (which can be seen from 3.26). The action of the Wick renormalised model on this object yields a sum of terms in different homogeneous chaos. Each of them can be expressed by a graph of the type

$$\epsilon^{\frac{1}{2}(p+q+n-4-\delta_\tau)}$$



, (4.14)

where $\delta_\tau = 0$ if $k + \ell$ is even, and $\delta_\tau = 1$ if $k + \ell$ is odd. Here, k is the total number of edges connected to $v_\star$, and ℓ is the total number of connected to $v^\star$ (both excluding the one from $v^\star$ to $v_\star$). The gray cloudy area again stands for a product of cumulant functions over totally n points. The above graph represents the Wick renormalised object, and the effect of the mass renormalisation $M_0^{(\epsilon)}$ is not included, but when $p + q \geq 2$, the graphs will be the same with or without $M_0^{(\epsilon)}$. In this section, we focus on the case $p + q \geq 2$ (i.e. "higher chaos").

We first introduce some notations. Let $v^\star$ and $v_\star$ denote the starting and end nodes of the barred arrow, respectively, and $v_\star$ is connected to the node 0 by a test function. We let

$$\mathcal{P} = \{\text{external (non-contracted noise) vertices connected to } v^\star\} ,$$
$$\mathcal{P}' = \{\text{contracted (noise) vertices connected to } v^\star\}.$$

We define similarly $\mathcal{Q}$ and $\mathcal{Q}'$ except that the vertices are connected to $v_\star$, and let $\mathcal{N} = \mathcal{P}' \cup \mathcal{Q}'$. Thus, $\mathcal{N}$ is the set of all contracted vertices. If we let p, p', q, q' denote the cardinalities of the corresponding sets, then we have

$$p' + q' = n , \qquad p + q + n = k + \ell ,$$

and the total number of external (noise) vertices is $p + q$.

For simplicity of notations, we will sometimes use the same letter $\mathcal{N}$ also for edges that connects vertices in $\mathcal{N}$ with $v^\star$ or $v_\star$. We also make similar use of the notations $\mathcal{P}, \mathcal{P}'$, $\mathcal{Q}$ and $\mathcal{Q}'$.

Similar as before, in order to make use of the fact that each occurrence of $\mathcal{E}$ increases the homogeneity by 1, we need to assign powers of ϵ's to the edges connecting to the noise vertices (including the contracted ones) to reduce their degrees.[3] If $(k, \ell) = (1, 3), (2, 2)$ or $(2, 3)$, these are standard Φ^4_3 graphs and there will be no powers of ϵ to assign. In all other cases, there is always positive powers of ϵ to assign to the edges, and we do it in the following way.

1. Divide the total powers of ϵ's into $(p + q + n - 4 - \delta_\tau)$ pieces, each with power $(\frac{1}{2} - \delta)$ for some sufficiently small δ. We further divide these pieces into two groups such that
$$p + q + n - 4 - \delta_\tau = \big(p + p' - 2 - \mathbf{1}_{\{\ell \text{ odd}\}}\big) + \big(q + q' - 1 - \mathbf{1}_{\{k \text{ even}\}}\big).$$
This always holds since $n = p' + q'$, and $\mathbf{1}_{\{k \text{ even}\}} + \mathbf{1}_{\{\ell \text{ odd}\}} = 1 + \delta_\tau$ since we cannot have the situation when k is odd and ℓ is even. If $k + \ell - 4 - \delta_\tau \geq 1$, then there will always be positive powers of ϵ's left.
2. We assign the $\big(p + p' - 2 - \mathbf{1}_{\{\ell \text{ odd}\}}\big)$ pieces of $(\frac{1}{2} - \delta)$-power of ϵ to edges in $\mathcal{P}$ and $\mathcal{P}'$ in the following way. Assign one piece to each edge in $\mathcal{P}'$ until using up all $\big(p + p' - 2 - \mathbf{1}_{\{\ell \text{ odd}\}}\big)$ pieces. If there are still pieces left, we continue assigning one to each of the edges in $\mathcal{P}$ until finished.
3. We assign the rest $\big(q + q' - 1 - \mathbf{1}_{\{k \text{ even}\}}\big)$ to edges in $\mathcal{Q} \cup \mathcal{Q}'$ in the same way.

We then have the following proposition.

Proposition 4.10 *The above way of assigning ϵ's yields the object with the graph representation of the same type as in (4.14) but such that*

1. *Among the $p + p'$ edges from $\mathcal{P} \cup \mathcal{P}'$, $2 + \mathbf{1}_{\{\ell \text{ odd}\}}$ of them have degree $a_e = 3$, and all the rest have degree $\frac{5}{2} + \delta$. In addition, if there is some edge from $\mathcal{P}$ that has degree $\frac{5}{2} + \delta$, then all edges from $\mathcal{P}'$ have degree $\frac{5}{2} + \delta$.*
2. *Among the $q + q'$ edges from $\mathcal{Q} \cup \mathcal{Q}'$, $1 + \mathbf{1}_{\{k \text{ even}\}}$ of them have degree 3, and all the rest have degree $\frac{5}{2} + \delta$. In addition, if any edge from $\mathcal{Q}$ has degree $\frac{5}{2} + \delta$, then all edges from $\mathcal{Q}'$ have degree $\frac{5}{2} + \delta$.*
3. *There is a quantity ϵ^θ for some $\theta \geq 0$ multiplying the graph, and $\theta = 0$ if and only if $p + q + n - 4 - \delta_\tau = 0$.*

Proof Property 3 is obvious. Properties 1 and 2 come from the fact that we assign ϵ's to edges in $\mathcal{P}$ (or $\mathcal{Q}$) only after all edges in $\mathcal{P}'$ (or $\mathcal{Q}'$) are assigned. □

[3] By reducing degree of the kernel represented by an edge we mean implementing the bound $\epsilon^\alpha (|x| + \epsilon)^{-\beta} \lesssim (|x| + \epsilon)^{-\beta+\alpha}$.

Note that after the allocation of ϵ's with the above procedure, the degrees of edges in the graph satisfies

$$5|H_{\text{in}} \setminus H_\star| + \frac{5}{2}|H_{\text{ex}}| - \sum_e a_e = -\frac{1}{2}\delta_\tau - (p+q+n-4-\delta_\tau)\delta,$$

which is at the correct homogeneity (since $|\tau|$ is below $-\frac{1}{2}\delta_\tau$, so the above quantity is slightly bigger than $|\tau|$ if δ is small enough). Similar to the case for first-order objects, in view of Theorem 4.6, it now suffices to check that the graph in (4.14) with the properties in Proposition 4.10 does satisfy Assumption 4.4. In the following we still call this graph $(H, \mathcal{E})$.

But this time, the graph H is more complicated, and it is hard to check all the conditions in Assumption 4.4 for all sub-graphs in a straightforward way. The following lemma gives a simpler procedure in the verification of this assumption. It roughly states that it suffices to check a very small set of sub-graphs $\bar{H}$, and all other sub-graphs will automatically satisfy the assumption if that small set does. To state the lemma we define the following sets of vertices

$$\mathcal{P}^\star = \{v^\star\} \cup \mathcal{P}\;, \quad \mathcal{Q}_\star = \{v_\star\} \cup \mathcal{Q}\;, \quad \mathcal{N}_\star^\star = \{v_\star, v^\star\} \cup \mathcal{N}.$$

Lemma 4.11 *Let $(H, \mathcal{E})$ be the labelled graph defined above.*

- *If for every $\bar{H} \subset H_0$ of cardinality at least 3 such that*

$$\bar{H} \cap \mathcal{P}^\star \in \{\emptyset, \{v^\star\}, \mathcal{P}^\star\}\;, \quad \bar{H} \cap \mathcal{Q}_\star \in \{\emptyset, \{v_\star\}, \mathcal{Q}_\star\}\;, \quad \bar{H} \cap \mathcal{N}_\star^\star \in \{\emptyset, \mathcal{N}_\star^\star\}$$

  satisfies (4.5), then item 2 of Assumption 4.4 is satisfied.
- *If for every $\bar{H} \subset H$ containing 0 of cardinality at least 2 such that*

$$\bar{H} \cap \mathcal{P}^\star \in \{\emptyset, \mathcal{P}^\star\}\;, \quad \bar{H} \cap \mathcal{Q}_\star \in \{\emptyset, \mathcal{Q}_\star\}\;, \quad \bar{H} \cap \mathcal{N}_\star^\star \in \{\emptyset, \mathcal{N}_\star^\star\}$$

  satisfies (4.6), then item 3 of Assumption 4.4 is satisfied.
- *If for every non-empty $\bar{H} \subset H_\star$ such that*

$$\bar{H} \cap \mathcal{Q}_\star = \emptyset\;, \quad \bar{H} \cap (\mathcal{P}^\star \cup \mathcal{N}_\star^\star) \in \{\emptyset, \mathcal{P}^\star \cup \mathcal{N}_\star^\star\}$$

  satisfies (4.7), then item 4 of Assumption 4.4 is satisfied.

Remark 4.12 The above lemma should be understood in the following way: in each case, the description of the set $\bar{H}$ given above is the "worst" case corresponding to the condition, and if the bound is satisfied for that "worst" case, then it will automatically be satisfied for all other sub-graphs. For example, for condition (4.6), the lemma states that if $v^\star \in \bar{H}$ but $v_\star$ not, and the sub-graph $\{v^\star\} \cup \mathcal{P}$ satisfies bound (4.6), then all other sub-graphs $\bar{H}$ that contains $v^\star$ but not $v_\star$ automatically satisfies (4.6).

Note that in the Item 1 above (for Condition 4.5), we impose further restrictions only for the case when $v^\star$ or $v_\star$ is not in $\bar{H}$, but not the case that they are in $\bar{H}$ (unlike Item 2 for Condition 4.6); this is due to the additional term $\mathbf{1}_{\bar{H}_{\text{ex}}=\emptyset}$ in (4.5).

Proof We first claim that if some but not all vertices from $\mathcal{P}$ (resp. $\mathcal{Q}$, or one connected component of $\mathcal{N}$) are in $\bar{H}$, then we can always worsen the bounds for the corresponding conditions by adding to $\bar{H}$ or removing from $\bar{H}$ vertices of $\mathcal{P}$ (resp. $\mathcal{Q}$, or one connected component of $\mathcal{N}$). (By "worsening" a bound or an inequality, we mean adding or subtracting numbers on both sides so that the difference between the values of the two sides becomes smaller). This will imply that one *never* needs to consider the case when $\bar{H}$ contains some but not all of the vertices from $\mathcal{P}$, or $\mathcal{Q}$, or a connected component of $\mathcal{N}$.

To see this, we first consider the set $\mathcal{P}$. Suppose that $\bar{H} \cap \mathcal{P} \notin \{\emptyset, \mathcal{P}\}$. For (4.5) and (4.6):

- If $v^\star \in \bar{H}$, then adding one more vertex from $\mathcal{P}$ into $\bar{H}$ will increase the left hand sides of (4.5) and (4.6) by 3 or $\frac{5}{2} + \delta$, while increase the right hand sides by only $\frac{5}{2}$. Thus, adding more vertices of $\mathcal{P}$ into $\bar{H}$ makes both bounds worse.
- If $v^\star \notin \bar{H}$, then removing one vertex in $\mathcal{P}$ from $\bar{H}$ will not change the left hand sides, but decrease both right hand sides (by at least $\frac{5}{2}$), thus also worsen the bounds.

For the bound (4.7),

- if $v^\star \in \bar{H}$, then adding each other vertex in $\mathcal{P}$ into $\bar{H}$ does not increase the left hand side, but increases the right hand side by $\frac{5}{2}$;
- if $v^\star \notin \bar{H}$, then removing each vertex in $\mathcal{P}$ from $\bar{H}$ decreases the left hand side by 3 or $\frac{5}{2} + \delta$, but decreases the right hand side by $\frac{5}{2}$ only.

So in either case the bound becomes worse. This shows that for all the three conditions, the worst case is $\bar{H} \cap \mathcal{P} \in \{\emptyset, \mathcal{P}\}$. The conclusion for $\mathcal{Q}$ follows from exactly the same argument.

For the set of contracted vertices $\mathcal{N}$, let N be one of the hyper-edges in $\mathcal{N}$, and suppose that $\bar{H} \cap N \notin \{\emptyset, N\}$. Then, for bounds (4.5) and (4.6), removing one vertex of N from $\bar{H}$ decreases both left hand sides by at most 3, but decreases the right hand sides by 5, which worsens the bounds. For (4.7), adding additional vertices from N into $\bar{H}$ will increase the left hand side by at most 3 (depending on whether $v^\star \in \bar{H}$ or not), but will increase the right hand side by 5, which also worsens the bound. Therefore, for each hyper-edge in $\mathcal{N}$, we also only need to consider the whole set instead of part of it.

With the above claimed fact we proceed the proof as follows.

For (4.5), it suffices to show that when $v^\star \notin \bar{H}$ the bound without $\mathcal{P}$ is worse than the one with $\mathcal{P}$. Indeed if $v^\star \notin \bar{H}$, then adding all of $\mathcal{P}$ does not make a difference to the left hand side, but yields an increment of $\frac{5}{2}(p+1)$ on the right hand side. (Note that the increment of the right hand side is $\frac{5}{2}(p+1)$ rather than $\frac{5p}{2}$ due to the additional term $\mathbf{1}_{\bar{H}_{\mathrm{ex}}=\emptyset}$). Similarly, one can verify that it suffices to check the case $\mathcal{Q}$ is not in $\bar{H}$ if $v_\star \notin \bar{H}$.

As for noise vertices $\mathcal{N}$, there are two situations:

- If $\{v^\star, v_\star\} \subset \bar{H}$, then adding $\mathcal{N}$ into $\bar{H}$ increases the LHS of (4.5) by

$$\sum_{e\in\mathcal{N}} a_e + \frac{5n}{2}, \tag{4.15}$$

where we have used the notation $\mathcal{N}$ also for edges that connect contracted vertices in $\mathcal{N}$ with $v^\star$ and $v_\star$. Since $|\mathcal{N}| = n$, and each edge in it has degree at least $\frac{5}{2} + \delta$, (4.15) is clearly larger than the increment of the RHS (which is $5n$). So the bound with $\mathcal{N}$ in $\bar{H}$ implies the one with $\mathcal{N}$ not in $\bar{H}$.

- If either $v^\star$ or $v_\star$ is not in $\bar{H}$, then the increment of the RHS of (4.5) is still $5n$, but that of the LHS takes the form (4.15) with $\sum_{e\in\mathcal{N}}$ replaced by $\sum_{e\in\tilde{\mathcal{N}}}$ where $\tilde{\mathcal{N}}$ is a strict subset of $\mathcal{N}$ (so $|\tilde{\mathcal{N}}| \leq n-1$). Now, each edge has degree either $\frac{5}{2} + \delta$ or 3, and there are at most 5 edges with degree 3 according to 4.10, so the increment of the LHS would be at most

$$5 \cdot 3 + (n-6) \cdot \left(\frac{5}{2} + \delta\right) + \frac{5n}{2} = 5n + (n-6)\delta.$$

  If condition (4.5) holds with $\mathcal{N}$ not in $\bar{H}$, then adding a multiple of δ which can be chosen arbitrarily small to the LHS will not change the validity of the bound, and therefore (4.5) also holds with all vertices of $\mathcal{N}$ being in $\bar{H}$.

The case for (4.6) is essentially the same, except for the noise vertices $\mathcal{P}$ and $\mathcal{Q}$ when $v^\star$ or $v_\star$ are in $\bar{H}$. If $v^\star \in \bar{H}$, then putting $\mathcal{P}$ into $\bar{H}$ yields an increment of $\sum_{e\in\mathcal{P}} a_e$ on the left hand side, which is strictly larger than $\frac{5p}{2}$. Thus, the bound is worse if $\mathcal{P} \subset \bar{H}$. Note that the conclusion here does not hold for (4.5) since the increment on the right hand side would be $\frac{5}{2}(p+1)$ instead of $\frac{5p}{2}$. The case for $v_\star$ is the same.

We finally turn to (4.7). Since $v_\star \notin \bar{H}$, including $\mathcal{Q}$ in $\bar{H}$ would yield an increment of $\sum_{e\in\mathcal{Q}} a_e$ on the left hand side, which is strictly bigger than $\frac{5q}{2}$. Since the inequality in this condition is reversed, the bound actually becomes better. So we should only consider the worst case that $\mathcal{Q}$ is not in $\bar{H}$. One can argue in the same way that we should include $\mathcal{P}$ in $\bar{H}$ if and only if $v^\star \in \bar{H}$. For contracted vertices $\mathcal{N}$, if $v^\star \in \bar{H}$, then adding $\mathcal{N}$ increases the left hand side by $\frac{5n}{2}$ but the right hand side by $5n$. If $v^\star \notin \bar{H}$, then the increment of the left hand side by adding $\mathcal{N}$ is exactly the quantity (4.15), which is bigger than $5n$. Thus, we should also include $\mathcal{N}$ into $\bar{H}$ if and only if $v^\star \in \bar{H}$.

This completes the proof. □

Proposition 4.13 *If $p + q \geq 2$, then the graph* (4.14) *satisfies Assumption* 4.4.

Proof The first condition of Assumption 4.4 is trivial as each edge has degree at most 3. We now give some details in the verification of the other three.

1. Condition (4.5). We first consider the case $\{v^\star, v_\star\} \subset \bar{H}$. According to Lemma 4.11, we only need to look at the situation when $\mathcal{N} \subset \bar{H}$. But Item 1 in that lemma does not give specifications in this case whether to include $\mathcal{P}$, $\mathcal{Q}$ or not, so we need to consider all four possibilities for $\mathcal{P}$ and $\mathcal{Q}$. If both $\mathcal{P}$ and $\mathcal{Q}$ are in $\bar{H}$, then condition (4.5) reads

$$\sum_{e\in\mathcal{P}\cup\mathcal{Q}\cup\mathcal{N}} a_e + 3 + \frac{5n}{2} < 5\big(n + 2 + \frac{1}{2}(p+q-1)\big),$$

which certainly holds since $|\mathcal{P} \cup \mathcal{Q} \cup \mathcal{N}| = p + q + n$ and all a_e in that sum have degree $\frac{5}{2} + \delta$ except at most 5 of them which have degree 3. If neither $\mathcal{P}$ nor $\mathcal{Q}$ is in $\bar{H}$, then if we let r denote the number of contracted edges that have degree 3, the condition reads

$$\left(\frac{5}{2} + \delta\right) n + \left(\frac{1}{2} - \delta\right) r + 3 + \frac{5n}{2} < 5(n + 1). \tag{4.16}$$

Since $p + q \geq 2$, by the assumption on the ϵ-allocation, we necessarily have $r \leq 3$, so the condition holds for all small enough δ. If one of $\mathcal{P}$ and $\mathcal{Q}$ is in $\bar{H}$ but the other not, then it is easy to see that the increment of the right hand side of (4.16) is larger than that of the left hand side (the right hand side has an additional increment $\frac{5}{2}$ since $\bar{H}_{\text{ex}} \neq \emptyset$), so Condition (4.5) also holds.
We now turn to the situation $\{v^\star, v_\star\} \cap \bar{H} = \emptyset$. In this case, the "worst" sub-graph $\bar{H}$ according to Lemma 4.11 is that $\bar{H} = \emptyset$. Note that the bound (4.5) only requires $|\bar{H}| \geq 3$, and in this case adding any three or more vertices into $\bar{H}$ keeps the left hand side 0 but yields a positive quantity on the right hand side (except adding all of $\mathcal{N}$, but then the increments are $\frac{5n}{2} < 5n$, which is still fine). Finally, the case when either $v^\star$ or $v_\star$ is in $\bar{H}$ but the other not is easy to verify. Thus, we conclude that the condition (4.5) is satisfied for all sub-graphs $\bar{H}$ with at least three vertices.
2. Condition (4.6). The case $\{v^\star, v_\star\} \subset \bar{H}$ is automatically implied by the verified bound (4.5), as here the left hand sides for both conditions are the same since neither term involving r_e is counted, but the right hand side of (4.6) is larger than that of (4.5). If neither $v^\star$ nor $v_\star$ is in $\bar{H}$, then as long as $\bar{H}$ contains one vertex other than 0, the right hand side is strictly positive while the left hand side is still 0, so the bound is also verified.
If $v^\star \in \bar{H}$ but $v_\star \notin \bar{H}$, then by Lemma 4.11, we only need to include $\mathcal{P}$ in $\bar{H}$, so the condition reads

$$\sum_{e \in \mathcal{P}} a_e + (3 + 1 - 1) < 5\left(1 + \frac{p}{2}\right).$$

Since the number of edges in $\mathcal{P}$ of degree 3 is at most three, the above bound obviously holds. The case $v_\star \in \bar{H}$ but $v^\star \notin \bar{H}$ can be checked in the same way. This completes the verification for (4.6).
3. We finally turn to Condition 4.7. Here we assume $v_\star \notin \bar{H}$. If $v^\star \notin \bar{H}$, then the worst situation is $\bar{H} = \emptyset$, and we have $0 = 0$. Since any other case will yield strictly better bound than this one, the bound then holds for any non-empty $\bar{H}$ that does not contain $v^\star$ and $v_\star$. If $v^\star \in \bar{H}$, then the worst case is that

$$\bar{H} = \{v^\star\} \cup \mathcal{P} \cup \mathcal{N},$$

so the condition becomes

$$\sum_{e \in \mathcal{P} \cup \mathcal{N}} a_e + \frac{5n}{2} + 3 + 1 > 5\left(n + 1 + \frac{p}{2}\right).$$

Again, let r denote the total number of edges in the above sum (in $\mathcal{P} \cup \mathcal{N}$) that have degree 3, then the left hand side above is

$$\left(\frac{5}{2}+\delta\right)(p+n)+\left(\frac{1}{2}-\delta\right)r+\frac{5n}{2}+4.$$

If $\ell \geq 3$, then $r \geq 3$ so the bound always holds. The bound also holds if $\ell = 2$ but $p+n \geq 3$. In the case of the 4-th homogeneous chaos of $\Psi^2\mathcal{I}(\Psi^2)$, we have $p+n=r=2$, so one gets an equality instead of a strict inequality in Condition (4.7). However, we can treat the barred arrow as having degree $a_3 = 3+\delta$ rather than 3. This will not violate any of the assumption on our model and give us a strictly inequality in (4.7).

We have now completed the verification for second order objects with chaos order $p+q \geq 2$. □

Remark 4.14 The assumption $p+q \geq 2$ is essential, as Condition (4.5) is indeed violated when $p+q \leq 1$. For example, for the graph representing the logarithmic divergence in (4.9), if we take $\bar{H} = H_0$, then $|\bar{H}_{\text{in}}| = 6$ and $\bar{H}_{\text{ex}} = \emptyset$, so both sides of (4.5) are 25, and the *inequality* does not hold. In fact, the same is true for all such graphs with $p+q \leq 1$, and the mass renormalisation M_0 is needed in order for them to satisfy Assumption 4.4.

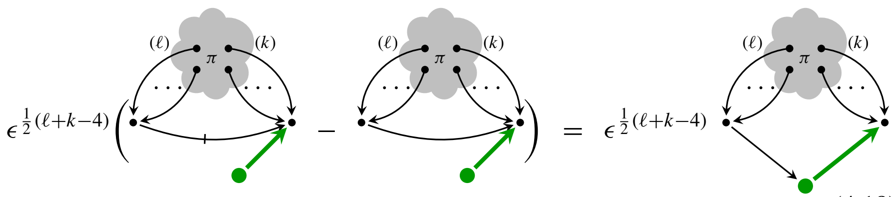
4.5 Bounds for second order objects in 0-th and 1-st homogeneous chaos

We now turn to bounding the objects in the 0-th and 1-st order chaos (that is, when $p+q \leq 1$). We will need the following lemma taken from Lemma 4.7 in [18].

Lemma 4.15 *Given space-time points $y_1, \ldots, y_n$ and $-5 < \alpha_i < 0$, one has*

$$\int \prod_{i=1}^{n} |y_i - x_i|^{\alpha_i} |\mathfrak{C}_n^{(\epsilon)}(x_1,\ldots,x_n)|\, dx_1 \cdots dx_n \lesssim \epsilon^{5(n/2-1)} \int_{\mathbf{R}^4} \prod_{i=1}^{n} \big(|y_i - x| + \epsilon\big)^{\alpha_i}\, dx. \tag{4.17}$$

The 0-th order objects only occur when $k+\ell$ is even. After renormalisations by $M_0^{(\epsilon)}$ (subtraction of the constants $C_{k,\ell}^{(\epsilon)}$'s as defined in 4.11), they are given by



(4.18)

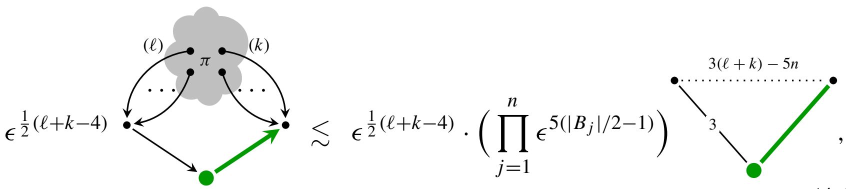
which represents a deterministic quantity. We have the following bound for it. This in particular implies the first bound of (4.3).

Proposition 4.16 *The quantity on the right hand side of* (4.18) *is bounded by* $\epsilon^{\theta}\lambda^{-\delta}$ *for some* $\theta \geq 0$ *and all sufficiently small* δ. *Moreover,* $\theta = 0$ *if and only if* $k = \ell = 2$ *and* π *is the pair-wise contraction.*

Proof Suppose $\ell + k \geq 6$, or $\ell + k = 4$ but π is the partition that contracts all four points together. Then, we have

$$\pi = B_1 \cup \cdots \cup B_n$$

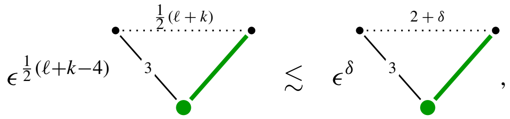
such that either $|B_j| \geq 4$ for some j or $\sum_j |B_j| \geq 6$. By consecutively applying Lemma 4.15 and [15, Lemma 10.14][4], the object can then be bounded by



(4.19)

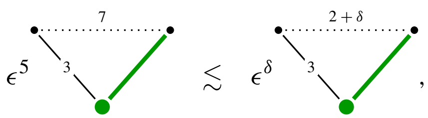
where the dotted line ········· with degree a_e denotes a kernel that is bounded by $(|z| + \epsilon)^{-a_e}$, where z is the difference between two end points of the line.

If $\ell + k \geq 6$, then the right hand side above is bounded by



where the first quantity is obtained by putting the product $\prod \epsilon^{5(|B_j|/2-1)}$ into the top edge of the graph and using $\sum_j |B_j| = \ell + k$, and the bound uses the assumption that $\ell + k \geq 6$. This gives the convergence to 0 at the desired topology when $\ell + k \geq 6$.

If $\ell + k = 4$ but π is the single partition that contracts all four points together, then $\pi = B$ with $|B| = 4$. The right hand side of (4.19) could then be reduced to

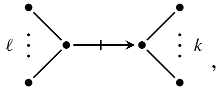


which also gives the convergence to 0 at the expected scale.

The case when $k = \ell = 2$ and π being the pair-wise contraction is the same, except that no positive power of ϵ could be created. This completes the proof. □

[4] We actually need a modified version of Lemma 10.14 in [15]: $\int (|x-y|+\epsilon)^{-a}(|y|+\epsilon)^{-b} dy \lesssim (|x|+\epsilon)^{-a-b+|s|}$ for all large a and b (no need to assume that a and b are smaller than the dimension).

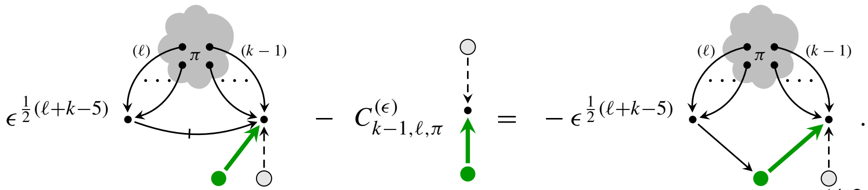
We finally turn to the objects in the first chaos. They are obtained from the terms of the form



(4.20)

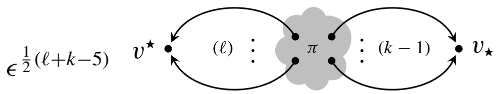
multiplied by $\epsilon^{\frac{1}{2}(\ell+k-5)}$. Here, ℓ is odd, k is even, and $\ell \geq 3$, $k \geq 2$. The elements in the first chaos of this object are obtained either from contracting ℓ vertices on the left with $k-1$ vertices on the right, or contracting $\ell-1$ vertices on the left with k vertices on the right. It turns out that we have a similar statement as in the 0-th chaos: after mass renormalisation, the first chaos of the quantity $(\widehat{\Pi}_0^\epsilon \tau)(\varphi_0^\lambda)$ has p-th moment bounded by $\epsilon^\theta \lambda^{(|\tau|+\delta)}$ for $\theta \geq 0$. Furthermore, $\theta = 0$ if and only if $k=2$, $\ell=3$ and π is the pair-wise contraction (which could only happen in contraction of the second type).

To see this, we first consider the contraction of the first type. The object in the first chaos is given by



(4.21)

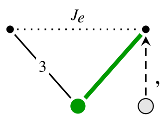
Again, by applying Lemma 4.15 above and [15, Lemma 10.14], it is not hard to check that the sub-graph



represents a kernel $J(v^\star - v_\star)$ such that

$$\|J_e\|_{2+\delta;\beta} \lesssim \epsilon^\theta \tag{4.22}$$

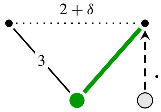
for all $\delta, \beta > 0$ and some $\theta > 0$ depending on δ. The degree of the kernel $2+\delta$ is relatively easy to check. To see that the norm is bounded by some positive power of ϵ, we first note if $k+\ell \geq 7$, then there is already some ϵ's multiplying the graph, and one could allocate these powers to the edges in such a way that there is always some positive power left. In the case $k=2$ and $\ell=3$, then π must be the contraction that groups together all three vertices from one side and one from the other side. In this case, one could also create a positive power of ϵ as in Proposition 4.16. Thus, in all cases, the right hand side of (4.21) could be reduced to the graph



(4.23)

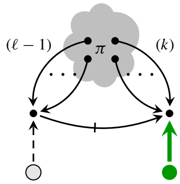
where the J_e satisfies the bound (4.22). Since its degree a_e is $2+\delta$, it is straightforward to check Assumption 4.4. By Theorem 4.6 and the positive power of ϵ in front of $\|J_e\|_{2+\delta}$, we see that the L^p norm of the quantity on the right hand side of (4.21) is bounded by $\epsilon^\theta \lambda^{-\frac{1}{2}-\delta}$. Again, we can take δ small enough so that it vanishes with the correct homogeneity.

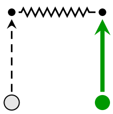
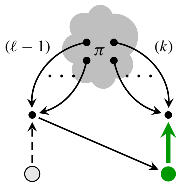
Remark 4.17 In what follows, we will represent the graph in (4.23) and similar graphs by

ϵ^θ  . (4.24)

The use of this graph here is of course ambiguous as it *does not* suggest that the quantity in (4.21) is bounded by such a graph. In fact, the correct bound is (4.23), which is different from the one above. But for simplicity of notations, we choose to write ϵ^θ outside the graph and to regard the norm of the upper-edge being $\mathcal{O}(1)$.

We now turn to the second type of contractions. For those contractions, we have

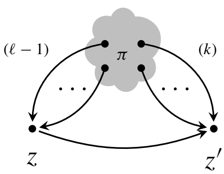
$\epsilon^{\frac{1}{2}(\ell+k-5)}$  $- \; C^{(\epsilon)}_{k,\ell-1,\pi}$ 

$= \; \epsilon^{\frac{1}{2}(\ell+k-5)} \Bigg($  $-$  $\Bigg)$, (4.25)

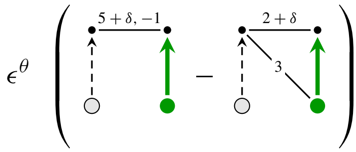
where the symbol  denotes the renormalised distribution

$$Q^{(\epsilon)}(\varphi) := \int Q^{(\epsilon)}(z)\big(\varphi(z) - \varphi(0)\big)dz,$$

where the kernel Q is given by

$Q^{(\epsilon)}(z-z') = Q^{(\epsilon)}_{k,\ell-1,\pi}(z-z') :=$  .

It is clear that the waved line  carries the renormalisation label $r_e = -1$. If $\ell+k \geq 7$, or $\ell = 3$, $k = 2$ but π contracts four points altogether, then either the graph itself carries ϵ's (when $\ell + k \geq 7$), or we can create a positive power of ϵ using the same argument as in Proposition 4.16. As a consequence, we can assign powers of ϵ's to the edge (wavy edge) so that it has degree $a_e = 5 + \delta$. It then follows that the second line of (4.25) is bounded by



for some $\theta > 0$. It is straightforward to check that both graphs satisfy Assumption 4.4, so their p-th moments are bounded by $\epsilon^{\theta} \lambda^{-(\frac{1}{2}+\delta)p}$. The case when $\ell = 3$, $k = 2$ and π being the pair-wise contraction is the same except that $\theta = 0$, since there would be no room to create any positive power of ϵ. Finally, we can choose δ small enough such that $-(\frac{1}{2} + \delta) > |\tau|$ so these bounds imply the first bound in Theorem 4.1. The second bound for these objects can be obtained in the same way by considering the difference of the kernels $K_{\bar{\epsilon}} - K$ and applying (4.2).

5 Convergence to the limit

In this section, we collect all the results from the previous sections to identify the limiting process for u_ϵ. Recall from (2.10) that our rescaled process u_ϵ satisfies the equation

$$\partial_t u_\epsilon = \Delta u_\epsilon - \sum_{j=0}^{m} \widehat{a}_j(\theta) \epsilon^{j-1} W_{2j+1,\mu_\epsilon}(u_\epsilon) + \zeta_\epsilon, \tag{5.1}$$

and we would like to show that u_ϵ converges to the solution of $\Phi^4_3(\widehat{a}_1)$ family.

We first give a convergence result for renormalised models, which is the following theorem. We will denote by $\mathfrak{M}$ the model which yields the solution of the Φ^4_3 equation, that is, the limiting model obtained in [15, Theorem 10.22] (called $\hat{Z}$ therein), [5] and call $\mathfrak{M}$ the "Φ^4_3 model". It is also the same as the limiting model $\mathfrak{M}$ obtained in [19, Theorem 4.3, Theorem 5.1] since as discussed in [19, Proof of Theorem 5.1], $\mathfrak{M}$ acting on any element which contains a positive power of $\mathcal{E}$ yields zero, and otherwise yield the same result as the Φ^4_3 model.

Theorem 5.1 *Let $\mathfrak{M}_\epsilon = M_\epsilon \mathscr{L}_\epsilon(\zeta_\epsilon)$ be the renormalised models defined in Sect. 3.3 and Theorem 4.1, and $\mathfrak{M}$ be the Φ^4_3 model. Then, we have $|||\mathfrak{M}_\epsilon; \mathfrak{M}|||_{\epsilon;0} \to 0$ where $|||-;\, -|||_{\epsilon;0}$ is defined in (3.13).*

[5] Strictly speaking $\mathfrak{M}$ depends on the particular choice of a solution in the $\Phi^4_3(\widehat{a}_1)$ family, but we do not reflect this in the notation.

Proof By the definition (3.13), we have

$$\begin{aligned} \|\mathfrak{M}_\epsilon ; \mathfrak{M}\|_{\epsilon;0} =& \|\mathfrak{M}_\epsilon ; \mathfrak{M}\| + \|\mathfrak{M}_\epsilon\|_\epsilon \\ &+ \sup_{t\in[0,1]} \|\mathcal{R}^\epsilon \mathcal{I}(\Xi)(t,\cdot) - \mathcal{R}\mathcal{I}(\Xi)(t,\cdot)\|_{\mathcal{C}^\eta}, \end{aligned} \tag{5.2}$$

where $\mathcal{R}^\epsilon$ and $\mathcal{R}$ denotes the reconstruction operators for $\mathfrak{M}_\epsilon$ and $\mathfrak{M}$ respectively. We first deal the second term on the right hand side of (5.2). By the same calculations as in Section 4 of [19], there exists $\theta > 0$ such that

$$\mathbf{E}|\widehat{f}_z^\epsilon(\mathscr{E}_\ell^k \tau)| \lesssim \epsilon^{|\tau|+k-|\ell|+\theta} \tag{5.3}$$

for every $\tau \in \mathcal{W}$ such that $|\mathcal{E}^k(\tau)| > 0$, and

$$\mathbf{E}|\widehat{\Pi}_z^\epsilon \tau(\psi_z^\lambda)| \lesssim \lambda^{\beta+\theta}\epsilon^{|\tau|-\beta}, \qquad \beta = \frac{6}{5} \tag{5.4}$$

for all test function ψ that annihilates affine functions. Note that here the non-Gaussianity of the noise does not create additional difficulty since we are only concerned about the first moment. In addition, both bounds (5.3) and (5.4) are uniform over all space-time points z in compact sets and all $\epsilon, \lambda < 1$. By definition of $\|\cdot\|_\epsilon$ (see 3.12), it follows immediately that

$$\|\mathfrak{M}_\epsilon\|_\epsilon \to 0 \tag{5.5}$$

in probability.

For the third term on the right hand side of (5.2), we have the desired bound if the models are built from Gaussian noise (which is the case studied in [19]). In order to make use of result in the Gaussian case, we follow a diagonal argument as in [18, Section 6]. Recall that for $\epsilon, \bar{\epsilon} \geq 0$, we have set

$$\zeta_{\epsilon,\bar{\epsilon}} := \zeta_\epsilon * \rho_{\bar{\epsilon}}$$

for some space-time mollifier ρ at scale $\bar{\epsilon}$. By triangle inequality, we have for every $\eta < -\frac{1}{2}$ the bound

$$\begin{aligned} \|\mathcal{R}^\epsilon \mathcal{I}(\Xi)(t,\cdot) - \mathcal{R}\mathcal{I}(\Xi)(t,\cdot)\|_{\mathcal{C}^\eta} \leq& \|K * \zeta_\epsilon(t,\cdot) - K * \zeta_{\epsilon,\bar{\epsilon}}(t,\cdot)\|_{\mathcal{C}^\eta} \\ &+ \|K * \zeta_{\epsilon,\bar{\epsilon}}(t,\cdot) - K * \zeta_{0,\bar{\epsilon}}(t,\cdot)\|_{\mathcal{C}^\eta} + \|K * \zeta_{0,\bar{\epsilon}}(t,\cdot) - K * \xi(t,\cdot)\|_{\mathcal{C}^\eta}. \end{aligned} \tag{5.6}$$

The last term in (5.6) only involves Gaussian objects, and is independent of ϵ. Thus, it follows from [15, Proposition 9.5] that the expectation of this term is bounded by $\bar{\epsilon}^\theta$ for some positive θ. Also, for the first term on the right hand side of (5.6), since $K * \zeta_{\epsilon,\bar{\epsilon}} = \rho_{\bar{\epsilon}} * K * \zeta_\epsilon$, it follows from [15, (9.16)] that

$$\|K * \zeta_\epsilon(t,\cdot) - K * \zeta_{\epsilon,\bar{\epsilon}}(t,\cdot)\|_{\mathcal{C}^\eta} \lesssim \bar{\epsilon}^\theta \|K * \zeta_\epsilon\|_{\mathcal{X}},$$

where $\mathcal{X} = \mathcal{C}^{\frac{\theta}{2}}(\mathbf{R}, \mathcal{C}^{\eta+\theta}(\mathbf{T}^3))$, and the proportionality constant is independent of $\epsilon < 1$ and ζ_ϵ. By taking θ small enough, we get

$$\mathbf{E}\|K * \zeta_\epsilon(t,\cdot) - K * \zeta_{\epsilon,\bar{\epsilon}}(t,\cdot)\|_{\mathcal{C}^\eta} \lesssim \bar{\epsilon}^\theta,$$

uniformly over all $\epsilon < 1$. Regarding the second term on the right hand side of (5.6), for any fixed $\bar{\epsilon}$, the convergence to 0 in $\mathcal{C}^\eta$ as $\epsilon \to 0$ is obvious since everything is smooth once $\bar{\epsilon}$ is fixed. Thus, by sending $\epsilon \to 0$ first, and then $\bar{\epsilon} \to 0$, we deduce from these arguments and (5.6) that

$$\mathbf{E}\|\mathcal{R}^\epsilon \mathcal{I}(\Xi)(t,\cdot) - \mathcal{R}\mathcal{I}(\Xi)(t,\cdot)\|_{\mathcal{C}^\eta} \to 0.$$

We finally turn to the first term on the right hand side of (5.2), $|||\mathfrak{M}_\epsilon; \mathfrak{M}|||$. For this term, we also use the diagonal argument as above and apply the known results for Gaussian models. Let

$$\mathfrak{M}_{\epsilon,\bar{\epsilon}} := M_{\epsilon,\bar{\epsilon}} \mathscr{L}_\epsilon(\zeta_{\epsilon,\bar{\epsilon}}),$$

then we have

$$\mathbf{E}|||\mathfrak{M}_\epsilon; \mathfrak{M}||| \leq \mathbf{E}|||\mathfrak{M}_\epsilon; \mathfrak{M}_{\epsilon,\bar{\epsilon}}||| + \mathbf{E}|||\mathfrak{M}_{\epsilon,\bar{\epsilon}}; \mathfrak{M}_{0,\bar{\epsilon}}||| + \mathbf{E}|||\mathfrak{M}_{0,\bar{\epsilon}}; \mathfrak{M}|||. \tag{5.7}$$

Note that all of the norms above are usual norms on modelled distributions, and they do not depend on ϵ. The last term is the distance between the Gaussian model $\mathfrak{M}_{0,\bar{\epsilon}}$ and $\mathfrak{M}$, and is bounded by $\bar{\epsilon}^\theta$ for some positive θ by the convergence result [19, Theorem 4.3]. For the first term, by Theorem 4.1 and then the arguments in the proof of [15, Theorem 10.7], we can show that it is uniformly (in ϵ) bounded by $\bar{\epsilon}^\theta$ for some positive θ. Thus, by sending $\epsilon \to 0$, (5.7) reduces to

$$\limsup_{\epsilon \to 0} \mathbf{E}|||\mathfrak{M}_\epsilon; \mathfrak{M}||| < C\bar{\epsilon}^\theta + \limsup_{\epsilon \to 0} \mathbf{E}|||\mathfrak{M}_{\epsilon,\bar{\epsilon}}; \mathfrak{M}_{0,\bar{\epsilon}}|||,$$

where C is independent of ϵ. Now, the only remaining term involves two smooth models. For any fixed $\bar{\epsilon}$, by the arguments in [18, Proof of Theorem 6.5], this term also vanishes as $\epsilon \to 0$. By further sending $\bar{\epsilon}$ to 0, we deduce that

$$\mathbf{E}|||\mathfrak{M}_\epsilon; \mathfrak{M}||| \to 0.$$

This concludes the proof of the theorem. □

We are now ready to state and prove our final result.

Theorem 5.2 *Let ζ and V_θ satisfy Assumptions 1.3 and 1.4. Let $\phi_0^{(\epsilon)} \in \mathcal{C}_\epsilon^{\gamma,\eta}$ be such that $\|\phi_0^{(\epsilon)}; \phi_0\|_{\gamma,\eta;\epsilon} \to 0$ for some $\phi_0 \in \mathcal{C}^\eta$. Let u_ϵ solves the PDE (5.1) with initial condition $\phi_0^{(\epsilon)}$, and let $T > 0$ be an arbitrary deterministic time. Then, there exists $a < 0$ such that for $\theta = a\epsilon|\log\epsilon| + \mathcal{O}(\epsilon)$, u_ϵ converges in law to the $\Phi_3^4(\widehat{a}_1)$ family* *of solutions with initial data* ϕ_0*, where* $\widehat{a}_1 = \frac{\partial^4 \langle V \rangle}{\partial x^4}(0,0)$*, and the convergence takes place in* $\mathcal{C}([0,T], \mathcal{C}^\eta(\mathbf{T}^3))$.

Proof Fix $T > 0$ arbitrary. By the global existence of Φ^4_3 in [16,24], the limiting model and hence the limiting solution u it represents exist up to time T.

Now, let $C^{(\epsilon)}_{k,\ell}$'s be the renormalisation constants defined in (4.11), and M_ϵ be the renormalisation map defined with these constants. By Theorem 3.5, if $\Phi \in \mathcal{D}^{\gamma,\eta}_\epsilon$ solves the fixed point Eq. (3.15), then the function $v_\epsilon = \mathcal{R}^\epsilon \Phi^{(\epsilon)}$ solves the equation

$$\partial_t v_\epsilon = \Delta v_\epsilon - \sum_{j=1}^m \lambda_j^{(\epsilon)} \epsilon^{j-1} W_{2j+1,\mu_\epsilon}(v_\epsilon) - \lambda_0^{(\epsilon)} v_\epsilon - C_\epsilon v_\epsilon + \zeta_\epsilon \tag{5.8}$$

with initial data $\phi_0^{(\epsilon)}$, and

$$C_\epsilon = \sum_{k,\ell=1}^m \lambda_k^{(\epsilon)} \lambda_\ell^{(\epsilon)} \Big((2k+1)(2k) C^{(\epsilon)}_{2k-1,2\ell+1} + (2k+1)(2\ell+1) C^{(\epsilon)}_{2k,2\ell} \Big). \tag{5.9}$$

Comparing (5.1) and (5.8), we see that as long as we choose

$$\lambda_j^{(\epsilon)} = \widehat{a}_j(\theta) \quad (j \geq 1), \qquad \lambda_0^{(\epsilon)} = \epsilon^{-1} \widehat{a}_0(\theta) - C_\epsilon,$$

then the right hand sides of (5.1) and (5.8) are identical, and we have $u_\epsilon = v_\epsilon = \mathcal{R}^\epsilon \Phi^{(\epsilon)}$. Also, with the above choice of $\lambda_j^{(\epsilon)}$, it is straightforward to see that $\lambda_j^{(\epsilon)} \to \widehat{a}_j$ for each $j \geq 1$ as long as $\theta = \theta(\epsilon) \to 0$. As for $\lambda_0^{(\epsilon)}$, we see from (5.9), the convergence of $\lambda_1^{(\epsilon)}$ to $\widehat{a}_1$, and the behaviour of the constants $C^{(\epsilon)}_{k,\ell}$ as in (4.9) and (4.10) (and $C^{(\epsilon)}_{k,\ell}$ is uniformly bounded in ϵ if $(k,\ell) \neq (2,2)$) that we have $C_\epsilon = 9\widehat{a}_1^2 C_{\log} |\log \epsilon| + \mathcal{O}(1)$. Thus, if we take

$$\theta = \frac{9\widehat{a}_1^2 C_{\log}}{\widehat{a}_0'} \epsilon |\log \epsilon| + \mathcal{O}(\epsilon)$$

where $\widehat{a}_0'$ is the derivative of $\widehat{a}_0(\theta)$ at $\theta = 0$, and recall that $\widehat{a}_0 = \widehat{a}_0(0) = 0$, then the logarithmic divergences in $\lambda_0^{(\epsilon)}$ cancel out, and thus $\lambda_0^{(\epsilon)}$ converges to a finite limit. It then follows from Theorem 3.3 and Theorem 5.1 that $u_\epsilon = \mathcal{R}^\epsilon \Phi^{(\epsilon)}$ converges to the $\Phi^4_3(\widehat{a}_1)$ family of solutions. □

Remark 5.3 If ζ is not symmetric, then the object (4.20) (after multiplication of suitable ϵ's) with k even and ℓ odd will have a divergent 0-th chaos component of order $\epsilon^{-\frac{1}{2}}$. Thus, in order for the renormalised models to converge, one needs to further subtract a constant of the same order. However, such a direct subtraction cannot be attained by merely adjusting the value θ, and will change the assumption on the potential V. We expect that in such a case, similar to [19] where the authors consider Gaussian noise but asymmetric potential, one needs to re-center and rescale the process u (at a scale depending on θ, but in general smaller than ϵ^{-1}) to kill that divergence and obtain either Φ_3^3 or O.U. process in the limit.

Acknowledgements We thank Ajay Chandra for discussions on general moment bounds, and the referee for carefully reading the manuscript and providing suggestions on improving the presentation. WX acknowledges the Mathematical Sciences Research Institute in Berkeley, California for its warm hospitality, and the National Science Foundation for supporting his stay here.

References

1. Aizenman, M.: Geometric analysis of φ^4 fields and Ising models. I, II. Commun. Math. Phys. **86**(1), 1–48 (1982)
2. Anshelevich, M.: Appell polynomials and their relatives. Int. Math. Res. Not. **65**, 3469–3531 (2004). https://doi.org/10.1155/S107379280413345X
3. Appell, P.: Sur une classe de polynômes. Ann. Sci. École Norm. Sup. (2) **9**, 119–144 (1880)
4. Albeverio, S., Röckner, M.: Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. Probab. Theory Relat. Fields **89**(3), 347–386 (1991). https://doi.org/10.1007/BF01198791
5. Avram, F., Taqqu, M.S.: Noncentral limit theorems and Appell polynomials. Ann. Probab. **15**(2), 767–775 (1987)
6. Bauerschmidt, R., Brydges, D., Slade, G.: Scaling limits and critical behaviour of the 4-dimensional n-component $|\phi|^4$ spin model. J. Stat. Phys. **157**(4–5), 692–742 (2014)
7. Bertini, L., Presutti, E., Rüdiger, B., Saada, E.: Dynamical fluctuations at the critical point: convergence to a nonlinear stochastic PDE. Teor. Veroyatnostei ee Primeneniya **38**(4), 689–741 (1993)
8. Catellier, R., Chouk, K.: Paracontrolled Distributions and the 3-dimensional Stochastic Quantization Equation (2013). arXiv:1310.6869
9. Chandra, A., Shen, H.: Moment bounds for SPDEs with non-Gaussian fields and application to the Wong-Zakai problem. arXiv:1605.05683 (2016)
10. Da Prato, G., Debussche, A.: Strong solutions to the stochastic quantization equations. Ann. Probab. **31**(4), 1900–1916 (2003)
11. Fröhlich, J.: On the triviality of $\lambda\varphi_d^4$ theories and the approach to the critical point in $d>4$ dimensions. Nucl. Phys. B **200**(2), 281–296 (1982). https://doi.org/10.1016/0550-3213(82)90088-8
12. Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. Forum Math. Pi **3**(e6), 75 (2015). https://doi.org/10.1017/fmp.2015.2
13. Glimm, J., Jaffe, A.: Quantum physics: A functional integral point of view, 2nd edn. Springer-Verlag, New York (1987). https://doi.org/10.1007/978-1-4612-4728-9
14. Giacomin, G., Lebowitz, J. L., Presutti, E.: Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In: Stochastic partial differential equations: six perspectives, vol. 64 of Math. Surveys Monogr., 107–152. Amer. Math. Soc., Providence, RI, (1999)
15. Hairer, M.: A theory of regularity structures. Invent. Math. **198**(2), 269–504 (2014). https://doi.org/10.1007/s00222-014-0505-4. arXiv:1303.5113
16. Hairer, M., Matetski, K.: Discretisations of Rough Stochastic PDEs (2015). arXiv:1511.06937
17. Hairer, M., Quastel, J.: A class of growth models rescaling to KPZ. arXiv preprint arXiv:1512.07845 (2015)
18. Hairer, M., Shen, H.: A central limit theorem for the KPZ equation. ArXiv e-prints, to appear in Ann. Probab. (2015). arXiv:1507.01237
19. Hairer, M., Xu, W.: Large scale behaviour of 3D continuous phase coexistence models. Arxiv e-prints (2016). arXiv:1601.05138
20. Kupiainen, A. (2014) Renormalization group and stochastic PDEs. Ann. Henri Poincaré. https://doi.org/10.1007/s00023-015-0408-y

21. Lukkarinen, J., Marcozzi, M.: Wick polynomials and time-evolution of cumulants (2015). arXiv:1503.05851
22. Mourrat, J.-C., Weber, H.: Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2. ArXiv e-prints, to appear in Comm. Pure Appl. Math. (2014). arXiv:1410.1179
23. Mourrat, J.-C., Weber, H.: Global well-posedness of the dynamic Φ^4 model in the plane. ArXiv e-prints, to appear in Ann. Probab. (2015). arXiv:1501.06191
24. Mourrat, J.-C., Weber, H.: Global well-posedness of the dynamic Φ^4_3 model on the torus (2016). arXiv:1601.01234
25. Peccati, G., Taqqu, M.S.: *Wiener Chaos: Moments, Cumulants and Diagrams, Vol. 1 of Bocconi & Springer Series*. Springer, Milan; Bocconi University Press, Milan: A Survey with Computer Implementation. Supplementary material available online (2011). https://doi.org/10.1007/978-88-470-1679-8
26. Parisi, G., Wu, Y.-S.: Perturbation theory without gauge fixing. Sci. Sin. **4**, 483–496 (1981)
27. Shen, H., Weber, H.: Glauber dynamics of 2D Kac-Blume-Capel model and their stochastic PDE limits (2016). arXiv:1608.06556