JEMS

Yu Deng · Zaher Hani

Propagation of chaos and higher order statistics in wave kinetic theory

Abstract. The wave kinetic equation describes the effective dynamics of the second moments of the Fourier modes of the NLS solution at the *kinetic timescale*, and in the *kinetic limit* in which the size of the system diverges to infinity and the strength of the nonlinearity vanishes to zero, according to a specified *scaling law*. This equation was fully and rigorously derived from the nonlinear Schrödinger (NLS) equation in dimensions $d \geq 3$ in our earlier work.

Here, we investigate the behavior of the joint distribution of these Fourier modes and derive their effective limit dynamics at the kinetic timescale. In particular, we prove *propagation of chaos* in the wave setting: initially independent Fourier modes retain this independence in the kinetic limit. Such statements are central to the formal derivations of all kinetic theories, dating back to the work of Boltzmann (Stosszahlansatz). We obtain this by deriving the asymptotics of the higher Fourier moments, which are given by solutions of the wave kinetic heirarchy (WKH) with factorized initial data. As a byproduct, we also provide a rigorous justification of this hierarchy for general (not necessarily factorized) initial data.

We treat both Gaussian and non-Gaussian initial distributions. In the Gaussian setting, we prove *propagation of Gaussianity* as we show that the asymptotic distribution retains the Gaussianity of the initial data in the limit. In the non-Gaussian setting, we derive the limiting equations for the higher order moments, as well as for the density function (PDF) of the solution. Some of the results we prove were conjectured in the physics literature, others appear to be new. This gives a complete description of the statistics of the solutions in the kinetic limit.

Keywords: wave turbulence thoery, wave kinetic theory, propagation of chaos, wave kinetic heirarchy, Feynman diagrams, Stosszahlansatz.

1. Introduction

Propagation of chaos is a central theme in all kinetic theories in statistical physics. Roughly speaking, it states that for a microscopic system with many interacting objects

Yu Deng: Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA; yudeng@usc.edu

Zaher Hani: Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA; zhani@umich.edu

(particles or waves), two distinct objects should be statistically independent in the kinetic limit. Of course, this independence is not true before taking the limit, even if it is true at initial time, because naturally the dynamics produces correlations between the objects. Nonetheless, the fact that this independence is resurrected in the limit is a cornerstone of the whole kinetic description, in both particle and wave kinetic theories. In fact, almost all formal derivations of kinetic models, dating back to founding work of Boltzmann, *assume* propagation of chaos to hold in order to get a closed kinetic equation for the lowest non-trivial marginal or moment of the solution.

Mathematically speaking, propagation of chaos can be phrased in terms of the asymptotics of appropriate correlations or joint distributions of the solution. In wave kinetic theory, also called wave turbulence theory, these are given by the (second and higher order) moments of the Fourier modes of the solution to the dispersive equation that describes the microscopic system. If $u(t)$ is this solution, the second moment $\mathbb{E} |\hat{u}(t, k)|^2$ is the central quantity whose asymptotics, in the kinetic limit, is given by the *wave kinetic equation* (WKE), which acts as the wave analog of Boltzmann's equation. The formal derivations of this equation in the physics literature, dating back to the pioneering works of Peierls, Hasselman, and others [22,23,28–30,34], are based on the unjustified assumption of propagation of chaos, which effectively allows one to represent higher order mixed moments by products of second order ones, thus yielding a closed equation for the second moments.

A rigorous derivation of the WKE at the kinetic timescale, starting from the nonlinear Schrödinger (NLS) equation with random initial data, has been given in our recent work [12]. This is the first result of its kind for any dispersive system (we will review some of the literature below). The derivation is done via a delicate analysis of the iterates of the NLS equation and their second order correlations, which are represented by ternary trees (and couples of such trees) often called Feynman diagrams. The analysis of such diagrams involves (a) identifying the leading order diagrams called *regular couples*, (b) proving that all remaining diagrams lead to negligible contributions, and (c) controlling the remainder term in the iteration. This outline is rather simplistic; in reality there are other almost-leading diagrams whose contributions have to be analyzed separately. Moreover, the problem of estimating the diagrams is probabilistically critical in the sense of [14], which is added to the factorial growth of the number of diagrams, to make the execution of this outline far from trivial. We will review some elements of that proof in Section 3 below, and also refer the reader to [12, Section 3] for a more detailed exposition.

In particular, the proof in [12] does not require establishing propagation of chaos for the higher moments of the solution in order to obtain the effective equation for the second moment, in sharp contrast with the earlier works that make use of the BBGKY and other similar hierarchies. This brings us to the main goal of this manuscript, which is to establish propagation of chaos and the corresponding (wave kinetic) hierarchy *a posteriori* relying on the analysis introduced in [12]. Highly interesting results and unique features will appear, for the higher order statistics, depending on the initial distribution of the data, as we discuss both Gaussian and non-Gaussian initial distributions (for concreteness, only the Gaussian case was treated in [12]). In the former case, we will prove propagation of

Gaussianity, which states that the asymptotic distribution of the modes remains Gaussian as it is initially. In the latter case, we will derive the limiting equations for the probability density function. We remark that this gives a *complete description of the statistics* of the solutions in the kinetic limit, for both Gaussian and non-Gaussian initial distributions.

1.1. The kinetic setup

To state our results more precisely, let us first recall the wave kinetic setup starting with the microscopic system given by the cubic nonlinear Schrödinger equation. In dimension $d \geq 3$, we set this equation on a large torus of size L. The torus may be rational or irrational, which can always be rescaled to the square torus $\mathbb{T}_L^d = [0, L]^d$ but with the twisted Laplacian

$$\Delta_\beta = (2\pi)^{-1}(\beta^1 \partial_1^2 + \cdots + \beta^d \partial_d^2),$$

where $\beta = (\beta^1, \ldots, \beta^d) \in (\mathbb{R}^+)^d$ determines the aspect ratios of the torus. Consider the cubic NLS equation

$$\begin{cases} (i\partial_t - \Delta_\beta)u + \lambda^2 |u|^2 u = 0, \\ u(0, x) = u_{\text{in}}(x), \end{cases} \tag{NLS}$$

with random initial data $u(0) = u_{\text{in}}$, and

$$u_{\text{in}}(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \widehat{u_{\text{in}}}(k) e^{2\pi i k \cdot x}, \quad \widehat{u_{\text{in}}}(k) = \sqrt{n_{\text{in}}(k)} \eta_k(\omega). \tag{DAT}$$

Here $\mathbb{Z}_L^d = (L^{-1}\mathbb{Z})^d$, n_{in} is a nonnegative Schwartz function on $\mathbb{R}^d$, and $\eta_k(\omega)$ are i.i.d. random variables satisfying

$$\mathbb{E}\eta_k = 0, \qquad \mathbb{E}|\eta_k|^2 = 1.$$

This distribution of initial data will be called Gaussian if the law of each η_k is a standard complex Gaussian, and will be called non-Gaussian otherwise. Define

$$\alpha = \lambda^2 L^{-d}, \quad T_{\text{kin}} = \frac{1}{2\alpha^2} = \frac{1}{2} \cdot \frac{L^{2d}}{\lambda^4}.$$

The parameter α stands for the strength of the nonlinearity[1] and T_{kin} is the kinetic timescale at which the NLS dynamics is approximated by that of the WKE. The *kinetic limit* is taken by letting $L \to \infty$ (large box limit) and $\alpha \to 0$ (weak nonlinearity limit), according to some scaling law that specifies the relative rate of those two limits.

---

[1]With overwhelming probability for large L, it can be shown that the size of the nonlinearity (say in L^2 norm) is comparable to α. This follows from the probabilistic analysis performed in [12], but can also be seen by simple heuristic considerations (cf. the introduction of [12]). We also note that it is common in the physics literature to use a different parametrization of the Fourier series in (DAT) by replacing the L^{-d} factor in (DAT) with $L^{-d/2}$, in which case α would be defined as λ^2 and $T_{\text{kin}} = 1/(2\lambda^4)$.

The general form of a scaling law is $\alpha = L^{-\gamma}$ where $0 \leq \gamma \leq \infty$, with the understanding that if $\gamma = 0$ then the $\alpha \to 0$ limit is taken after the $L \to \infty$ limit, and vice versa for $\gamma = \infty$. As explained in the introduction of [12], not all scaling laws are admissible for the kinetic theory, and the admissibility range can depend on the shape of the torus (i.e. the Diophantine properties of β). Indeed, without any Diophantine conditions on β, the admissible range of γ is $0 \leq \gamma \leq 1$, and one can show (e.g. [9]) that if $\gamma > 1$, then the kinetic description does not hold, for example when $\beta = (1, \ldots, 1)$. Imposing generic Diophantine conditions on β by removing a set of "bad" vectors of zero Lebesgue measure, widens the admissible range of γ to $0 \leq \gamma \leq d/2$.

In [12], we treated scaling laws of the form $\alpha = L^{-\gamma}$ with $\gamma \leq 1$ but sufficiently close to 1. When $\gamma < 1$ no requirements on the shape of the torus are needed, but for the endpoint $\gamma = 1$, the torus needs to have generic shape, i.e. β should belong to the complement of a fixed Lebesgue null set $\mathfrak{Z}$ defined by a set of explicit Diophantine conditions (Lemma 2.1). We remark that the result in [12] has been extended to the full range $\gamma \in (0, 1)$ in [13]. In the current paper, for the sake of concreteness, we will stick to the setup in [12] and adopt the scaling law $\alpha = L^{-1}$, with the understanding that the result also applies to the same range of γ as in [13] and without any Diophantine condition on β. Hence, throughout the proof we will assume β is generic in the above sense, $\lambda = L^{(d-1)/2}$, and $T_{\text{kin}} = L^2/2$.

For $0 < \delta \ll 1$ depending on n_{in}, define the solution $n = n(t, k)$, for $t \in [0, \delta]$ and $k \in \mathbb{R}^d$, to the wave kinetic equation

$$\begin{cases} \partial_t n(t, k) = \mathcal{K}(n(t), n(t), n(t))(k), \\ n(0, k) = n_{\text{in}}(k), \end{cases} \tag{WKE}$$

with the nonlinearity

$$\mathcal{K}(\phi_1, \phi_2, \phi_3)(k)$$
$$= \int_{(\mathbb{R}^d)^3} \{\phi_1(k_1)\phi_2(k_2)\phi_3(k_3) - \phi_1(k)\phi_2(k_2)\phi_3(k_3) + \phi_1(k_1)\phi_2(k)\phi_3(k_3)$$
$$- \phi_1(k_1)\phi_2(k_2)\phi_3(k)\}$$
$$\times \delta(k_1 - k_2 + k_3 - k) \cdot \delta(|k_1|_\beta^2 - |k_2|_\beta^2 + |k_3|_\beta^2 - |k|_\beta^2) \, dk_1 \, dk_2 \, dk_3. \tag{KIN}$$

Here δ is the Dirac delta, and for $k = (k^1, \ldots, k^d)$ and $\ell = (\ell^1, \ldots, \ell^d)$ we denote

$$|k|_\beta^2 := \langle k, k \rangle_\beta, \quad \langle k, \ell \rangle_\beta := \beta^1 k^1 \ell^1 + \cdots + \beta^d k^d \ell^d.$$

The following theorem is the main result of [12], which describes the evolution of the variance $\mathbb{E}|\hat{u}(t, k)|^2$ in the limit. Here and below, the expectation $\mathbb{E}$ is always taken under the assumption that (NLS) has a smooth solution on $[0, \delta \cdot T_{\text{kin}}]$, which happens with overwhelming probability.

Theorem 1.1 ([12, Theorem 1.1]). *Fix $A \geq 40d$, $\beta \in (\mathbb{R}^+)^d \setminus \mathfrak{Z}$, and a function $n_{\text{in}} \geq 0$ such that*

$$\|n_{\text{in}}\|_{\mathfrak{S}^{40d}} := \max_{|\alpha|, |\beta| \leq 40d} \|k^\alpha \partial_k^\beta n_{\text{in}}\|_{L^2} \leq C_1 < \infty.$$

Assume the law of each η_k is Gaussian. Let δ be small enough depending on (A, β, C_1), and L be sufficiently large depending on δ. Set $\lambda = L^{(d-1)/2}$ so $\alpha = L^{-1}$ and $T_{\text{kin}} = L^2/2$. Then the equation (NLS), *with random initial data* (DAT), *has a smooth solution up to time*

$$T = \frac{\delta L^2}{2} = \delta \cdot T_{\text{kin}}$$

with probability $\geq 1 - L^{-A}$. *Moreover,*

$$\lim_{L \to \infty} \sup_{t \in [0,T]} \sup_{k \in \mathbb{Z}_L^d} \left| \mathbb{E} |\hat{u}(t,k)|^2 - n\left(\frac{t}{T_{\text{kin}}}, k\right) \right| = 0, \tag{1.1}$$

where $n(t,k)$ is the solution to (WKE).

Remark 1.2. Theorem 1.1 is stated in [12] for Schwartz n_{in}. A closer look at the proof shows that it remains true as long as $n_{\text{in}} \in \mathfrak{S}^{40d}$, and δ should only depend on the $\mathfrak{S}^{40d}$ norm of n_{in}; see [12, remarks after Theorem 1.1]. The same comment also applies to all the main results of the current paper.

1.2. Propagation of chaos: The Gaussian case

As mentioned above, the proof of Theorem 1.1 does not require obtaining asymptotics on the higher Fourier moments. Such information is provided in our first main result, which can be viewed as an extension of Theorem 1.1.

Theorem 1.3 (Propagation of chaos and Gaussianity). *Under the same assumptions as Theorem 1.1 above, fix a positive integer r and nonnegative integers $p_1, \ldots, p_r$ and $q_1, \ldots, q_r$. If $p_j \neq q_j$ for at least one j with $1 \leq j \leq r$, then*

$$\lim_{L \to \infty} \sup_{t \in [0,T]} \sup_{\substack{(k_1,\ldots,k_r) \in (\mathbb{Z}_L^d)^r \\ k_i \neq k_j \ (\forall i \neq j)}} \left| \mathbb{E}\left(\prod_{j=1}^{r} (\hat{u}(t,k_j))^{p_j} (\overline{\hat{u}(t,k_j)})^{q_j} \right) \right| = 0. \tag{1.2}$$

Here, as in Theorem 1.1, the expectation is taken only when (NLS) *has a smooth solution on $[0, T]$ where $T = \delta \cdot T_{\text{kin}}$ (which has probability $\geq 1 - L^{-A}$). If $p_j = q_j$ for each $1 \leq j \leq r$, then*

$$\lim_{L \to \infty} \sup_{t \in [0,T]} \sup_{\substack{(k_1,\ldots,k_r) \in (\mathbb{Z}_L^d)^r \\ k_i \neq k_j \ (\forall i \neq j)}} \left| \mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t,k_j)|^{2p_j} \right) - \prod_{j=1}^{r} (p_j)! \cdot n\left(\frac{t}{T_{\text{kin}}}, k_j\right)^{p_j} \right| = 0.$$
$$\tag{1.3}$$

A key feature of Theorem 1.3 is that, up to error terms that vanish as $L \to \infty$, we have

$$\mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t,k_j)|^{2p_j} \right) \approx \prod_{j=1}^{r} \mathbb{E} |\hat{u}(t,k_j)|^{2p_j}, \quad \text{and all other moments are} \approx 0. \tag{1.4}$$

This means that, for fixed t, the random variables $\hat{u}(t, k)$ for different k become independent in the limit (at least in terms of the marginal distributions of any finitely many of them), which justifies rigorously the *propagation of chaos* assumption in the literature, as described at the beginning of this paper.

Note that these coefficients cannot be independent without taking limits, because correlations will always be produced by the nonlinear interactions in the NLS equation. Nonetheless, this independence reappears in the kinetic limit as $L \to \infty$ and $\alpha \to 0$, for the same subtle and deep reason that makes the kinetic approximation in (1.1) hold. Namely, the only nonvanishing interactions contributing to the expectations in (1.1)–(1.3) are those obtained by concatenating blocks of basic interactions called $(1, 1)$-mini couples and mini trees (see Figures 1–3), thus forming what we call *regular couples* (for second moments) or *regular multi-couples* (for higher order moments, see Section 1.6). Such interactions can only be built if $p_j = q_j$ in the notation of Theorem 1.3; moreover, in the higher order case, the associated structure actually decouples into second order structures, hence (1.4) naturally occurs. The same reasoning also holds in the non-Gaussian case below (Section 1.3), for which (1.4) remains valid.

In addition, in this Gaussian setting we have

$$\mathbb{E}\,|\hat{u}(t, k)|^{2p} \approx p! \cdot n\left(\frac{t}{T_{\mathrm{kin}}}, k\right)^p, \tag{1.5}$$

which means that the law of $\hat{u}(t, k)$ in the limit is Gaussian with variance $n(t/T_{\mathrm{kin}}, k)$ as long as the initial state at $t = 0$ is Gaussian. This has been conjectured in the physics literature under the name of *propagation of Gaussianity* (see also the discussion following Theorem 1.5).

1.3. The non-Gaussian case

Highly interesting results appear in the non-Gaussian case, where unlike Theorems 1.1 and 1.3, the law of η_k may not be Gaussian. While the second moments still follow the WKE in this setting, the non-Gaussianity of the initial law starts to manifest itself at the higher (≥ 4) order moments and statistics. We will assume the law of η_k is rotationally symmetric,[2] and has exponential tails. Then we have the following modification to Theorem 1.3:

Theorem 1.4 (Evolution of moments). *Suppose the i.i.d. random variables $\{\eta_k(\omega)\}$ have a law that is rotation symmetric, and satisfies*

$$\mu_r := \mathbb{E}\,|\eta_k|^{2r} \leq (C_0 r)!, \quad \mu_1 = 1.$$

---

[2]Though rotation symmetry seems to be always assumed in physics literature, it would be interesting to see what happens without this assumption, in particular if (1.10) remains true. Here the loss of gauge invariance may lead to additional contributions, but probably they will be error terms in the end.

for some constant C_0 (this is equivalent to $\mathbb{E}(e^{|\eta_k|^\beta}) < \infty$ for small $\beta > 0$). Then the same limits in (1.1) and (1.2) remain true. Moreover, instead of (1.3), we have

$$\lim_{L\to\infty} \sup_{t\in[0,T]} \sup_{\substack{(k_1,\ldots,k_r)\in(\mathbb{Z}_L^d)^r \\ k_i\neq k_j \, (\forall i\neq j)}} \left| \mathbb{E}\Big(\prod_{j=1}^r |\hat{u}(t,k_j)|^{2p_j}\Big) - \prod_{j=1}^r \mu_{p_j}\Big(\frac{t}{T_{\text{kin}}}, k_j\Big) \right| = 0. \quad (1.6)$$

Here the functions $\mu_r(t,k)$ is defined as follows: recall $n(t,k)$ is the solution to (WKE). Let $n_0(t,k)$ be the solution to the equation

$$\begin{cases} \partial_t n_0(t,k) = \mathcal{K}_0(t,k), \\ n(0,k) = n_{\text{in}}(k), \end{cases} \quad \text{(WKE-0)}$$

where

$$\mathcal{K}_0(t,k) = \int_{(\mathbb{R}^d)^3} n_0(t,k)\{n(t,k_1)n(t,k_3) - n(t,k_2)n(t,k_3) - n(t,k_1)n(t,k_2)\}$$
$$\times \delta(k_1 - k_2 + k_3 - k) \cdot \delta(|k_1|_\beta^2 - |k_2|_\beta^2 + |k_3|_\beta^2 - |k|_\beta^2)\, dk_1\, dk_2\, dk_3. \quad \text{(KIN-0)}$$

Define also $n_+(t,k) = n(t,k) - n_0(t,k)$. Then

$$\mu_q(t,k) = \sum_{p=0}^q \binom{q}{p}^2 (q-p)!\mu_p \cdot (n_0(t,k))^p (n_+(t,k))^{q-p}. \quad (1.7)$$

Note that if $\{\eta_k\}$ is Gaussian, then $\mu_p = p!$, so (1.7) yields $\mu_q(t,k) = q!(n(t,k))^q$, and we recover Theorem 1.3. Similarly, for $q = 1$ we have $\mu_1(t,k) = n(t,k)$, so Theorem 1.1 remains true in the non-Gaussian case.

Note that in Theorem 1.4 we still have (1.4), thus propagation of chaos remains true in the non-Gaussian case. In addition, instead of (1.5) we have

$$\mathbb{E}\,|\hat{u}(t,k)|^{2p} \approx \mu_p(t/T_{\text{kin}},k)$$

where $\mu_p(t,k)$ is defined as in (1.7). As far as we know, these expressions for higher order moments are new.

We remark that Theorems 1.3 and 1.4 actually hold for moments whose degree (given by $\sum_{j=1}^r (p_j + q_j)$ in the notation of (1.2)) may diverge as $L \to \infty$. Indeed, we will see in the proof that this degree can be taken as high as $\log L$ (for Theorem 1.3) or $\frac{\log L}{(\log\log L)^2}$ (for Theorem 1.4).

Under slightly stronger assumptions, Theorem 1.4 allows us to describe the evolution of the law of individual Fourier modes in terms of the density function, which then provides a full description of the statistics of the NLS solution in the limit. This is summarized in our next theorem below.

Theorem 1.5 (Evolution of density). *In Theorem 1.4, assume further that $\mu_r \leq C^r(2r)!$ for some constant C (this is equivalent to $\mathbb{E}(e^{\beta|\eta_k|}) < \infty$ for small $\beta > 0$). Recall the*

solution $n = n(t, k)$ to (WKE), and define

$$\sigma_k(t) = \int_{(\mathbb{R}^d)^3} n(t, k_1) n(t, k_2) n(t, k_3) \delta(k_1 - k_2 + k_3 - k)$$

$$\times \delta(|k_1|_\beta^2 - |k_2|_\beta^2 + |k_3|_\beta^2 - |k|_\beta^2)\, dk_1\, dk_2\, dk_3, \qquad (1.8)$$

$$\gamma_k(t) = \int_{(\mathbb{R}^d)^3} \{n(t, k_1) n(t, k_3) - n(t, k_2) n(t, k_3) - n(t, k_1) n(t, k_2)\}$$

$$\times \delta(k_1 - k_2 + k_3 - k) \delta(|k_1|_\beta^2 - |k_2|_\beta^2 + |k_3|_\beta^2 - |k|_\beta^2)\, dk_1\, dk_2\, dk_3. \qquad (1.9)$$

Let the density function of each $\eta_k(\omega)$ be $\rho_* = \rho_*(v)$, where $v \in \mathbb{C}$ is also viewed as an $\mathbb{R}^2$ vector; assume ρ_* is a radial function. Let $\rho_k = \rho_k(t, v)$ be the solution to the linear equation

$$\begin{cases} \partial_t \rho_k = \dfrac{\sigma_k(t)}{4} \Delta \rho_k - \dfrac{\gamma_k(t)}{2} \nabla \cdot (v \rho_k), \\[2mm] \rho_k(0) = \dfrac{1}{n_{\mathrm{in}}(k)} \rho_* \left(\dfrac{v}{\sqrt{n_{\mathrm{in}}(k)}} \right). \end{cases} \qquad (1.10)$$

Clearly each ρ_k is also radial. Fix $t \in [0, \delta]$, a positive integer r and distinct vectors $k_j \in \mathbb{R}^d$ ($1 \le j \le r$). Let $k_j^{(L)} \in \mathbb{Z}_L^d$ ($1 \le j \le r$) be such that $k_j^{(L)} \to k_j$ ($1 \le j \le r$) as $L \to \infty$. Then the random variables

$$(\hat{u}(t \cdot T_{\mathrm{kin}}, k_1^{(L)}), \ldots, \hat{u}(t \cdot T_{\mathrm{kin}}, k_r^{(L)})) \qquad (1.11)$$

converge in law, as $L \to \infty$, to the random variable with density function

$$\rho_{k_1}(t, v_1) \cdot \rho_{k_2}(t, v_2) \cdots \rho_{k_r}(t, v_r). \qquad (1.12)$$

The factorization structure in (1.12) is a consequence of propagation of chaos, which has been established in Theorem 1.4; thus the main feature of Theorem 1.5 is the evolution of the individual density (1.10). It appears that this equation has only been discovered fairly recently in the physics literature (see [8, 25], and [28, Section 6.6]).

Note that in the Gaussian case (Theorem 1.3) we have $\rho_*(v) = \pi^{-1} e^{-|v|^2}$. Then the solution to (1.10) equals

$$\rho_k(t, v) = (\pi n(t, k))^{-1} e^{-|v|^2 / n(t, k)},$$

so by (1.12), the limit distribution is given by independent Gaussians with variance $n(t, k)$, which provides another manifestation of the propagation of Gaussianity. Other solutions to (1.10) can be obtained and analyzed using the method of characteristics in Fourier space (see [7]).

1.4. The wave kinetic hierarchy

By taking $p_j = 1$ in (1.3) or (1.6) we obtain the limits

$$n_r(t, k_1, \ldots, k_r) := \lim_{L \to \infty} \mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t, k_j)|^2 \right). \qquad (1.13)$$

These limit quantities are conjectured to solve an infinite hierarchy of equations called the *wave kinetic hierarchy* (WKH), which is a linear system for symmetric functions $n_r = n_r(t, k_1, \ldots, k_r)$, and has the form

$$\partial_t n_r(t, k_1, \ldots, k_r)$$

$$= \sum_{j=1}^{r} \int_{(\mathbb{R}^d)^3} \boldsymbol{\delta}(\ell_1 - \ell_2 + \ell_3 - k_j) \cdot \boldsymbol{\delta}(|\ell_1|_\beta^2 - |\ell_2|_\beta^2 + |\ell_3|_\beta^2 - |k_j|_\beta^2) \, d\ell_1 \, d\ell_2 \, d\ell_3$$

$$\times \big[n_{r+2}(t, k_1, \ldots, k_{j-1}, \ell_1, \ell_2, \ell_3, k_{j+1}, \ldots, k_r)$$

$$+ n_{r+2}(t, k_1, \ldots, k_{j-1}, \ell_1, k_j, \ell_3, k_{j+1}, \ldots, k_r)$$

$$- n_{r+2}(t, k_1, \ldots, k_{j-1}, k_j, \ell_2, \ell_3, k_{j+1}, \ldots, k_r)$$

$$- n_{r+2}(t, k_1, \ldots, k_{j-1}, \ell_1, \ell_2, k_j, k_{j+1}, \ldots, k_r) \big]. \quad \text{(WKH)}$$

This hierarchy is the analog of Boltzmann and Gross–Pitaevskiĭ hierarchies, and is formally derived in recent works such as Chibarro et al. [5, 6], Eyink–Shi [18] and Newell–Nazarenko–Biven [29], though it also follows from much earlier works including the foundational work of Peierls (see [3, 28, 30]).

The key property of (WKH) is *factorizability*: factorized initial data of the form $(n_r)_{in}(k_1, \ldots, k_r) = \prod_{j=1}^{r} n_{in}(k_j)$ leads to factorized solutions of the form $n_r(t, k_1, \ldots, k_r) = \prod_{j=1}^{r} n(t, k_j)$ where $n(t, k)$ solves (WKE) with initial data n_{in}. This follows from direct calculations together with a suitable uniqueness theorem, which is recently proved by Rosenzweig and Staffilani [31].

In the above sense, we can view (WKH) as a generalization of (WKE) that allows for dependent Fourier modes. Indeed, suppose the initial data u_{in} of (NLS) is given by (DAT) with $\widehat{u_{in}}(k)$ being independent for different k. Then Theorem 1.4 implies that the limit (1.13) will be a factorized solution to (WKH) with factorized initial data, which is in fact the tensor product of the solution to (WKE). However, if u_{in} does not have independent Fourier modes, then the initial data $(n_r)_{in}(k_1, \ldots, k_r)$, i.e. (1.13) at time 0, will not have factorized form, in which case (1.13) at time t is conjectured to be a more general solution to (WKH).

Such scenario may arise, as discussed in [31, Section 1.3], if one considers a hybrid, or "twice randomized data" problem of (NLS) as follows: Instead of taking n_{in} deterministic in (DAT), we choose it randomly according to a probability measure ζ defined on the space of all nonnegative functions n_{in}, in such a way that the new random function n_{in} is *independent* of the pre-fixed i.i.d. random variables $\{\eta_k\}$. When η_k are random phases ($\eta_k(\omega) = e^{i\theta_k(\omega)}$ with θ_k uniformly distributed on the circle), this process of randomization is referred to as the "Random Phase and Amplitude" assumption in the wave turbulence theory literature, where in this general setup different amplitudes are not necessarily independent.

In other words, we are choosing a random initial data whose law of distribution (as a probability measure) is given by a suitable average of those specific probability measures which are laws of distribution of random data of the form (DAT), i.e. having independent Fourier coefficients. This averaging is achieved by first generating a random nonnega-

tive function n_{in} according to the probability measure ζ on the space of all nonnegative functions, and then selecting the random initial data as (DAT) with some pre-fixed i.i.d. random variables $\{\eta_k\}$. Since independent Fourier modes in (DAT) correspond to factorized solutions to (WKH), we know, using also the linearity of (WKH), that the above process will result in a solution to (WKH) which is an average of certain factorized solutions. These are referred to as "super-statistical solutions" in Eyink–Shi [18] and may provide a possible explanation of intermittency in wave turbulence.

Just like (WKE), the rigorous derivation of (WKH) has been an outstanding open problem. In fact, these two problems are closely related; as mentioned at the beginning of this paper, there are many earlier works on similar problems that first derive the corresponding hierarchies and then restrict to factorized solutions to obtain the kinetic equations. In the wave turbulence context, such an approach is theoretically possible but has not yet been successful. Instead, we are following the *exactly opposite* route: we first derive the kinetic equation (WKE) in [12], then apply the same techniques to derive the hierarchy (WKH) *a posterori*, in the current paper. So our last main result is the rigorous derivation of (WKH) for general nonfactorized initial data, which we state as follows.

Theorem 1.6 (Derivation of (WKH)). *Fix a positive number $\mathfrak{X} > 0$ and a sequence of i.i.d. random variables $\{\eta_k\}$ as in Section 1.1 that satisfy the requirements of Theorem 1.4. Suppose $(n_r)_{\text{in}} = (n_r)_{\text{in}}(k_1, \ldots, k_r)$ are nonnegative symmetric functions of $k_j \in \mathbb{R}^d$ $(1 \leq j \leq r)$ such that*

$$\|(n_r)_{\text{in}}\|_{\mathfrak{G}^{40d;r}} := \sup_{|\alpha_j|,|\beta_j| \leq 40d} \|k_1^{\alpha_1} \cdots k_r^{\alpha_r} \partial_{k_1}^{\beta_1} \cdots \partial_{k_r}^{\beta_r} (n_r)_{\text{in}}(k_1, \ldots, k_r)\|_{L^2} \leq C_1^r \quad (1.14)$$

for some large constant C_1 (note $C_1 \gtrsim \mathfrak{X}$). We say $(n_r)_{\text{in}}$ is admissible *if for any $r \geq 2$ we have*

$$\int_{\mathbb{R}^d} (n_r)_{\text{in}}(k_1, \ldots, k_r) \, dk_r = \mathfrak{X} \cdot (n_{r-1})_{\text{in}}(k_1, \ldots, k_{r-1}), \quad \int_{\mathbb{R}^d} (n_1)_{\text{in}}(k_1) \, dk_1 = \mathfrak{X}. \quad (1.15)$$

Consider a probability measure ζ on the set $\mathcal{A}$ of nonnegative functions $m = m(k)$ on $\mathbb{R}^d$, which is defined by

$$\mathcal{A} := \left\{ m : \mathbb{R}^d \to \mathbb{R}_{\geq 0} : \|m\|_{\mathfrak{G}^{40d}} \leq C_1, \int_{\mathbb{R}^d} m(k) \, dk = \mathfrak{X} \right\}. \quad (1.16)$$

For this ζ, consider the hybrid initial data u_{in} which is given by (DAT), except that n_{in} should be replaced by m, which is another random variable with values in $\mathcal{A}$, such that m is independent with all the η_k and the law of m is given by ζ. We say $(n_r)_{\text{in}}$ is hybrid *if there exists a ζ such that for the above choice of u_{in},*

$$\mathbb{E}\left(\prod_{j=1}^{r} |\widehat{u_{\text{in}}}(k_j)|^2 \right) = (n_r)_{\text{in}}(k_1, \ldots, k_r) \quad (1.17)$$

for any L and any distinct $k_j \in \mathbb{Z}_L^d$ $(1 \leq j \leq r)$.

Let $T = \delta \cdot T_{\text{kin}}$ where δ is as in Theorem 1.1 (except C_1 is now defined by (1.14)); the other parameters are as in Theorem 1.1. Then we have the following:

(1) *The sequence $(n_r)_{\text{in}}$ is hybrid if and only if it is admissible; in this case the measure ζ is unique.*

(2) *Assume $(n_r)_{\text{in}}$ is admissible. Then with the hybrid initial data defined above, the equation (NLS) has a smooth solution up to time T with probability $\geq 1 - L^{-A}$. Moreover, for any fixed r we have*

$$\lim_{L \to \infty} \sup_{t \in [0,T]} \sup_{\substack{(k_1,\ldots,k_r) \in (\mathbb{Z}_L^d)^r \\ k_i \neq k_j \ (\forall i \neq j)}} \left| \mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t, k_j)|^2 \right) - n_r\left(\frac{t}{T_{\text{kin}}}, k_1, \ldots, k_r \right) \right| = 0, \quad (1.18)$$

where $n_r(t, k_1, \ldots, k_r)$ is the unique solution to (WKH) constructed in [31] with initial data $(n_r)_{\text{in}}$. For any $0 \leq t \leq \delta$, this solution $(n_r)(t)$ is admissible in the sense of (1.15) for the same $\mathfrak{X}$.

We make two remarks regarding Theorem 1.6. First, the $\mathfrak{S}^{40d;r}$ norms defined in (1.14) are much stronger than the $\mathfrak{L}_{s,\epsilon}^{\infty}$ norms defined in [31], because of the strong $\mathfrak{S}^{40d}$ norm used in Theorem 1.1. It may be possible to relax this regularity assumption to match [31], but this requires refining the proof of Theorem 1.1 (and Theorems 1.3–1.5), which we are not doing here.

Second, the admissibility requirement (1.15) seems natural in view of the conclusion (1): anything that actually arises from these hybrid initial data must be admissible. Nonadmissible solutions to (WKH) do exist, but they are probably not physically meaningful, as pointed out in [31].

1.5. Background literature

The proof of Theorems 1.3–1.6 are based on the framework introduced in [12] to prove Theorem 1.1. That work comes as a culmination of an extensive research effort over the past years to provide a rigorous justification of the wave kinetic equation starting from the nonlinear dispersive PDEs as first principle [4, 9–11, 15–17, 19, 27]. This is Hilbert's sixth problem for *waves*; its *particle* analog is the rigorous derivation of the Boltzmann equation from Newtonian mechanics (see [2, 20, 21, 26] and references therein). We refer the reader to the introduction of [12] for a discussion of the developments leading up to it.

We should remark on the progress that has happened since the submission of [12]. First, we mention the work of Staffilani and Tran [33], who consider a high (≥ 14) dimensional discrete KdV-type equation, with a Stratonovich-type stochastic multiplicative noise, which has the effect of regularly randomizing the phases of the Fourier modes. In the presence of this noise, the authors derive the associated kinetic equation at the kinetic timescale T_{kin} and in the scaling law $\alpha = L^{-0}$. They also have a conditional result in the absence of noise, which assumes that some a priori estimates hold for the solution, and they verify that these conditions are met for some more restrictive sets of initial data.

Another work in this direction is due to Ampatzoglou–Collot–Germain [1] which considers the problem of deriving the WKE in an inhomogeneous setting. The authors derive this equation from a quadratic NLS-type equation for short (asymptotically vanishing) timescales, which, similar to [11], is a subcritical version of the critical setting considered here and in [12].

Note that the works [4, 9–12, 15–17, 27] concern cubic nonlinearities or 4-wave interactions, while the works [1,19,33] concern quadratic nonlinearities or 3-wave interactions. Both models represent a lot of important physical scenarios. Although the cubic case is considered in the current paper and in [12], we believe that the quadratic case can be treated in the same way without much difference in strategy (as exhibited by [1]).

1.6. Idea of the proof

Before discussing the main ideas, we first review the proof of Theorem 1.1 in [12]. The basic strategy is to perform a high order expansion of the NLS solution in Fourier space as

$$\widehat{u}(t,k) = \sum_{n=0}^{N} \mathcal{J}_n(t,k) + \mathcal{R}_N(t,k), \quad k \in \mathbb{Z}_L^d. \tag{1.19}$$

Here, N is the order of the expansion which diverges appropriately with the size L of the domain, $\mathcal{J}_n$ is the n-th Picard iterate, and $\mathcal{R}_N$ is the remainder. The iterates $\mathcal{J}_n$ can be written as the sum of $\mathcal{J}_{\mathcal{T}}$, where $\mathcal{T}$ runs over all ternary trees that have n branches; these are often called Feynman diagrams. To derive (WKE) in [12], one has to compute the asymptotics of the second moments $\mathbb{E} \, |\widehat{u}(t,k)|^2$ which leads to the analysis of the correlations $\mathbb{E}(\mathcal{J}_{\mathcal{T}_1} \overline{\mathcal{J}_{\mathcal{T}_2}})$ for trees $\mathcal{T}_1$ and $\mathcal{T}_2$ of at most N branches. These expressions naturally lead to the notion of *couples* which consist of *two* ternary trees whose leaves are paired to each other. The key observation is that the leading couples in the expansion take a very special form, which we call *regular couples*, namely they are obtained by appropriately concatenating $(1, 1)$-mini couples and mini trees (see Figures 1–3). The proof in [12], as described before, then reduces to (a) establishing the precise asymptotics of the regular couples, which is made possible by their precise, albeit highly complex, structure, (b) showing that the remaining couples are of lower order, which constitutes the heart of the proof, and (c) showing the remainder $\mathcal{R}_N$ is also of lower order.

Now, in Theorem 1.3, we are interested in the higher order moments of the solutions, where the order R can be arbitrarily large (or even grow to infinity with L). If we perform the same expansion (1.19), then we need to consider expressions of the form

$$\mathbb{E}\big(\mathcal{J}_{\mathcal{T}_1}(t,k_1)^{\pm} \ldots \mathcal{J}_{\mathcal{T}_R}(t,k_R)^{\pm}\big)$$

where, as usual, a minus superscript denotes complex conjugation. This leads to the key new concept in the current paper, which we call *gardens*,[3] that are formed by R trees whose leaves are paired to each other.

---

[3]This name is partly inspired by the song *Spiritual Garden* of Yukari Tamura (2005).

In the Gaussian setting of Theorem 1.3, gardens are the only new structures that emerge. Since R can be arbitrarily large and may even grow to infinity with L, the analysis of gardens of R trees will be a lot more complicated than that of couples of two trees. However, the methodology introduced in [12], originally designed to treat couples, is in fact so robust that it can be extended to gardens—even for very large R—with some additional twists. Indeed, the leading contributions here come from those gardens that are formed by putting together $R/2$ couples (we call them *multi-couples*), which can be analyzed using the results of [12]. In particular, as shown in [12], only the *regular* multi-couples, where each of the $R/2$ couples is regular, provide the top order contributions; these can be explicitly calculated as in [12] to match the desired right hand side expressions, and the rest is of lower order.

As for the gardens that are not multi-couples, we apply the procedure of [12] (which are defined for couples but can be easily generalized to gardens) and conclude that they are of lower order (Proposition 4.7). A few technical differences occur here (such as in combinatorics, cf. Proposition 6.4 and [12, Proposition 9.6]), but the most important one, which is also the reason why these terms are of lower order, comes from the structure of the molecules (see Section 6) associated with such gardens. This is stated in Proposition 6.3 (for comparison, we have $\chi = m$ instead of $\chi \leq m - R/2$ for multi-couples), which can be used to establish a power gain in the counting estimates (Proposition 6.8, note the $m - R/2$ in the exponent), and subsequently the lower order bounds.

In the non-Gaussian setting (Theorem 1.4), we need to introduce even more general structures. In fact, gardens appear from dividing the leaves of the R trees as above into two-leaf pairs. In the Gaussian case, due to Isserlis' theorem, only expressions associated with gardens need to be considered; in the non-Gaussian case, we have a substitute of Isserlis' theorem (Lemma 9.1), which is reminiscent of the cumulant expansions of the moments of random variables, but with the important quantitative estimates included. This leads to the notion of *over-gardens* which are basically the same as gardens but allow pairings of more than two leaves. Again, in this setting, we identify the leading over-gardens (called regular ones) and prove that the complementary set is of lower order. It is here that the non-Gaussianity starts to manifest itself, as regular over-gardens contribute to the leading terms in addition to regular gardens, which explains the difference between (1.3) and (1.6).

In all the proofs above, as well as in [12], the leading structures (regular couples, multi-couples and over-gardens) are still highly complex objects, whose number grows exponentially (rather than factorially) in their size. However, their redeeming feature is that one can write down *exact* expressions for them in the kinetic limit which allows one to match their contribution, order by order, with the solutions of the kinetic equations that appear in (WKE), (WKE-0), or (WKH).

Finally, Theorem 1.5 is a direct consequence of (1.3) and (1.7), and uniqueness of the moment problem in this setting (i.e. the moments uniquely define the law, see Lemma 9.6) and Theorem 1.6 basically follow from averaging the results of Theorem 1.3 in different scenarios, and applying the Hewitt–Savage theorem (see Lemma 10.1) to represent arbitrary densities by tensor products.

We remark that the proof in this paper relies heavily on the notions and framework introduced in [12]. On the other hand, despite a few places where we briefly go over the results and proofs of [12], the majority of this paper is devoted to the *new* components needed in the higher order setting. In particular, the gardens we introduce are fundamental objects with important new features (such as Proposition 6.3), which will play significant roles in future studies of wave turbulence.

1.7. *Organization of the paper*

The paper is organized as follows: In Section 2 we review the setup and present some reductions of the problem. In Section 3, we review the argument in [12] and the needed results from there. In Section 4, we introduce the notion of gardens, their elementary combinatorial properties, and state the needed estimates to prove Theorem 1.3. These estimates are then proved in Sections 5–8. In Section 9 we deal with the non-Gaussian case and prove Theorems 1.4 and 1.5, and in Section 10 we prove Theorem 1.6.

2. Preliminary reductions

2.1. *Reduction of* (NLS)

As in [12], we make the following reductions. Suppose u is a solution to (NLS), and define $a = a_k(t)$ such that

$$a_k(t) = e^{-\delta \pi i L^2 |k|_\beta^2 t} \cdot e^{-2i\lambda^2 M \delta T_{\text{kin}} t} \cdot \widehat{u}(\delta T_{\text{kin}} \cdot t, k), \tag{2.1}$$

where M is the conserved mass of u. Then it solves the equation

$$\begin{cases} \partial_t a_k = \mathcal{C}_+(a, \overline{a}, a)_k(t), \\ a_k(0) = (a_k)_{\text{in}} = \sqrt{n_{\text{in}}(k)} \eta_k(\omega), \end{cases} \tag{2.2}$$

with the nonlinearity

$$\mathcal{C}_+(f, g, h)_k(t)$$
$$:= \frac{\delta}{2L^{d-1}} \cdot i \sum_{k_1 - k_2 + k_3 = k} \epsilon_{k_1 k_2 k_3} e^{\delta \pi i L^2 \Omega(k_1, k_2, k_3, k) t} f_{k_1}(t) g_{k_2}(t) h_{k_3}(t). \tag{2.3}$$

In (2.3) and below, the summation is taken over $(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3$, and

$$\epsilon_{k_1 k_2 k_3} = \begin{cases} 1 & \text{if } k_2 \notin \{k_1, k_3\}, \\ -1 & \text{if } k_1 = k_2 = k_3, \\ 0 & \text{otherwise}, \end{cases} \tag{2.4}$$

and the resonance factor is

$$\Omega = \Omega(k_1, k_2, k_3, k) := |k_1|_\beta^2 - |k_2|_\beta^2 + |k_3|_\beta^2 - |k|_\beta^2 = 2\langle k_1 - k, k - k_3 \rangle_\beta. \tag{2.5}$$

Note that $\epsilon_{k_1 k_2 k_3}$ is always supported in the nondegenerate set

$$\mathfrak{S} := \{(k_1, k_2, k_3) : \text{either } k_2 \notin \{k_1, k_3\}, \text{ or } k_1 = k_2 = k_3\}. \tag{2.6}$$

Below we will focus on the system (2.2)–(2.3), with the relevant terms defined in (2.4)–(2.6), for time $t \in [0, 1]$.

2.2. Reduction of Theorem 1.3

By plugging in (2.1) we can reduce Theorem 1.3 to proving the following bounds:

$$\left| \mathbb{E}\left(\prod_{j=1}^{r} (a_{k_j}(t))^{p_j} \, \overline{(a_{k_j}(t))}^{q_j} \right) \right| \lesssim_R L^{-\nu} \tag{2.7}$$

if $p_j \neq q_j$ for some $1 \leq j \leq r$, and

$$\left| \mathbb{E}\left(\prod_{j=1}^{r} |a_{k_j}(t)|^{2p_j} \right) - \prod_{j=1}^{r} p_j! \, n(\delta t, k_j)^{p_j} \right| \lesssim_R L^{-\nu}, \tag{2.8}$$

uniformly in $t \in [0, 1]$ and in $k_1, \ldots, k_r \in \mathbb{Z}_L^d$ satisfying $k_i \neq k_j$, with $\nu > 0$ being an absolute constant and the implicit constants depending on $R := (p_1 + \cdots + p_r + q_1 + \cdots + q_r)/2$.

Note that if $a_k(t)$ solves (2.2) then $e^{i\theta} a_k(t)$ solves the same equation, with the initial data obeying the same law. From this it is easy to deduce that

$$\mathbb{E}\left(\prod_{j=1}^{r} (a_{k_j}(t))^{p_j} \, \overline{(a_{k_j}(t))}^{q_j} \right) = 0 \quad \text{if } p_1 + \cdots + p_r \neq q_1 + \cdots + q_r.$$

Below we will always assume $p_1 + \cdots + p_r = q_1 + \cdots + q_r = R$. As we consider the limit $L \to \infty$ with R fixed, we may assume $R \leq \log L$. We shall introduce a simpler notation as follows. For $1 \leq j \leq r$, take p_j copies of the variable k_j with sign $+$ and q_j copies of the variable k_j with sign $-$, and rename them as $(k_1^*, \ldots, k_{2R}^*)$ with associated signs ζ_j ($1 \leq j \leq 2R$). For simplicity we will write k_j instead of k_j^* below. Then (2.7) and (2.8) result from the following unified and more precise estimate:

$$\left| \mathbb{E}\left(\prod_{j=1}^{2R} a_{k_j}^{\zeta_j}(t) \right) - \sum_{\mathcal{P}} \prod_{\{j,j'\} \in \mathcal{P}} \mathbf{1}_{k_j = k_{j'}} \prod_{j}^{(+)} n(\delta t, k_j) \right| \lesssim R! \cdot M_{\mathrm{kin}}^R \cdot L^{-\nu}. \tag{2.9}$$

Here we denote $z^+ = z$ and $z^- = \bar{z}$, and the sum is taken over all partitions $\mathcal{P}$ of $\{1, \ldots, 2R\}$ into two-element subsets $\{j, j'\}$ such that $\zeta_{j'} = -\zeta_j$. The first product is taken over all $\{j, j'\} \in \mathcal{P}$, and the second is over all $1 \leq j \leq 2R$ such that $\zeta_j = +$. Finally, M_{kin} is defined as

$$M_{\mathrm{kin}} := 1 + \sup_{t \in [0,1], \, k \in \mathbb{R}^d} |n(\delta t, k)|, \tag{2.10}$$

and the implicit constant in (2.9) depends only on $(d, \beta, n_{\mathrm{in}})$ but not on R.

The goal of the rest of the paper is to prove (2.9).

2.3. Parameters and notations

Most of our parameters and notations are taken from [12]. First, we fix $\beta \in (\mathbb{R}^+)^d \setminus \mathfrak{Z}$, where $\mathfrak{Z}$ is defined by the following lemma.

Lemma 2.1 ([12, Lemma A.1]). *There exists a Lebesgue null set $\mathfrak{Z} \subset (\mathbb{R}^+)^d$ such that the following hold for any $\beta = (\beta^1, \ldots, \beta^d) \in (\mathbb{R}^+)^d \setminus \mathfrak{Z}$:*

(1) *For any integers $(K^1, K^2) \neq (0, 0)$, we have*

$$|\beta^1 K^1 + \beta^2 K^2| \gtrsim (1 + |K^1| + |K^2|)^{-1} \log^{-4}(2 + |K^1| + |K^2|). \qquad (2.11)$$

(2) *The numbers $\beta^1, \ldots, \beta^d$ are algebraically independent over $\mathbb{Q}$, and for any R we have*

$$\#\{(X, Y, Z) \in (\mathbb{Z}^d)^3 : |X|, |Y|, |Z| \leq R, \ X \neq 0, \ \max(|\langle X, Y \rangle_\beta|, |\langle X, Z \rangle_\beta|) \leq 1\}$$
$$\lesssim R^{3d-4+1/6}. \qquad (2.12)$$

Throughout this paper, we will use C to denote any large constant that depends only on the dimension d, and use C^+ to denote any large constant that depends on $(d, \beta, n_{\mathrm{in}})$; these may differ from line to line, and note in particular that they do *not* depend on the value of R in (2.9). The notations $X \lesssim Y$ and $X = O(Y)$ will mean $X \leq C^+ Y$ unless otherwise stated.

Recall that $A \geq 40d$ and δ, which is small enough depending on A and C^+, are fixed as in Theorem 1.3. We also fix $\nu = (100d)^{-1} \ll 1$ and define $N = \lfloor (\log L)^4 \rfloor$. Note that the value of N is different from the one in [12]. As above, we assume $R \leq \log L$. For later purposes we may need slightly larger values (like $2R$), but all our proofs work equally fine as long as $R \leq 2 \log L$, which will be satisfied throughout the paper. Note that we do not assume any inequality between δ and R.

We adopt the shorthand notation $k[A] = (k_j)_{j \in A}$ and similarly for other vectors, and also define $d\alpha[A] = \prod_{j \in A} d\alpha_j$. We also use multi-indices ρ with the usual notations. Define the time Fourier transform (the meaning of $\hat{\cdot}$ later may depend on the context)

$$\hat{u}(\lambda) = \int_{\mathbb{R}} u(t) e^{-2\pi i \lambda t} \, dt, \quad u(t) = \int_{\mathbb{R}} \hat{u}(\lambda) e^{2\pi i \lambda t} \, d\lambda.$$

Define the X^κ norm for functions $F = F(t, k)$ or $G = G(t, s, k)$ by

$$\|F\|_{X^\kappa} = \int_{\mathbb{R}} \langle \lambda \rangle^{1/9} \sup_k \langle k \rangle^\kappa |\hat{F}(\lambda, k)| \, d\lambda,$$

$$\|G\|_{X^\kappa} = \int_{\mathbb{R}^2} (\langle \lambda \rangle + \langle \mu \rangle)^{1/9} \sup_k \langle k \rangle^\kappa |\hat{G}(\lambda, \mu, k)| \, d\lambda \, d\mu,$$

where $\hat{\cdot}$ denotes the Fourier transform in t or (t, s). When F or G does not depend on k, this norm will not depend on κ and will be denoted by X. Define the localized version X^κ_{loc}

(and similarly X_{loc}) as

$$\|F\|_{X_{\text{loc}}^{\kappa}} = \inf\{\|\tilde{F}\|_{X^{\kappa}} : \tilde{F} = F \text{ for } 0 \leq t \leq 1\},$$
$$\|G\|_{X_{\text{loc}}^{\kappa}} = \inf\{\|\tilde{G}\|_{X^{\kappa}} : \tilde{G} = G \text{ for } 0 \leq t, s \leq 1\}.$$

If we only use the value of G in some set (for example $\{t > s\}$ in Proposition 3.8), then in the above definition we may only require $\tilde{G} = G$ in this set. Define the Z norm for a function $a = a_k(t)$ by

$$\|a\|_Z^2 = \sup_{0 \leq t \leq 1} L^{-d} \sum_{k \in \mathbb{Z}_L^d} \langle k \rangle^{10d} |a_k(t)|^2. \qquad (2.13)$$

3. A brief summary of [12]

The results of this section are proved in [12]. Here we state the relevant propositions and definitions that will be needed in the proof below.

3.1. Trees, couples, and decorations

We first recall the definitions of trees, couples, and decorations, which are drawn directly from [12].

Definition 3.1 ([12, Definition 2.1]). A *ternary tree* $\mathcal{T}$ (we will simply say a *tree* below) is a rooted tree where each nonleaf (or *branching*) node has exactly three children nodes, which we shall distinguish as the *left*, *mid* and *right* ones. We say $\mathcal{T}$ is *trivial* (and write $\mathcal{T} = \bullet$) if it consists only of the root, in which case this root is also viewed as a leaf.

   We denote generic nodes by $\mathfrak{n}$, generic leaves by $\mathfrak{l}$, the root by $\mathfrak{r}$, the set of leaves by $\mathscr{L}$ and the set of branching nodes by $\mathscr{N}$. The *scale* of a tree $\mathcal{T}$ is defined by $n(\mathcal{T}) = |\mathscr{N}|$, so if $n(\mathcal{T}) = n$ then $|\mathscr{L}| = 2n + 1$ and $|\mathcal{T}| = 3n + 1$.

   A tree $\mathcal{T}$ may have sign $+$ or $-$. If its sign is fixed then we decide the signs of its nodes as follows: the root $\mathfrak{r}$ has the same sign as $\mathcal{T}$, and for any branching node $\mathfrak{n} \in \mathscr{N}$, the signs of the three children nodes of $\mathfrak{n}$ from left to right are $(\zeta, -\zeta, \zeta)$ if $\mathfrak{n}$ has sign $\zeta \in \{\pm\}$. Once the sign of $\mathcal{T}$ is fixed, we will denote the sign of $\mathfrak{n} \in \mathcal{T}$ by $\zeta_{\mathfrak{n}}$. Define the conjugate $\overline{\mathcal{T}}$ of a tree $\mathcal{T}$ to be the same tree but with opposite sign.

Definition 3.2 ([12, Definition 2.2]). A *couple* $\mathcal{Q}$ is an unordered pair $(\mathcal{T}^+, \mathcal{T}^-)$ of two trees $\mathcal{T}^{\pm}$ with signs $+$ and $-$ respectively, together with a partition $\mathscr{P}$ of the set $\mathscr{L}^+ \cup \mathscr{L}^-$ into $n + 1$ pairwise disjoint two-element subsets, where $\mathscr{L}^{\pm}$ is the set of leaves for $\mathcal{T}^{\pm}$, and $n = n^+ + n^-$ where $n^{\pm}$ is the scale of $\mathcal{T}^{\pm}$. This n is also called the *scale* of $\mathcal{Q}$, denoted by $n(\mathcal{Q})$. The subsets $\{\mathfrak{l}, \mathfrak{l}'\} \in \mathscr{P}$ are referred to as *pairs*, and we require that $\zeta_{\mathfrak{l}'} = -\zeta_{\mathfrak{l}}$, i.e. the signs of paired leaves must be opposite. If both $\mathcal{T}^{\pm}$ are trivial, we call $\mathcal{Q}$ the *trivial couple* (and write $\mathcal{Q} = \times$).

   For a couple $\mathcal{Q} = (\mathcal{T}^+, \mathcal{T}^-, \mathscr{P})$ we denote the set of branching nodes by $\mathscr{N}^* = \mathscr{N}^+ \cup \mathscr{N}^-$, and the set of leaves by $\mathscr{L}^* = \mathscr{L}^+ \cup \mathscr{L}^-$; for simplicity we will abuse notation

and write $\mathcal{Q} = \mathcal{T}^+ \cup \mathcal{T}^-$. We also define a *paired tree* to be a tree where *some* leaves are paired to each other, according to the same pairing rule for couples. We say a paired tree is *saturated* if there is only one unpaired leaf (called the *lone leaf*). In this case the tree forms a couple with the trivial tree $\bullet$.

Definition 3.3 ([12, Definition 2.3]). A *decoration* $\mathcal{D}$ of a tree $\mathcal{T}$ is a set of vectors $(k_\mathfrak{n})_{\mathfrak{n} \in \mathcal{T}}$ such that $k_\mathfrak{n} \in \mathbb{Z}_L^d$ for each node $\mathfrak{n}$, and

$$k_\mathfrak{n} = k_{\mathfrak{n}_1} - k_{\mathfrak{n}_2} + k_{\mathfrak{n}_3}, \quad \text{or equivalently} \quad \zeta_\mathfrak{n} k_\mathfrak{n} = \zeta_{\mathfrak{n}_1} k_{\mathfrak{n}_1} + \zeta_{\mathfrak{n}_2} k_{\mathfrak{n}_2} + \zeta_{\mathfrak{n}_3} k_{\mathfrak{n}_3},$$

for each branching node $\mathfrak{n} \in \mathcal{N}$, where $\zeta_\mathfrak{n}$ is the sign of $\mathfrak{n}$ as in Definition 3.1, and $\mathfrak{n}_1, \mathfrak{n}_2, \mathfrak{n}_3$ are the three children nodes of $\mathfrak{n}$ from left to right. Clearly a decoration $\mathcal{D}$ is uniquely determined by the values of $(k_\mathfrak{l})_{\mathfrak{l} \in \mathcal{L}}$. For $k \in \mathbb{Z}_L^d$, we say $\mathcal{D}$ is a k-*decoration* if $k_\mathfrak{r} = k$ for the root $\mathfrak{r}$.

Given a decoration $\mathcal{D}$, we define the coefficient

$$\epsilon_\mathcal{D} := \prod_{\mathfrak{n} \in \mathcal{N}} \epsilon_{k_{\mathfrak{n}_1} k_{\mathfrak{n}_2} k_{\mathfrak{n}_3}}, \tag{3.1}$$

where $\epsilon_{k_1 k_2 k_3}$ is as in (2.4). Note that in the support of $\epsilon_\mathcal{D}$ we have $(k_{\mathfrak{n}_1}, k_{\mathfrak{n}_2}, k_{\mathfrak{n}_3}) \in \mathfrak{S}$ for each $\mathfrak{n} \in \mathcal{N}$. We also define the *resonance factor* $\Omega_\mathfrak{n}$ for each $\mathfrak{n} \in \mathcal{N}$ by

$$\Omega_\mathfrak{n} = \Omega(k_{\mathfrak{n}_1}, k_{\mathfrak{n}_2}, k_{\mathfrak{n}_3}, k_\mathfrak{n}) = |k_{\mathfrak{n}_1}|_\beta^2 - |k_{\mathfrak{n}_2}|_\beta^2 + |k_{\mathfrak{n}_3}|_\beta^2 - |k_\mathfrak{n}|_\beta^2. \tag{3.2}$$

A decoration $\mathcal{E}$ of a couple $\mathcal{Q} = (\mathcal{T}^+, \mathcal{T}^-, \mathcal{P})$ is a set $(k_\mathfrak{n})_{\mathfrak{n} \in \mathcal{Q}}$ of vectors such that $\mathcal{D}^\pm := (k_\mathfrak{n})_{\mathfrak{n} \in \mathcal{T}^\pm}$ is a decoration of $\mathcal{T}^\pm$, and moreover $k_\mathfrak{l} = k_{\mathfrak{l}'}$ for each pair $\{\mathfrak{l}, \mathfrak{l}'\} \in \mathcal{P}$. We define $\epsilon_\mathcal{E} := \epsilon_{\mathcal{D}^+} \epsilon_{\mathcal{D}^-}$, and define the resonance factors $\Omega_\mathfrak{n}$ for $\mathfrak{n} \in \mathcal{N}^*$ as in (3.2). Note that we must have $k_{\mathfrak{r}^+} = k_{\mathfrak{r}^-}$ where $\mathfrak{r}^\pm$ is the root of $\mathcal{T}^\pm$; again we say $\mathcal{E}$ is a k-decoration if $k_{\mathfrak{r}^+} = k_{\mathfrak{r}^-} = k$. Finally, we can define decorations $\mathcal{D}$ of paired trees, as well as $\epsilon_\mathcal{D}$ and $\Omega_\mathfrak{n}$ etc., similar to the above.

Definition 3.4 ([12, Definition 4.2]). Define a *regular couple* to be a couple formed from the trivial couple $\times$ by repeatedly applying one of the steps $\mathbb{A}$ and $\mathbb{B}$, where in step $\mathbb{A}$ one replaces a pair of leaves with a $(1, 1)$-*mini couple*, and in step $\mathbb{B}$ one replaces a node with a mini tree. Here a $(1, 1)$-mini couple is a couple formed by two trees each of scale 1 such that no siblings are paired, and a *mini tree* is a saturated paired tree of scale 2 such that no siblings are paired. See Figures 1–3. We also define a *regular tree* to be a saturated paired tree $\mathcal{T}$ that forms a *regular* couple with the trivial tree. This is equivalent to the definition in [12, Remark 4.15], namely that $\mathcal{T}$ can be obtained from a regular chain by replacing each leaf pair with a regular couple. Here a *regular chain* (see [12, Definition 4.6]) is defined to be the result of repeatedly applying step $\mathbb{B}$ at a branching node or the lone leaf starting from the trivial tree $\bullet$. Note that the scale of a regular couple or a regular tree is always even.

Proposition 3.5. *The number of regular couples and regular trees of scale n is at most C^n.*

*Proof.* See [12, Corollary 4.9]. ∎

Fig. 1. A $(1, 1)$-mini couple. Here and below, two leaves of same color are paired There are two possibilities indicated by codes 00 and 01 as in [12].
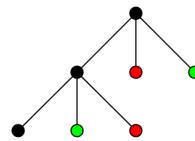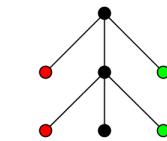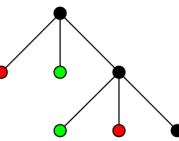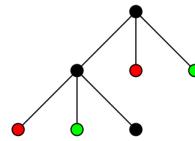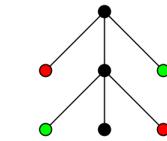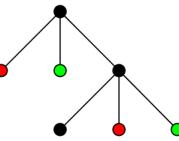

**Fig. 2.** A mini tree. There are six possibilities indicated by codes 10, 20, 30, 11, 21, 31 as in [12].
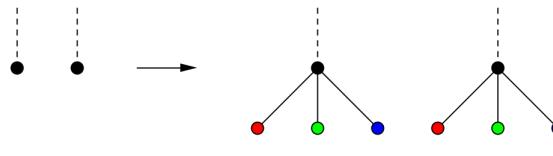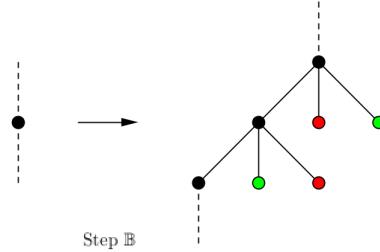
Fig. 3. Steps $\mathbb{A}$ and $\mathbb{B}$ as in Definition 3.4.

Lemma 3.6. *Let $\mathcal{T}$ be a tree of scale n. For any node $\mathfrak{n} \in \mathcal{T}$ define $\mu_{\mathfrak{n}}$ to be the number of leaves in the subtree rooted at $\mathfrak{n}$. Then, for any $\mathfrak{n} \in \mathcal{N}$, consider the values of $\mu_{\mathfrak{m}}$ where $\mathfrak{m}$ is a child of $\mathfrak{n}$, and let the* second maximum *of these values be $\mu_{\mathfrak{n}}^{\circ}$. Then*

$$\prod_{\mathfrak{n} \in \mathcal{N}} \mu_{\mathfrak{n}}^{\circ} \leq \frac{3^n}{2n+1}. \tag{3.3}$$

Proof. See [12, Lemma 6.6]. ∎

3.2. Expansion ansatz and regular couples

The following results are taken from [12].

Proposition 3.7. *For any tree $\mathcal{T}$, define*

$$(\mathcal{J}_{\mathcal{T}})_k(t) = \left(\frac{\delta}{2L^{d-1}}\right)^n \widetilde{\zeta}(\mathcal{T}) \sum_{\mathscr{D}} \epsilon_{\mathscr{D}} \int_{\mathfrak{D}} \prod_{\mathfrak{n} \in \mathcal{N}} e^{\pi i \zeta_{\mathfrak{n}} \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}} \prod_{\mathfrak{l} \in \mathcal{L}} \sqrt{n_{\mathrm{in}}(k_{\mathfrak{l}})} \cdot \eta_{k_{\mathfrak{l}}}^{\zeta_{\mathfrak{l}}}(\omega). \tag{3.4}$$

Here n is the scale of $\mathcal{T}$, $\widetilde{\zeta}(\mathcal{T}) = \prod_{\mathfrak{n} \in \mathcal{N}}(i\zeta_{\mathfrak{n}})$, $\mathscr{D}$ runs over all k-decorations of $\mathcal{T}$, and $\mathfrak{D}$ is the domain

$$\mathfrak{D} = \{t[\mathcal{N}] : 0 < t_{\mathfrak{n}'} < t_{\mathfrak{n}} < t \text{ whenever } \mathfrak{n}' \text{ is a child of } \mathfrak{n}\}. \tag{3.5}$$

We may expand $a_k(t)$ as

$$a_k(t) = \sum_{n=0}^{N} (\mathcal{J}_n)_k(t) + b_k(t), \quad (\mathcal{J}_n)_k(t) = \sum_{n(\mathcal{T}^+)=n} (\mathcal{J}_{\mathcal{T}^+})_k(t), \tag{3.6}$$

where the second sum is taken over all trees $\mathcal{T}^+$ of sign $+$ such that $n(\mathcal{T}^+) = n$.

The remainder b satisfies the equation

$$b = \mathscr{R} + \mathscr{L}b + \mathscr{B}(b,b) + \mathscr{C}(b,b,b), \tag{3.7}$$

where the terms on the right hand side are defined by

$$\mathscr{R} = \sum_{(0)} \mathscr{I}\mathscr{C}_+(u,\overline{v},w), \quad \mathscr{L}b = \sum_{(1)} \mathscr{I}\mathscr{C}_+(u,\overline{v},w),$$

$$\mathscr{B}(b,b) = \sum_{(2)} \mathscr{I}\mathscr{C}_+(u,\overline{v},w), \quad \mathscr{C}(b,b,b) = \mathscr{I}\mathscr{C}_+(b,\overline{b},b). \tag{3.8}$$

In (3.8) the summations are taken over (u,v,w), each of which being either b or $\mathcal{J}_n$ for some $0 \leq n \leq N$; moreover in the summation $\sum_{(j)}$ for $0 \leq j \leq 2$, exactly j inputs in (u,v,w) equal b, and in the summation $\sum_{(0)}$ we require that the sum of the three n's in the $\mathcal{J}_n$'s is at least N.

Lastly, uniformly in $t \in [0,1]$ and $k \in \mathbb{Z}_L^d$, we have

$$\left| \sum_{0 \leq n_1, n_2 \leq N} \mathbb{E}\left((\mathcal{J}_{n_1})_k(t)\overline{(\mathcal{J}_{n_2})_k(t)}\right) - n(\delta t, k) \right| \lesssim L^{-\nu}. \tag{3.9}$$

Proof. The expansion (3.6) is introduced in [12, Sections 2.2.1, 2.2.2], and (3.4) follows by combining the formulas in [12, Section 5.1]. Equation (3.7) for b is deduced in [12, Section 2.2.2]. Finally, (3.9) is a qualitative version of Theorem 1.1, which is proved in Section 12 of [12]. Note that here we are choosing $N = \lfloor (\log L)^4 \rfloor$ instead of $N = \lfloor \log L \rfloor$, but the proof is not affected as long as (say) $N \ll L^{\delta^2}$. ∎

Proposition 3.8. *For any couple $\mathcal{Q}$, define*

$$\mathcal{K}_{\mathcal{Q}}(t,s,k) = \left(\frac{\delta}{2L^{d-1}}\right)^n \zeta^*(\mathcal{Q}) \sum_{\mathscr{E}} \epsilon_{\mathscr{E}} \int_{\mathcal{E}} \prod_{\mathfrak{n}\in\mathcal{N}^*} e^{\pi i \zeta_{\mathfrak{n}} \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}} \prod_{\mathfrak{l}\in\mathcal{L}^*}^{(+)} n_{\mathrm{in}}(k_{\mathfrak{l}}). \tag{3.10}$$

Here n is the scale of $\mathcal{Q}$, $\zeta^*(\mathcal{Q}) = \prod_{\mathfrak{n}\in\mathcal{N}^*}(i\zeta_{\mathfrak{n}})$, $\mathscr{E}$ runs over all k-decorations of $\mathcal{Q}$, the last product is taken over all $\mathfrak{l}\in\mathcal{L}^*$ with sign $+$, and $\mathcal{E}$ is the domain

$$\mathcal{E} = \{t[\mathcal{N}^*] : 0 < t_{\mathfrak{n}'} < t_{\mathfrak{n}} \text{ when } \mathfrak{n}' \text{ is a child of } \mathfrak{n};$$
$$t_{\mathfrak{n}} < t \text{ for } \mathfrak{n}\in\mathcal{N}^+, \, t_{\mathfrak{n}} < s \text{ for } \mathfrak{n}\in\mathcal{N}^-\}. \tag{3.11}$$

Now suppose $\mathcal{Q}$ is a regular *couple with scale $2n$ where $n \leq (\log L)^{50}$, then there exist a function $(\mathcal{K}_{\mathcal{Q}})_{\mathrm{app}}(t,s,k)$, which is the sum of at most 2^n terms such that each term has the form $\delta^n \cdot \mathcal{J}\mathcal{A}(t,s) \cdot \mathcal{M}(k)$ (with possibly different $\mathcal{J}\mathcal{A}$ and $\mathcal{M}$ for different terms), and that*

$$\|\mathcal{J}\mathcal{A}\|_{X_{\mathrm{loc}}} \leq (C^+)^n, \quad \sup_{|\rho|\leq 40d} |\partial^\rho \mathcal{M}(k)| \leq (C^+)^n \langle k\rangle^{-40d} \quad \text{for each term,} \tag{3.12}$$

$$\|\mathcal{K}_{\mathcal{Q}}(t,s,k) - (\mathcal{K}_{\mathcal{Q}})_{\mathrm{app}}(t,s,k)\|_{X_{\mathrm{loc}}^{40d}} \leq (C^+\delta)^n L^{-2\nu}. \tag{3.13}$$

Similarly, for any regular tree $\mathcal{T}$ with lone leaf $\mathfrak{l}_*$, define

$$\mathcal{K}_{\mathcal{T}}^*(t,s,k) = \left(\frac{\delta}{2L^{d-1}}\right)^n \widetilde{\zeta}(\mathcal{T}) \sum_{\mathscr{D}} \int_{\mathcal{D}} \prod_{\mathfrak{n}\in\mathcal{N}} e^{\pi i \zeta_{\mathfrak{n}} \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}} \prod_{\mathfrak{l}\in\mathcal{L}\setminus\{\mathfrak{l}_*\}}^{(+)} n_{\mathrm{in}}(k_{\mathfrak{l}}). \tag{3.14}$$

Here n is the scale of $\mathcal{T}$, $\widetilde{\zeta}(\mathcal{T}) = \prod_{\mathfrak{n}\in\mathcal{N}}(i\zeta_{\mathfrak{n}})$, $\mathscr{D}$ runs over all k-decorations of $\mathcal{Q}$, the last product is taken over all $\mathfrak{l}\in\mathcal{L}\setminus\{\mathfrak{l}_*\}$ with sign $+$, and $\mathcal{D}$ is the domain

$$\mathcal{D} = \{t[\mathcal{N}] : 0 < t_{\mathfrak{n}'} < t_{\mathfrak{n}} < t \text{ when } \mathfrak{n}' \text{ is a child of } \mathfrak{n}; \, t_{(\mathfrak{l}_*)^p} > s\} \tag{3.15}$$

where $(\mathfrak{l}_*)^p$ is the parent of $\mathfrak{l}_*$. Suppose $\mathcal{T}$ has scale $2n$ where $n \leq (\log L)^{20}$. Then there exists a function $(\mathcal{K}_{\mathcal{T}}^*)_{\mathrm{app}}(t,s,k)$, which is the sum of at most 2^n terms, such that each term has the form $\delta^n \cdot \mathcal{J}\mathcal{A}^*(t,s) \cdot \mathcal{M}^*(k)$ (with possibly different $\mathcal{J}\mathcal{A}^*$ and $\mathcal{M}^*$ for different terms), and

$$\|\mathcal{J}\mathcal{A}^*\|_{X_{\mathrm{loc}}} \leq (C^+)^n, \quad \sup_{|\rho|\leq 40d} |\partial^\rho \mathcal{M}^*(k)| \leq (C^+)^n \quad \text{for each term,} \tag{3.16}$$

$$\|\mathcal{K}_{\mathcal{T}}^*(t,s,k) - (\mathcal{K}_{\mathcal{T}}^*)_{\mathrm{app}}(t,s,k)\|_{X_{\mathrm{loc}}^0} \leq (C^+\delta)^n L^{-2\nu}. \tag{3.17}$$

Proof. This follows from [12, Propositions 6.7, 6.10]. Whether the upper bound for n is $(\log L)^6$ or $(\log L)^{50}$ does not affect the proof (again, as long as $n \ll L^{\delta^2}$). ∎

4. Gardens

4.1. Structure of gardens

The key concept of this paper is a generalization of couples, which we call *gardens*.

Definition 4.1. Given a sequence $(\zeta_1, \ldots, \zeta_{2R})$, where $\zeta_j \in \{\pm\}$ and exactly R of them are $+$, we define a *garden* $\mathcal{G}$ of signature $(\zeta_1, \ldots, \zeta_{2R})$ to be an ordered collection $(\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ of trees such that $\mathcal{T}_j$ has sign ζ_j for $1 \leq j \leq 2R$, together with a partition $\mathscr{P}$ of the set of leaves in all $\mathcal{T}_j$ into two-element subsets (again called pairings) such that the two paired leaves have opposite signs; see Figure 4. The *width* of the garden is defined to be $2R$, which is always an even number. The *scale* $n(\mathcal{G})$ of a garden $\mathcal{G}$ is the sum of scales of all $\mathcal{T}_j$ ($1 \leq j \leq 2R$). We denote by $\mathscr{L}^* = \mathscr{L}_1 \cup \cdots \cup \mathscr{L}_{2R}$ the set of leaves and by $\mathscr{N}^* = \mathscr{N}_1 \cup \cdots \cup \mathscr{N}_{2R}$ the set of branching nodes, where $\mathscr{L}_j$ and $\mathscr{N}_j$ are the sets of leaves and branching nodes of $\mathcal{T}_j$.

Note that a garden of width 2 is just a couple. If the set $\{1, \ldots, 2R\}$ can be partitioned into two-element subsets such that for each such subset $\{j, j'\}$, the leaves in $\mathcal{T}_j$ and $\mathcal{T}_{j'}$ are all paired with each other (in particular $\zeta_{j'} = -\zeta_j$), then we say this garden is a *multi-couple*. In this case, this garden is formed by R couples $(\mathcal{T}_j, \mathcal{T}_{j'})$. If each of them is a regular couple then we say the multi-couple is *regular*. A *trivial* garden is a garden when all $\mathcal{T}_j$ are trivial trees; note that it is always a regular multi-couple (formed by R trivial couples). If in a garden $\mathcal{G}$, no two trees $\mathcal{T}_j$ and $\mathcal{T}_{j'}$ have all their leaves paired with each other, then we say the garden is *mixed*.

Definition 4.2. Given a garden $\mathcal{G}$, a *decoration* of $\mathcal{G}$, denoted by $\mathscr{I}$, is a set $(k_{\mathfrak{n}})_{\mathfrak{n} \in \mathcal{G}}$ of vectors where $\mathfrak{n}$ runs over all nodes of $\mathcal{G}$ such that $(k_{\mathfrak{n}})_{\mathfrak{n} \in \mathcal{T}_j}$ is a decoration of $\mathcal{T}_j$ for each $1 \leq j \leq 2R$, and $k_{\mathfrak{n}'} = k_{\mathfrak{n}}$ for each pair $\{\mathfrak{n}, \mathfrak{n}'\}$ of leaves. Given vectors $(k_1, \ldots, k_{2R})$, we say a decoration is a $(k_1, \ldots, k_{2R})$-*decoration* if $k_{\mathfrak{r}_j} = k_j$ for each $1 \leq j \leq 2R$, where $\mathfrak{r}_j$ is the root of $\mathcal{T}_j$. For any branching node $\mathfrak{n} \in \mathcal{N}^*$, define $\Omega_{\mathfrak{n}}$ as in (3.2); see Figure 4. We also define $\epsilon_{\mathscr{I}} = \prod_{j=1}^{2R} \epsilon_{\mathscr{D}_j}$, where $\epsilon_{\mathscr{D}_j}$ is defined as in (3.1), with $\mathscr{D}_j$ being the restriction of $\mathscr{I}$ to $\mathcal{T}_j$.

Definition 4.3. Define steps $\mathbb{A}$ and $\mathbb{B}$ for gardens in the same way as for couples in Definition 3.4; see Figure 3. Define a garden $\mathcal{G}$ to be *prime* if it is not obtained from any other garden by performing steps $\mathbb{A}$ or $\mathbb{B}$.

Proposition 4.4. *For any garden $\mathcal{G}$ there exists a unique prime garden $\mathcal{G}_{\mathrm{sk}}$ such that $\mathcal{G}$ is obtained from $\mathcal{G}_{\mathrm{sk}}$ by applying steps $\mathbb{A}$ and $\mathbb{B}$. This $\mathcal{G}_{\mathrm{sk}}$ is called the* skeleton *of $\mathcal{G}$. Finally, $\mathcal{G}_{\mathrm{sk}}$ is a trivial garden, if and only if $\mathcal{G}$ is a regular multi-couple.*

Proof. The proof is the same as that of [12, Proposition 4.13]. For the convenience of the reader we present the proof here. Denote the inverse operations of $\mathbb{A}$ and $\mathbb{B}$ by $\bar{\mathbb{A}}$ and $\bar{\mathbb{B}}$, where one collapses a $(1, 1)$-mini couple or a mini tree to a leaf pair or a single node. To prove existence of $\mathcal{G}_{\mathrm{sk}}$, by definition, one just needs to repeatedly apply $\bar{\mathbb{A}}$ and $\bar{\mathbb{B}}$ until no such operation is possible.
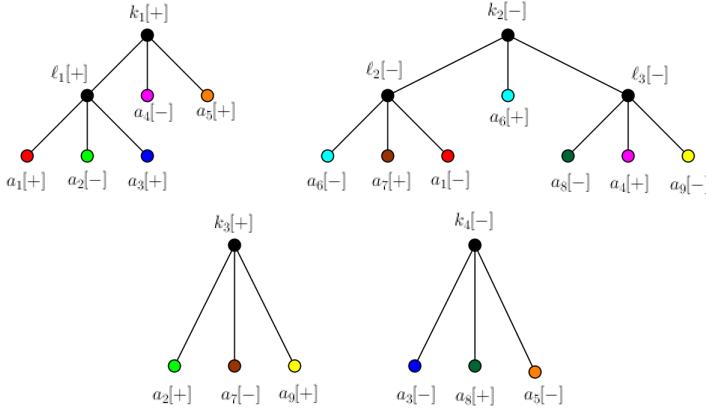
Fig. 4. An example of a garden (Definition 4.1) of width 4 and scale 7, together with a decoration (Definition 4.2). The signs of nodes are also indicated; the signature is $(+, -, +, -)$.

To prove uniqueness of $\mathcal{G}_{\mathrm{sk}}$, we just make one key observation: if $\mathcal{G}$ contains two basic objects (i.e. $(1, 1)$-mini couples or mini trees), and we let $\bar{\mathbb{D}}_1$ and $\bar{\mathbb{D}}_2$ be the inverse operations associated with them, then $\bar{\mathbb{D}}_1 \bar{\mathbb{D}}_2 = \bar{\mathbb{D}}_2 \bar{\mathbb{D}}_1$. In fact, this just shows that collapsing one of the basic objects does not affect the other, which can be directly verified by definition.

Now we can prove the uniqueness of $\mathcal{G}_{\mathrm{sk}}$ by induction. The base case is easy: suppose uniqueness is true for $\mathcal{G}$ of smaller scale; then for any $\mathcal{G}$ we shall look for $(1, 1)$-couples and mini trees (Definition 3.4). If there is none then $\mathcal{G}$ is already prime; if there is only one, then we apply $\bar{\mathbb{A}}$ or $\bar{\mathbb{B}}$ to collapse it and apply induction hypothesis for the resulting garden. Suppose there are more than one, then we apply $\bar{\mathbb{A}}$ or $\bar{\mathbb{B}}$ to collapse any one of them and apply induction hypothesis for the resulting garden. The final result does not depend on the first $\bar{\mathbb{A}}$ or $\bar{\mathbb{B}}$ we choose, because any two such steps, which can be performed for the original $\mathcal{G}$, must commute as proved above. Therefore $\mathcal{G}_{\mathrm{sk}}$ is unique. ∎

Proposition 4.5. *Suppose $\mathcal{G}$ is a garden with skeleton $\mathcal{G}_{\mathrm{sk}}$. Then $\mathcal{G}$ is formed from $\mathcal{G}_{\mathrm{sk}}$ by replacing each leaf pair with a regular couple and each branching node with a regular tree; see Figure 5. This representation is unique.*

Proof. The proof is basically the same as that of [12, Proposition 4.14]. To prove existence, we can induct on the scale of $\mathcal{G}$. The base case $\mathcal{G} = \mathcal{G}_{\mathrm{sk}}$ is obvious. Suppose the result is true for $\mathcal{G}$, and let $\mathcal{G}_+$ be obtained from $\mathcal{G}$ by applying $\mathbb{A}$ or $\mathbb{B}$. We know that $\mathcal{G}$ is obtained from $\mathcal{G}_{\mathrm{sk}}$ by replacing each branching node with a regular tree $\mathcal{T}_j$ ($1 \leq j \leq n$), and replacing each leaf pair by a regular couple $\mathcal{Q}_j$ ($1 \leq j \leq m$). Then:

(1) If one applies step $\mathbb{A}$, it must be applied either at a leaf pair belonging to some regular couple $\mathcal{Q}_i$ ($1 \leq i \leq m$), or at a leaf pair belonging to some regular tree $\mathcal{T}_i$ ($1 \leq i \leq n$). In this case the other regular trees and regular couples remain the same, and the regular tree $\mathcal{T}_i$ or regular couple $\mathcal{Q}_i$ is replaced by $\mathbb{A}\mathcal{T}_i$ or $\mathbb{A}\mathcal{Q}_i$.
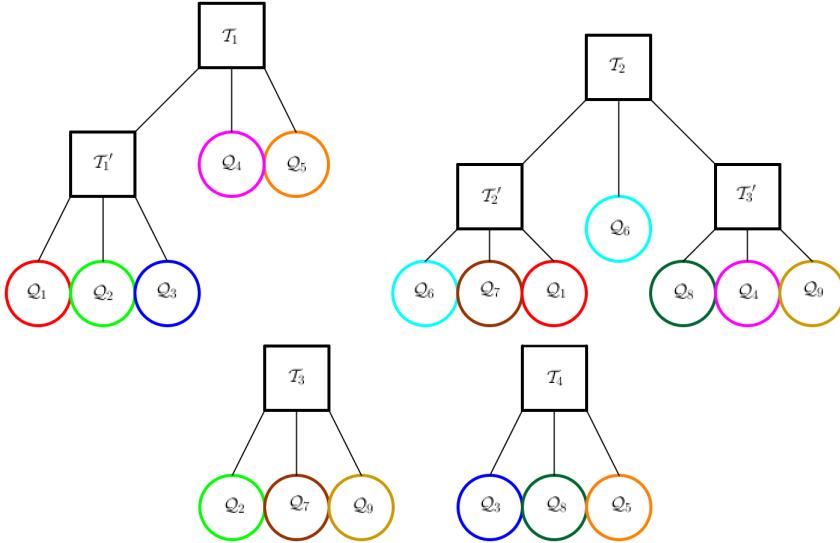
Fig. 5. A garden whose skeleton is the garden in Figure 4; see Proposition 4.5. Here each $\mathcal{T}_j$ and $\mathcal{T}_j'$ represents a regular tree, and each $\mathcal{Q}_j$ represents a regular couple.

(2) If one applies step $\mathbb{B}$, it must be applied either at a node belonging to some regular couple $\mathcal{Q}_i$ ($1 \leq i \leq m$), or at a node belonging to some regular tree $\mathcal{T}_i$ ($1 \leq i \leq n$). In this case the other regular trees and regular couples remain the same, and the regular tree $\mathcal{T}_i$ or regular couple $\mathcal{Q}_i$ is replaced by $\mathbb{B}\mathcal{T}_i$ or $\mathbb{B}\mathcal{Q}_i$.

In either case, notice that a regular tree or a regular couple still remains a regular tree or a regular couple after applying step $\mathbb{A}$ or $\mathbb{B}$. This proves existence.

Now to prove uniqueness of the representation, note that by Definition 3.4, the process of forming $\mathcal{G}$ from $\mathcal{G}_{sk}$ can also be described as follows: (i) first replace each branching node of $\mathcal{G}_{sk}$ by a *regular chain*, forming a garden $\mathcal{G}_{int}$; (ii) replacing each leaf pair in $\mathcal{G}_{int}$ by a regular couple to form $\mathcal{G}$. Given $\mathcal{G}_{sk}$, clearly $\mathcal{G}_{int}$ uniquely determines the regular chains in step (i), and also uniquely determines the regular couples in step (ii) replacing the leaf pairs in $\mathcal{G}_{int}$, so it suffices to show that $\mathcal{G}$ uniquely determines $\mathcal{G}_{int}$. Now we can show, via a case-by-case argument, that $\mathcal{G}_{int}$ contains no nontrivial regular subcouple (i.e. no two subtrees rooted at two nodes in $\mathcal{G}_{int}$ form a nontrivial regular couple). Since $\mathcal{G}$ is formed from $\mathcal{G}_{int}$ by replacing each leaf pair with a regular couple, we see that $\mathcal{G}_{int}$ can be reconstructed by collapsing each *maximal regular subcouple* (under inclusion) in $\mathcal{G}$ to a leaf pair (because any regular subcouple of $\mathcal{G}$ must be a subcouple of one of the regular couples in $\mathcal{G}$ replacing a leaf pair in $\mathcal{G}_{int}$). Clearly this collapsing process is commutative as explained in the proof of Proposition 4.4, hence the resulting couple $\mathcal{G}_{int}$ is unique. This completes the proof. ∎

Proposition 4.6. *Given any $\mathcal{G}_{sk}$, the number of gardens $\mathcal{G}$ that has scale m, width $2R$ and skeleton $\mathcal{G}_{sk}$ is at most C^{m+R}.*

Proof. This is basically the same as [12, Corollary 4.16]. If $\mathscr{G}$ has scale m and width $2R$, then $\mathscr{G}_{\mathrm{sk}}$ has scale at most m and width at most $2R$. Given $\mathscr{G}_{\mathrm{sk}}$, to construct $\mathscr{G}$, using Proposition 4.5, we just need to choose a regular tree at each branching node of $\mathscr{G}_{\mathrm{sk}}$, and a regular couple at each leaf pair of $\mathscr{G}_{\mathrm{sk}}$. Note that the number of branching nodes in $\mathscr{G}_{\mathrm{sk}}$ is at most m, and the number of leaf pairs in $\mathscr{G}_{\mathrm{sk}}$ is at most $m + R$. Thus the number of choices for $\mathscr{G}$ is at most

$$\sum_{n_1 + \cdots + n_{m'} \leq m} C_0^{n_1} \cdots C_0^{n_{m'}} \leq C^{m+R},$$

where $m' = 2m + R$, and C_0 is an absolute constant as in Proposition 3.5. ∎

4.2. Expressions $\mathcal{M}_{\mathscr{G}}$ for gardens $\mathscr{G}$

Given a garden $\mathscr{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ with width $2R$, signature $(\zeta_1, \ldots, \zeta_{2R})$ and scale m, and $k_j \in \mathbb{Z}_L^d$ for each $1 \leq j \leq 2R$, and time $t \in [0, 1]$, define

$$
\begin{aligned}
&\mathcal{M}_{\mathscr{G}}(t, k_1, \ldots, k_{2R}) \\
&\quad = \left(\frac{\delta}{2L^{d-1}}\right)^m \zeta^*(\mathscr{G}) \sum_{\mathscr{I}} \epsilon_{\mathscr{I}} \int_{\mathscr{J}} \prod_{\mathfrak{n} \in \mathcal{N}^*} e^{\pi i \zeta_{\mathfrak{n}} \cdot \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}} \cdot \prod_{\mathfrak{l} \in \mathcal{L}^*}^{(+)} n_{\mathrm{in}}(k_{\mathfrak{l}}). \quad (4.1)
\end{aligned}
$$

Here $\zeta^*(\mathscr{G}) = \prod_{\mathfrak{n} \in \mathcal{N}^*} (i \zeta_{\mathfrak{n}})$ and $\epsilon_{\mathscr{I}} = \prod_{j=1}^{2R} \epsilon_{\mathscr{D}_j}$ where $\mathscr{D}_j$ is the restriction of $\mathscr{I}$ to $\mathcal{T}_j$ (which is a k_j-decoration of $\mathcal{T}_j$), the sum is taken over all $(k_1, \ldots, k_{2R})$-decorations $\mathscr{I}$, the last product is taken over all $\mathfrak{l} \in \mathcal{L}^*$ with sign $+$, and $\mathscr{J}$ is the domain

$$\mathscr{J} = \{t[\mathcal{N}^*] : 0 < t_{\mathfrak{n}'} < t_{\mathfrak{n}} < t \text{ whenever } \mathfrak{n}' \text{ is a child of } \mathfrak{n}\}. \quad (4.2)$$

By using Isserlis' theorem [12, Lemma A.2] and repeating the arguments in [12, Section 2.2.3], we can find, for any tree $\mathcal{T}_j$ ($1 \leq j \leq 2R$) with sign ζ_j, that

$$\mathbb{E}\left(\prod_{j=1}^{2R} (\mathcal{J}_{\mathcal{T}_j})_{k_j}^{\zeta_j}(t)\right) = \sum_{\mathscr{P}} \mathcal{M}_{\mathscr{G}}(t, k_1, \ldots, k_{2R}). \quad (4.3)$$

Here the sum is taken over all possible pairings $\mathscr{P}$ that make $(\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ a garden, and $\mathscr{G}$ is the resulting garden.

We can reduce (2.9) to the following two propositions. Here Proposition 4.7 is the key component, and Proposition 4.8 follows from similar arguments. Note also that Proposition 4.8 is actually an improvement of [12, Propositions 12.1, 12.2], where the decay of exceptional probability is improved from L^{-A} to $e^{-(\log L)^3}$.

Proposition 4.7. *Fix R and $(\zeta_1, \ldots, \zeta_{2R})$ and $(k_1, \ldots, k_{2R})$ and $(m_1, \ldots, m_{2R})$. Assume $R \leq 2 \log L$, and $m_j \leq N$ ($1 \leq j \leq 2R$), and set $m := m_1 + \cdots + m_{2R}$. Let*

$$\mathscr{S} = \sum_{\mathscr{G}} \mathcal{M}_{\mathscr{G}}(t, k_1, \ldots, k_{2R}), \quad (4.4)$$

where the sum is taken over all mixed *gardens* $\mathcal{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ of width $2R$ and signature $(\zeta_1, \ldots, \zeta_{2R})$ such that the scale of $\mathcal{T}_j$ is m_j for $1 \leq j \leq 2R$. Then

$$|\mathscr{S}| \lesssim (C^+ \delta^{1/4})^m L^{-\nu R} \tag{4.5}$$

uniformly in t and in $(k_1, \ldots, k_{2R})$.

Proposition 4.8. *With probability* $\geq 1 - e^{-(\log L)^3}$, *we have*

$$|(\mathcal{J}_n)_k(t)| \lesssim \langle k \rangle^{-9d} (C^+ \sqrt{\delta})^{n/2} e^{(\log L)^3}, \quad |\mathcal{R}_k(t)| \lesssim \langle k \rangle^{-9d} (C^+ \sqrt{\delta})^{N/2} e^{(\log L)^3} \tag{4.6}$$

for all $0 \leq n \leq N^3$, *as well as*

$$\|\mathscr{L}^n\|_{Z \to Z} \lesssim (C^+ \sqrt{\delta})^{n/2} e^{(\log L)^3} \tag{4.7}$$

for all $0 \leq n \leq N$, *uniformly in any* $k \in \mathbb{Z}_L^d$ *and* $t \in [0, 1]$. *Here* $\mathcal{R}$ *and* $\mathscr{L}$ *are defined in* (3.8), *and the* Z *norm is defined in* (2.13).

Before proceeding, we illustrate how Propositions 4.7 and 4.8 imply Theorem 1.3.

Proof of Theorem 1.3. We only need to prove (2.9). Let E_1 be the event that (NLS) has a smooth solution on $[0, T]$, and $E \subset E_1$ be the event that Proposition 4.8 holds.c then $\mathbb{P}(E_1) \geq \mathbb{P}(E) \geq 1 - e^{-(\log L)^3}$.

Note that, under the assumption E, we can bound the remainder b defined in (3.6) by $\|b\|_Z \leq e^{-(\log L)^4}$. This can be proved similarly to [12, Proposition 12.3]. In fact, equation (3.7) satisfied by b can be written as

$$b = (1 - \mathscr{L}^N)^{-1}(1 + \mathscr{L} + \cdots + \mathscr{L}^{N-1})(\mathcal{R} + \mathscr{B}(b, b) + \mathscr{C}(b, b, b)). \tag{4.8}$$

We view this as the fixed point equation for a contraction mapping from the set $\{b \in Z : \|b\|_Z \leq e^{-(\log L)^4}\}$ to itself, hence the solution b is unique and satisfies the desired bound. The contraction mapping property follows from the estimates (using also the definition of $\mathscr{B}$ and $\mathscr{C}$, see (3.8))

$$\|\mathcal{R}\|_Z \leq e^{-2(\log L)^4}, \tag{4.9}$$

$$\|\mathcal{J}_n\|_Z \leq e^{2(\log L)^3} \ (\forall \, 0 \leq n \leq N^3), \tag{4.10}$$

$$\|\mathcal{J}\mathcal{C}_+(f, g, h)\|_Z \leq L^{10d} \|f\|_Z \|g\|_Z \|h\|_Z, \tag{4.11}$$

$$\|\mathscr{L}^n\|_{Z \to Z} \leq e^{2(\log L)^3} \ (\forall \, 0 \leq n \leq N), \tag{4.12}$$

$$\|(1 - \mathscr{L}^N)^{-1}\|_{Z \to Z} \leq 2. \tag{4.13}$$

Here (4.9)–(4.10) follow from (4.6) and our choice $N = \lfloor (\log L)^4 \rfloor$, (4.11) is elementary, and (4.12)–(4.13) follow from (4.7), our choice $N = \lfloor (\log L)^4 \rfloor$ and Neumann series expansions.

Now, to prove (2.9) we need to calculate

$$\mathbb{E}\Big(\mathbf{1}_{E_1} \cdot \prod_{j=1}^{2R} a_{k_j}^{\zeta_j}(t)\Big). \tag{4.14}$$

As assumed, we have $R \leq \log L$. Using mass conservation we can bound $|a_{k_j}(t)| \lesssim L^d$ for each k_j, so if E_1 is replaced by $E_1 \setminus E$ in (4.14), the corresponding contribution is bounded by

$$\left| \mathbb{E}\left(\mathbf{1}_{E_1 \setminus E} \cdot \prod_{j=1}^{2R} a_{k_j}^{\zeta_j}(t) \right) \right| \leq \mathbb{P}(E^c) L^{2dR} \leq e^{-(\log L)^3} L^{2d \log L} \leq L^{-10},$$

so we may replace E_1 by E in (4.14). Under the assumption E, we may expand $a_{k_j}(t)$ using (3.6), which leads to different combinations of terms.

Consider the terms where all factors are of the form $\mathcal{J}_n$. For such factors we will also replace $\mathbf{1}_E$ by 1, and deal with the resulting error term later. Hence, we get a contribution

$$\sum_{1 \leq m_1, \ldots, m_{2R} \leq N} \mathbb{E}\left(\prod_{j=1}^{2R} (\mathcal{J}_{m_j})_{k_j}^{\zeta_j}(t) \right). \tag{4.15}$$

For fixed $(m_1, \ldots, m_{2R})$, using the second expansion in (3.6) and (4.3), we can write

$$\mathbb{E}\left(\prod_{j=1}^{2R} ((\mathcal{J}_{m_j})_{k_j}(t))^{\zeta_j} \right) = \sum_{\mathcal{G}} \mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R}), \tag{4.16}$$

where the sum is taken over *all gardens* $\mathcal{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ of width $2R$ and signature $(\zeta_1, \ldots, \zeta_{2R})$ such that the scale of $\mathcal{T}_j$ is m_j for $1 \leq j \leq 2R$. Note that by definition, each $\mathcal{G}$ is uniquely expressed as the union of some couples and a mixed garden; suppose the number of couples is $R_1 \leq R$, and $R_2 := R - R_1$. If $R_2 = 0$, then there is a unique partition $\mathcal{P}$ of $\{1, \ldots, 2R\}$ into two-element subsets $\{j, j'\}$ such that $\zeta_{j'} = -\zeta_j$ and $\{\mathcal{T}_j, \mathcal{T}_{j'}\}$ is a couple for each pair $\{j, j'\}$; moreover, for $\mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R})$ to be nonzero one must have $k_j = k_{j'}$. For $\mathcal{P}$ fixed, the contribution of this part of sum equals

$$\prod_{\{j,j'\} \in \mathcal{P}} \mathbf{1}_{k_j = k_{j'}} \sum_{\mathcal{Q}} \mathcal{M}_{\mathcal{Q}}(t, k_j, k_j) = \prod_{\{j,j'\} \in \mathcal{P}} \mathbf{1}_{k_j = k_{j'}} \cdot \mathbb{E}\left((\mathcal{J}_{m_j})_{k_j}^{\zeta_j}(t) (\mathcal{J}_{m_{j'}})_{k_j}^{\zeta_{j'}}(t) \right), \tag{4.17}$$

where for fixed $\{j, j'\} \in \mathcal{P}$, the sum is taken over all couples $\mathcal{Q} = \{\mathcal{T}_j, \mathcal{T}_{j'}\}$ such that the two trees have signs ζ_j and $\zeta_{j'}$ and scales m_j and $m_{j'}$ respectively, and the equality in (4.17) follows from (4.3). Now, upon summing over all choices for $(m_1, \ldots, m_{2R})$ and using (3.9), we find that this contribution equals

$$\sum_{\mathcal{P}} \prod_{\{j,j'\} \in \mathcal{P}} \mathbf{1}_{k_j = k_{j'}} \prod_{j}^{(+)} (n(\delta t, k_j) + O(L^{-\nu}))$$

$$= \sum_{\mathcal{P}} \prod_{\{j,j'\} \in \mathcal{P}} \mathbf{1}_{k_j = k_{j'}} \prod_{j}^{(+)} n(\delta t, k_j) + O(R! M_{\mathrm{kin}}^R L^{-\nu}),$$

where in the last inequality we have used $1 + |n(\delta t, k_j)| \leq M_{\mathrm{kin}}$ for each j.

Next, consider the contribution where $R_2 > 0$. Up to a factor $\binom{2R}{2R_2} R_1! \leq (2R)^{2R_2} R!$ and a permutation, we may assume $\{\mathcal{T}_{2j-1}, \mathcal{T}_{2j}\}$ is a couple for $R_2 + 1 \leq j \leq R$ and

$\mathcal{G}_2 := (\mathcal{T}_1, \ldots, \mathcal{T}_{2R_2})$ is a mixed garden. Again we must have $k_{2j-1} = k_{2j}$ for $R_2 + 1 \leq j \leq R$; if we fix $(m_1, \ldots, m_{2R_2})$ and sum over the other m_j, then in the same way as above, we can bound the corresponding contribution by

$$(2R)^{2R_2} R! \prod_{j=R_2+1}^{R} \left(n(\delta t, k_{2j-1}) + O(L^{-\nu}) \right) \cdot \sum_{\mathcal{G}_2} \mathcal{K}_{\mathcal{G}_2}(t, k_1, \ldots, k_{2R_2}), \qquad (4.18)$$

where the sum is taken over all mixed gardens $\mathcal{G}_2 = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R_2})$ of width $2R_2$ and signature $(\zeta_1, \ldots, \zeta_{2R_2})$ such that the scale of $\mathcal{T}_j$ is m_j. By Proposition 4.7 we have

$$(4.18) \leq (2R)^{2R_2} R! \cdot M_{\text{kin}}^R \cdot \delta^{(m_1+\cdots+m_{2R_2})/8} L^{-\nu R_2}.$$

Upon summing over $(m_1, \ldots, m_{2R_2})$ and using $R \leq \log L$, we can bound this contribution by the right hand side of (2.9).

Finally, we show that all the remainder terms are bounded by the right hand side of (2.9). In fact, the above arguments imply that

$$\left| \mathbb{E}\left(\prod_{j=1}^{2R} (\widetilde{\mathcal{J}})_{k_j}^{\zeta_j}(t) \right) \right| \lesssim R! \cdot M_{\text{kin}}^R, \quad \text{where} \quad \widetilde{\mathcal{J}} = \sum_{n=0}^{N} \mathcal{J}_n;$$

in particular, we have

$$\mathbb{E}\left(\prod_{j=1}^{2R} |(\widetilde{\mathcal{J}})_{k_j}(t)|^2 \right) \lesssim (2R)! \cdot M_{\text{kin}}^{2R}. \qquad (4.19)$$

Since $R \leq \log L$, this allows us to control the terms where all factors are of the form $\mathcal{J}_n$, but with $\mathbf{1}_E$ replaced by $1 - \mathbf{1}_E$ (where we simply apply Cauchy–Schwarz and use the fact $\mathbb{P}(E^c) \leq e^{-(\log L)^3}$); similarly, if at least one factor in the expansion is the remainder b, then we can also apply Cauchy–Schwarz and use the bound $|b_k(t)| \leq e^{-(\log L)^4}$ together with (4.19) to control this term. This completes the proof. ∎

From now on we will focus on the proof of Propositions 4.7 and 4.8.

5. Irregular chains

5.1. Reduction to prime gardens

Let $\mathcal{G}_{\text{sk}}$ be the skeleton of a garden $\mathcal{G}$, which is then a prime garden. By Proposition 4.5, $\mathcal{G}$ can be obtained from $\mathcal{G}_{\text{sk}}$ by replacing each branching node $\mathfrak{m}$ with a regular tree $\mathcal{T}^{(\mathfrak{m})}$, and replacing each leaf pair $\{\mathfrak{m}, \mathfrak{m}'\}$ in $\mathcal{G}_{\text{sk}}$ with a regular couple $\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}$. Similar to [12, Section 8.1], using Proposition 3.8, we can reduce $\mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R})$ to an expression that has similar form to $\mathcal{M}_{\mathcal{G}_{\text{sk}}}(t, k_1, \ldots, k_{2R})$. For the sake of completeness we briefly recall the reduction process below.

Recall that

$$
\mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R})
$$
$$
= \left(\frac{\delta}{2L^{d-1}}\right)^m \zeta^*(\mathcal{G}) \sum_{\mathscr{I}} \epsilon_{\mathscr{I}} \int_{\mathcal{J}} \prod_{\mathfrak{n} \in \mathcal{N}^*} e^{\pi i \zeta_{\mathfrak{n}} \cdot \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}} \cdot \prod_{\mathfrak{l} \in \mathscr{L}^*}^{(+)} n_{\mathrm{in}}(k_{\mathfrak{l}}), \quad (5.1)
$$

where m is the scale of $\mathcal{G}$, $\mathcal{J}$ is the domain defined in (4.2), $\mathscr{I}$ is a $(k_1, \ldots, k_{2R})$-decoration and other objects are as before, all associated to the garden $\mathcal{G}$. By definition, the restriction of $\mathscr{I}$ to nodes in $\mathcal{G}_{\mathrm{sk}}$ forms a $(k_1, \ldots, k_{2R})$-decoration of $\mathcal{G}_{\mathrm{sk}}$, and the relevant quantities such as $\Omega_{\mathfrak{n}}$ are the same for both decorations (i.e. each $\Omega_{\mathfrak{n}}$ in the decoration of $\mathcal{G}_{\mathrm{sk}}$ uniquely corresponds to some $\Omega_{\mathfrak{n}}$ in the corresponding decoration of $\mathcal{G}$).

Now, let $\{\mathfrak{m}, \mathfrak{m}'\}$ be a leaf pair in $\mathcal{G}_{\mathrm{sk}}$, which becomes the roots of the regular subcouple $\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}$ in $\mathcal{G}$. We must have $k_{\mathfrak{m}} = k_{\mathfrak{m}'}$. In (5.1), consider the summation in the variables $k_{\mathfrak{n}}$, where $\mathfrak{n}$ runs over all nodes in $\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}$ other than $\mathfrak{m}$ and $\mathfrak{m}'$ (these variables, together with $k_{\mathfrak{m}}$ and $k_{\mathfrak{m}'}$, form a $k_{\mathfrak{m}}$-decoration of $\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}$), and the integration in the variables $t_{\mathfrak{n}}$, where $\mathfrak{n}$ runs over all branching nodes in $\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}$, with all the other variables fixed. By definition, this summation and integration equals, up to some sign $\zeta^*(\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')})$ and some power of $\delta(2L^{d-1})^{-1}$, the *exact* expression $\mathcal{K}_{\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}}(t_{\mathfrak{m}^p}, t_{(\mathfrak{m}')^p}, k_{\mathfrak{m}})$. Here we assume $\zeta_{\mathfrak{m}} = +$ and $\zeta_{\mathfrak{m}'} = -$, and $\mathfrak{m}^p$ is the parent of $\mathfrak{m}$ (if $\mathfrak{m}$ is the root of some tree then $t_{\mathfrak{m}^p}$ should be replaced by t; similarly for $(\mathfrak{m}')^p$). The relevant notations here and below are defined as in Proposition 3.8.

Similarly, let $\mathfrak{m}$ be a branching node in $\mathcal{G}_{\mathrm{sk}}$, which becomes the root $\mathfrak{p}$ and lone leaf $\mathfrak{q}$ of a regular tree $\mathcal{T}^{(\mathfrak{m})}$ in $\mathcal{Q}$. We must have $k_{\mathfrak{p}} = k_{\mathfrak{q}}$. In (5.1), consider the summation in the variables $k_{\mathfrak{n}}$, where $\mathfrak{n}$ runs over all nodes in $\mathcal{T}^{(\mathfrak{m})}$ other than $\mathfrak{p}$ and $\mathfrak{q}$ (these variables, together with $k_{\mathfrak{p}}$ and $k_{\mathfrak{q}}$, form a $k_{\mathfrak{m}}$-decoration of $\mathcal{T}^{(\mathfrak{m})}$), and the integration in the variables $t_{\mathfrak{n}}$, where $\mathfrak{n}$ runs over all branching nodes in $\mathcal{T}^{(\mathfrak{m})}$, with all the other variables fixed. In the same way, this summation and integration equals, up to some sign $\widetilde{\zeta}(\mathcal{T}^{(\mathfrak{m})})$ and some power of $\delta(2L^{d-1})^{-1}$, the *exact* expression $\mathcal{K}^*_{\mathcal{T}^{(\mathfrak{m})}}(t_{\mathfrak{p}^p}, t_{\mathfrak{q}}, k_{\mathfrak{p}})$. Here $\mathfrak{p}^p$ is the parent of $\mathfrak{p}$ (again, if $\mathfrak{p}$ is the root of some tree then $t_{\mathfrak{p}^p}$ should be replaced by t).

After performing this reduction for each leaf pair and branching node of $\mathcal{G}_{\mathrm{sk}}$, we can reduce the summation in (5.1) to the summation in $k_{\mathfrak{m}}$ for all leaves and branching nodes $\mathfrak{m}$ of $\mathcal{G}_{\mathrm{sk}}$, i.e. a $(k_1, \ldots, k_{2R})$-decoration of $\mathcal{G}_{\mathrm{sk}}$. Moreover, we can reduce the integration in (5.1) to the integration in $t_{\mathfrak{m}}$ for all branching nodes $\mathfrak{m}$ of $\mathcal{G}_{\mathrm{sk}}$ (for a regular tree, the time variables $t_{\mathfrak{p}^p}$ and $t_{\mathfrak{q}}$ for $\mathcal{G}$ correspond to $t_{\mathfrak{m}^p}$ and $t_{\mathfrak{m}}$ for $\mathcal{G}_{\mathrm{sk}}$ where $\mathfrak{m}^p$ is the parent of $\mathfrak{m}$). This implies that

$$
\mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R}) = \left(\frac{\delta}{2L^{d-1}}\right)^{m_0} \zeta^*(\mathcal{G}_{\mathrm{sk}}) \sum_{\mathscr{I}_{\mathrm{sk}}} \int_{\mathcal{J}_{\mathrm{sk}}} \epsilon_{\mathscr{I}_{\mathrm{sk}}} \prod_{\mathfrak{n} \in \mathcal{N}^*_{\mathrm{sk}}} e^{\zeta_{\mathfrak{n}} \pi i \cdot \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}}
$$
$$
\times \prod_{\mathfrak{m} \in \mathscr{L}^*_{\mathrm{sk}}}^{(+)} \mathcal{K}_{\mathcal{Q}^{(\mathfrak{m}, \mathfrak{m}')}}(t_{\mathfrak{m}^p}, t_{(\mathfrak{m}')^p}, k_{\mathfrak{m}}) \prod_{\mathfrak{m} \in \mathcal{N}^*_{\mathrm{sk}}} \mathcal{K}^*_{\mathcal{T}^{(\mathfrak{m})}}(t_{\mathfrak{m}^p}, t_{\mathfrak{m}}, k_{\mathfrak{m}}), \quad (5.2)
$$

where m_0 is the scale of $\mathcal{G}_{\mathrm{sk}}$, $\mathcal{J}_{\mathrm{sk}}$ is the domain defined in (4.2), $\mathscr{I}_{\mathrm{sk}}$ is a $(k_1, \ldots, k_{2R})$-decoration of $\mathcal{G}_{\mathrm{sk}}$, the other objects are as before but associated to the garden $\mathcal{G}_{\mathrm{sk}}$. More-

over in (5.2), the first product is taken over all leaves $\mathfrak{m}$ of sign $+$ with $\mathfrak{m}'$ being the leaf paired to $\mathfrak{m}$, the second product is taken over all branching nodes $\mathfrak{m}$, and $\mathfrak{m}^p$ is the parent of $\mathfrak{m}$.

Using Proposition 3.8, in (5.2) we can decompose

$$\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')} = (\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')})_{\text{app}} + \mathscr{R}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}, \quad \mathcal{K}^*_{\mathcal{T}(\mathfrak{m})} = (\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})})_{\text{app}} + \mathscr{R}^*_{\mathcal{T}(\mathfrak{m})}. \quad (5.3)$$

Here $(\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')})_{\text{app}}$ and $(\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})})_{\text{app}}$ are the leading terms in Proposition 3.8, and each is a linear combination of functions of (t,s) multiplied by functions of k, which in turn satisfy (3.12) and (3.16); the remainders $\mathscr{R}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$ and $\mathscr{R}^*_{\mathcal{T}(\mathfrak{m})}$ satisfy (3.13) and (3.17).

We may fix a *mark* in $\{\mathfrak{L}, \mathfrak{R}\}$ for each leaf pair and each branching node in $\mathcal{G}_{\text{sk}}$ which indicates whether we select the *leading* term $(\cdots)_{\text{app}}$ or the *remainder* term $\mathscr{R}$ or $\mathscr{R}^*$; for a general garden $\mathcal{G}$ we can do the same but only for the nodes of its skeleton $\mathcal{G}_{\text{sk}}$. In this way we can define *marked* gardens, which we still denote by $\mathcal{G}$, and expressions of the form (5.2) but with $\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$ and $\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}$ replaced by the corresponding leading or remainder terms, which we still denote by $\mathcal{M}_{\mathcal{G}}$. By definition, any sum of $\mathcal{M}_{\mathcal{G}}$ over unmarked gardens $\mathcal{G}$ equals the corresponding sum over marked gardens $\mathcal{G}$ for all possible unmarked gardens and all possible markings.

In the next section we will define the notion of *irregular chains* to exhibit the cancellation between $\mathcal{M}_{\mathcal{G}}$ for some different gardens $\mathcal{G}$ with specific symmetries.

5.2. Irregular chains and congruence

The notion of irregular chains for gardens is defined in the same way as for couples (see [12, Section 8.2]).

Definition 5.1 ([12, Definition 8.1]). Given a garden $\mathcal{G}$ (or a paired tree $\mathcal{T}$), we define an *irregular chain* to be a sequence $(\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ of nodes such that (i) $\mathfrak{n}_{j+1}$ is a child of $\mathfrak{n}_j$ for $0 \leq j \leq q-1$, and the other two children of $\mathfrak{n}_j$ are leaves, and (ii) for $0 \leq j \leq q-1$, there is a child $\mathfrak{m}_j$ of $\mathfrak{n}_j$ which has opposite sign to $\mathfrak{n}_{j+1}$, and is paired (as a leaf) to a child $\mathfrak{p}_{j+1}$ of $\mathfrak{n}_{j+1}$. We also define $\mathfrak{p}_0$ to be the child of $\mathfrak{n}_0$ other than $\mathfrak{n}_1$ and $\mathfrak{m}_0$.

Definition 5.2 ([12, Definition 8.2]). Consider any irregular chain $\mathcal{H} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$. By Definition 5.1, we know $\mathfrak{p}_j$ is the child of $\mathfrak{n}_j$ other than $\mathfrak{n}_{j+1}$ and $\mathfrak{m}_j$ for $0 \leq j \leq q-1$, thus $\mathfrak{p}_j$ has the same sign as $\mathfrak{n}_j$ (hence it is either its first or third child). Now for two irregular chains $\mathcal{H} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ and $\mathcal{H}' = (\mathfrak{n}'_0, \ldots, \mathfrak{n}'_q)$, with $\mathfrak{p}_j$ and $\mathfrak{p}'_j$ etc. defined accordingly, we say they are *congruent* if $\zeta_{\mathfrak{n}_0} = \zeta_{\mathfrak{n}'_0}$, and for each $0 \leq j \leq q-1$, either $\mathfrak{p}_j$ is the first child of $\mathfrak{n}_j$ and $\mathfrak{p}'_j$ is the first child of $\mathfrak{n}'_j$, or $\mathfrak{p}_j$ is the third child of $\mathfrak{n}_j$ and $\mathfrak{p}'_j$ is the third child of $\mathfrak{n}'_j$, counting from left to right.

In particular, if q and the congruence class (and hence $\zeta_{\mathfrak{n}_0}$) are fixed, then an irregular chain $\mathcal{H}$ is uniquely determined by the signs $\zeta_{\mathfrak{n}_j}$ for $1 \leq j \leq q$. We relabel the nodes $\mathfrak{n}_j, \mathfrak{p}_j$ ($0 \leq j \leq q$) by defining $\{\mathfrak{b}_j, \mathfrak{c}_j\} = \{\mathfrak{n}_j, \mathfrak{p}_j\}$, and that $\mathfrak{b}_j = \mathfrak{n}_j$ if and only if $\zeta_{\mathfrak{n}_j} = +$. Further, we label the two children of $\mathfrak{n}_q$ other than $\mathfrak{p}_q$ as $\mathfrak{e}$ and $\mathfrak{f}$, with $\zeta_{\mathfrak{e}} = +$ and $\zeta_{\mathfrak{f}} = -$.

Proposition 5.3 ([12, Proposition 8.3]). *Let $\mathcal{H} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ be an irregular chain. For any decoration $\mathcal{D}$ (or $\mathcal{E}$), its restriction to $\mathfrak{n}_j$ ($0 \leq j \leq q$) and their children are uniquely determined by $2(q + 2)$ vectors $k_j, \ell_j \in \mathbb{Z}_L^d$ ($0 \leq j \leq q + 1$) such that $k_{\mathfrak{b}_j} = k_j$ and $k_{\mathfrak{c}_j} = \ell_j$ for $0 \leq j \leq q$, and $k_{\mathfrak{e}} = k_{q+1}$ and $k_{\mathfrak{f}} = \ell_{q+1}$. These vectors satisfy*

$$k_0 - \ell_0 = k_1 - \ell_1 = \cdots = k_{q+1} - \ell_{q+1} =: h,$$

and for each $0 \leq j \leq q$ we have $\zeta_{\mathfrak{n}_j} \Omega_{\mathfrak{n}_j} = 2\langle h, k_{j+1} - k_j \rangle_\beta$. Moreover, $\epsilon_{k_{\mathfrak{n}_{j1}} k_{\mathfrak{n}_{j2}} k_{\mathfrak{n}_{j3}}} = \epsilon_{k_{j+1} \ell_{j+1} \ell_j}$, where $(\mathfrak{n}_{j1}, \mathfrak{n}_{j2}, \mathfrak{n}_{j3})$ are the children of $\mathfrak{n}_j$ from left to right. We say this decoration has small gap, large gap *or* zero gap *with respect to $\mathcal{H}$ if $0 < |h| \leq \frac{1}{100\delta L}$, $|h| \geq \frac{1}{100\delta L}$ or $h = 0$ respectively.*

Definition 5.4 ([12, Definition 8.4]). Let $\mathcal{H} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ be an irregular chain contained in a garden $\mathcal{G}$ or a paired tree $\mathcal{T}$. If we replace $\mathcal{H}$ by a congruent irregular chain $\mathcal{H}' = (\mathfrak{n}'_0, \ldots, \mathfrak{n}'_q)$, then we obtain a modified couple $\mathcal{G}'$ or paired tree $\mathcal{T}'$ by (i) attaching the same subtree of $\mathfrak{e}$ and $\mathfrak{f}$ in $\mathcal{G}$ (or $\mathcal{T}$) to the bottom of $\mathfrak{e}'$ and $\mathfrak{f}'$, and (ii) assigning to $\mathfrak{n}'_0$ the same parent of $\mathfrak{n}_0$ and keeping the rest of the couple unchanged.

Given a marked prime garden $\mathcal{G}_{\text{sk}}$, we identify all the maximal irregular chains $\mathcal{H} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ such that $q \geq 10^3 d$, and all $\mathfrak{n}_j$ and their children have mark $\mathfrak{L}$. For each such maximal irregular chain $\mathcal{H}$, consider $\mathcal{H}^\circ = (\mathfrak{n}_5, \ldots, \mathfrak{n}_{q-5})$ formed by omitting five nodes at both ends (so that it does not affect other possible irregular chains). We define another marked prime couple $\widetilde{\mathcal{G}}_{\text{sk}}$ to be *congruent* to $\mathcal{G}_{\text{sk}}$ if it can be obtained from $\mathcal{G}_{\text{sk}}$ by changing each of the irregular chains $\mathcal{H}^\circ$ to a congruent irregular chain, as described above.

Given a marked garden $\mathcal{G}$, we define $\widetilde{\mathcal{G}}$ to be congruent to $\mathcal{G}$ if it can be formed as follows. First obtain the (marked) skeleton $\mathcal{G}_{\text{sk}}$ and change it to a congruent marked prime couple $\widetilde{\mathcal{G}}_{\text{sk}}$. Then, we attach the regular couples $\mathcal{Q}^{(\mathfrak{m},\mathfrak{m}')}$ and regular trees $\mathcal{T}^{(\mathfrak{m})}$ from $\mathcal{G}$ to the relevant leaf pairs and branching nodes of $\widetilde{\mathcal{G}}_{\text{sk}}$. Note that if an irregular chain $\mathcal{H}^\circ = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$ in $\mathcal{G}_{\text{sk}}$ is replaced by $(\mathcal{H}^\circ)' = (\mathfrak{n}'_0, \ldots, \mathfrak{n}'_q)$ in $\widetilde{\mathcal{G}}_{\text{sk}}$, with relevant nodes $\mathfrak{m}_j, \mathfrak{p}_j$ etc. as in Definition 5.1, then for $0 \leq j \leq q - 1$, the same regular couple $\mathcal{Q}^{(\mathfrak{m}_j, \mathfrak{p}_{j+1})}$ is attached to the leaf pair $\{\mathfrak{m}'_j, \mathfrak{p}'_{j+1}\}$ in $\widetilde{\mathcal{G}}_{\text{sk}}$. Similarly, for $1 \leq j \leq q$, if $\zeta_{\mathfrak{n}'_j} = \zeta_{\mathfrak{n}_j}$ then the same regular tree $\mathcal{T}^{(\mathfrak{n}_j)}$ is placed at the branching node $\mathfrak{n}'_j$ in $\widetilde{\mathcal{G}}_{\text{sk}}$; otherwise the conjugate regular tree $\overline{\mathcal{T}^{(\mathfrak{n}_j)}}$ is placed at $\mathfrak{n}'_j$.

Note that the congruence relation preserves the scale of each tree of a garden, i.e. if $\mathcal{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$ and $\widetilde{\mathcal{G}} = (\widetilde{\mathcal{T}}_1, \ldots, \widetilde{\mathcal{T}}_{2R})$ are congruent, then the scale of $\mathcal{T}_j$ equals the scale of $\widetilde{\mathcal{T}}_j$ for $1 \leq j \leq 2R$.

5.3. Expressions associated with irregular chains

We shall analyze the expressions associated with irregular chains, in the same way as in [12, Section 8.3].

Given *one congruence class* $\mathscr{F}$ of marked gardens as in Definition 5.4, consider the sum

$$\sum_{\mathcal{G}\in\mathscr{F}} \mathcal{M}_{\mathcal{G}}(t, k_1, \ldots, k_{2R}), \tag{5.4}$$

which is taken over all *marked gardens* $\mathcal{G} \in \mathscr{F}$. Let the lengths of all the irregular chains $\mathcal{H}^{\circ}$ involved in the congruence class $\mathscr{F}$, as in Definition 5.4, be $q_1, \ldots, q_r$. Then $|\mathscr{F}| = 2^Q$ where $Q = q_1 + \cdots + q_r$. Since these irregular chains do not affect each other, we may focus on one individual chain, say $\mathcal{H}^{\circ} = (\mathfrak{n}_0, \ldots, \mathfrak{n}_q)$; that is, we only sum over $\mathcal{G} \in \mathscr{F}$ obtained by altering this irregular chain $\mathcal{H}^{\circ}$.

In the summation and integration in (5.2), we will first fix all the variables $k_{\mathfrak{n}}$ and $t_{\mathfrak{n}}$, *except $k_{\mathfrak{n}}$ with $\mathfrak{n} \in \{\mathfrak{n}_j, \mathfrak{p}_j, \mathfrak{m}_{j-1}\}$ ($1 \leq j \leq q$) and $t_{\mathfrak{n}}$ with $\mathfrak{n} = \mathfrak{n}_j$ ($1 \leq j \leq q - 1$)*, and sum and integrate over these variables. Note that we are fixing $k_{\mathfrak{n}_0}$ and $k_{\mathfrak{p}_0}$ as well as $k_{\mathfrak{e}}$ and $k_{\mathfrak{f}}$, in the notation of Definition 5.2, and are thus fixing $(k_0, \ell_0, k_{q+1}, \ell_{q+1})$ and $k_0 - \ell_0 = k_{q+1} - \ell_{q+1} = h$ as in Proposition 5.3. It is easy to see that in the summation and integration in (5.2) over the *fixed variables* (i.e. those $k_{\mathfrak{n}}$ and $t_{\mathfrak{n}}$ not in the above list), the summand and integrand do not depend on the way $\mathcal{H}^{\circ}$ is changed, because the rest of the couple is preserved under the change of $\mathcal{H}^{\circ}$, by Definition 5.4.

We thus only need to consider the sum and integral over the variables listed above. By Proposition 5.3, this is the same as the sum over the variables k_j ($1 \leq j \leq q$), with $\ell_j := k_j - h$, and integral over the variables $t_j := t_{\mathfrak{n}_j}$ ($1 \leq j \leq q - 1$), which satisfies $t_0 > t_1 > \cdots > t_{q-1} > t_q$ with $t_0 := t_{\mathfrak{n}_0}$ and $t_q := t_{\mathfrak{n}_q}$. For any possible choice of $\mathcal{H}^{\circ}$ (there are 2^q of them), the sum and integral can be written, using (5.2) and Proposition 5.3, as

$$\sum_{k_1,\ldots,k_q} \int_{t_0>t_1>\cdots>t_{q-1}>t_q} \left(\frac{\delta}{2L^{d-1}}\right)^q \prod_{j=1}^{q} (i\zeta_{\mathfrak{n}_j}) \prod_{j=0}^{q} \epsilon_{k_{j+1}\ell_{j+1}\ell_j}$$

$$\times \prod_{j=0}^{q} e^{2\pi i \delta L^2 \langle h, k_{j+1}-k_j\rangle_\beta t_j} \prod_{j=1}^{q} \mathcal{K}_{j,\mathcal{H}^{\circ}} \cdot \mathcal{K}_{j,\mathcal{H}^{\circ}}^* \, dt_1 \cdots dt_{q-1}. \tag{5.5}$$

Here we have

$$\mathcal{K}_{j,\mathcal{H}^{\circ}} = \mathcal{K}_j(t_j, t_{j-1}, k_j - h), \quad \mathcal{K}_{j,\mathcal{H}^{\circ}}^* = \mathcal{K}_j^*(t_{j-1}, t_j, k_j)$$

if $\zeta_{\mathfrak{n}_j} = +$, and

$$\mathcal{K}_{j,\mathcal{H}^{\circ}} = \mathcal{K}_j(t_{j-1}, t_j, k_j), \quad \mathcal{K}_{j,\mathcal{H}^{\circ}}^* = \overline{\mathcal{K}_j^*(t_{j-1}, t_j, k_j - h)}$$

if $\zeta_{\mathfrak{n}_j} = -$, where $\mathcal{K}_j = (\mathcal{K}_{\mathcal{Q}^{(\mathfrak{p}_j, \mathfrak{m}_{j-1})}})_{\mathrm{app}}$ and $\mathcal{K}_j^* = (\mathcal{K}_{\mathcal{T}^{(\mathfrak{n}_j)}}^*)_{\mathrm{app}}$ where $\mathcal{T}^{(\mathfrak{n}_j)}$ is chosen to have sign $+$; note that if $\overline{\mathcal{T}}$ is the regular tree conjugate to $\mathcal{T}$ then $\mathcal{K}_{\overline{\mathcal{T}}}^* = \overline{\mathcal{K}_{\mathcal{T}}^*}$, and the same holds for the leading contribution $(\cdots)_{\mathrm{app}}$.

Note that to calculate the above-mentioned contribution (i.e. the sum (5.4) with only $\mathcal{H}^{\circ}$ altered), we need to sum over all possible choices of $\mathcal{H}^{\circ}$ (i.e. all possible choices

of $\zeta_{\mathfrak{n}_j}$ $(1 \leq j \leq q)$), in addition to the summation and integration in (5.5). This results in the expression

$$\sum_{\zeta_{\mathfrak{n}_j} \in \{\pm\} \, (1 \leq j \leq q)} (5.5) = \text{some function of } (k_0, \ell_0, k_{q+1}, \ell_{q+1}, t_0, t_q). \qquad (5.6)$$

Now (5.6) is exactly the same expression that is explicitly calculated in [12, Sections 8.3.1, 8.3.2], so we shall take the results of such calculations from [12] and apply them below. There are three cases depending on the value of $h := k_0 - \ell_0$.

(1) The zero gap case ($h = 0$): This is very easy, as we have $k_j = \ell_j$, so in view of the $\epsilon_{k_{j+1}\ell_{j+1}\ell_j}$ factors we must have $k_1 = \cdots = k_q = k_0$, so the expression (5.5) gains a large negative power of L, and can be treated in the same way as the small gap term below.

(2) The small gap case ($0 < |h| \leq (100\delta L)^{-1}$): We have

$$(5.6) = (C^+\delta)^{m_{\text{tot}}} (i\delta/2)^q \int_{\mathbb{R}} \int_0^1 G(\lambda) \mathcal{P}(\lambda, \sigma, k_0, \ell_0) \cdot \delta(t_0 - t_q - \sigma) e^{\pi i \delta L^2 \Omega^* t_q}$$
$$\times e^{\pi i \lambda t_q} \, \mathrm{d}\sigma \, \mathrm{d}\lambda. \qquad (5.7)$$

Here m_{tot} is the sum of the scales of all regular couples $\mathcal{Q}_{(\mathfrak{p}_j, \mathfrak{m}_{j-1})}$ and regular trees $\mathcal{T}_{(\mathfrak{n}_j)}$, $\Omega^* := |k_{q+1}|_\beta^2 - |\ell_{q+1}|_\beta^2 + |\ell_0|_\beta^2 - |k_0|_\beta^2$, and the functions G and $\mathcal{P}$ satisfy

$$\|\langle\lambda\rangle^{1/18} G\|_{L^1} \lesssim (C^+)^{m_{\text{tot}}}, \qquad \sup_{\lambda, k_0, \ell_0} \int_0^1 |\mathcal{P}(\lambda, \sigma, k_0, \ell_0)| \, \mathrm{d}\sigma \lesssim L^{-40d}. \qquad (5.8)$$

(3) The large gap case ($|h| > (100\delta L)^{-1}$): We have the same expression (5.7) and the same bound (5.8), but the factor L^{-40d} on the right hand side of the second inequality of (5.8) should be replaced by 1.

Below we will ignore the zero gap case. In the other two cases, we define the new marked garden $\mathscr{G}_{\text{sk}}^<$ as follows. In the small gap case, and in the large gap case assuming also $k_0 \neq k_{q+1}$, we remove the whole chain $\mathscr{H}^\circ$ by setting $(\mathfrak{p}_0, e, \mathfrak{f})$ (see Definition 5.2) to be the three children nodes of $\mathfrak{n}_0$, with the order determined by their signs and the relative position of $\mathfrak{p}_0$, and remove the other nodes (i.e. $(\mathfrak{n}_j, \mathfrak{p}_j)$ for $1 \leq j \leq q$ and $\mathfrak{m}_j$ for $0 \leq j \leq q - 1$). In the large gap case assuming $k_0 = k_{q+1}$, we must have $k_0 \neq k_q$ since $k_q \neq k_{q+1}$ in view of the factor $\epsilon_{k_{q+1}\ell_{q+1}\ell_q}$ in (5.5), so in this case we remove the chain $(\mathfrak{n}_0, \ldots, \mathfrak{n}_{q-1})$, which is the chain $\mathscr{H}^\circ$ minus one node, in the same way as above.

In either case, denote the scale of $\mathscr{G}_{\text{sk}}^<$ by $m_0^<$. Note that $\mathscr{G}_{\text{sk}}^<$ does not depend on the choice of $\mathscr{H}^\circ$ in the fixed congruence class (except in the large gap case, where this dependence does not matter), and for the decoration of $\mathscr{G}_{\text{sk}}^<$ coming from the decoration of $\mathscr{G}_{\text{sk}}$, we have $\zeta_{\mathfrak{n}_0}\Omega_{\mathfrak{n}_0} = \Omega^*$ for each choice of $\mathscr{H}^\circ$. Then, we can reduce the expression

$$\sum_{\mathscr{G}} \mathcal{M}_{\mathscr{G}}^{(*)}(t, k_1, \ldots, k_{2R}) \qquad (5.9)$$

using (5.7), where in (5.9) the sum is taken over all marked gardens $\mathscr{G}$ formed by altering the irregular chain $\mathcal{H}^\circ$ in $\mathscr{G}_{\mathrm{sk}}$, and ($*$) represents either "sg" or "lg", where we restrict to the small gap or large gap case. In fact, using (5.7) we have

$$(5.9) = (C^+\delta)^{m_{\mathrm{tot}}}(i\delta/2)^q \cdot \left(\frac{\delta}{2L^{d-1}}\right)^{m_0^<}\zeta^*(\mathscr{G}_{\mathrm{sk}}^<)$$

$$\times \int_{\mathbb{R}} G(\lambda)\,\mathrm{d}\lambda \int_0^1 \mathrm{d}\sigma \cdot \sum_{\mathscr{I}_{\mathrm{sk}}^<} \int_{\widetilde{\mathscr{J}}_{\mathrm{sk}}^<} \widetilde{\epsilon}_{\mathscr{I}_{\mathrm{sk}}^<} \cdot \mathcal{P}(\lambda,\sigma,k_0,\ell_0)$$

$$\times\, e^{\pi i \lambda t_{\mathfrak{n}_0}} \prod_{\mathfrak{n}\in(\mathcal{N}_{\mathrm{sk}}^<)^*} e^{\zeta_{\mathfrak{n}}\pi i\cdot\delta L^2\Omega_{\mathfrak{n}} t_{\mathfrak{n}}}\,\mathrm{d}t_{\mathfrak{n}}$$

$$\times \prod_{\mathfrak{m}\in(\mathscr{L}_{\mathrm{sk}}^<)^*}^{(+)} \mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}(t_{\mathfrak{m}^p},t_{(\mathfrak{m}')^p},k_{\mathfrak{m}})$$

$$\times \prod_{\mathfrak{m}\in(\mathcal{N}_{\mathrm{sk}}^<)^*} \mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}(t_{\mathfrak{m}^p},t_{\mathfrak{m}},k_{\mathfrak{m}}). \tag{5.10}$$

Here the sum is taken over all $(k_1,\ldots,k_{2R})$-decorations $\mathscr{I}_{\mathrm{sk}}^<$ of $\mathscr{G}_{\mathrm{sk}}^<$, and the other notations are all associated with $\mathscr{G}_{\mathrm{sk}}^<$, except $\widetilde{\mathscr{J}}_{\mathrm{sk}}^<$ and $\widetilde{\epsilon}_{\mathscr{I}_{\mathrm{sk}}^<}$; instead, for $\widetilde{\mathscr{J}}_{\mathrm{sk}}^<$ we add one extra condition $t_{\mathfrak{n}_0^p} > t_{\mathfrak{n}_0} + \sigma$ (where $\mathfrak{n}_0^p$ is the parent of $\mathfrak{n}_0$) to the original definition (4.2). As for $\widetilde{\epsilon}_{\mathscr{I}_{\mathrm{sk}}^<}$, in the "sg" case we remove one factor $\epsilon_{k_{\mathfrak{n}_{01}}k_{\mathfrak{n}_{02}}k_{\mathfrak{n}_{03}}}$ (where $\mathfrak{n}_{0j}$ are the children of $\mathfrak{n}_0$ from left to right) from the original definition (3.1), while in the "lg" case we set it to be the same as $\epsilon_{\mathscr{I}_{\mathrm{sk}}^<}$. Moreover, the variables (k_0,ℓ_0) are defined as in Definition 5.3, and the functions G and $\mathcal{P}$ etc. are as in (5.7); they satisfy either (5.8) or the alternative version in the "lg" case. We also insert the corresponding "sg" or "lg" cutoffs restricting to $0 < |h| \leq 1/(100\delta L)$ or $|h| > 1/(100\delta L)$ in (5.10). Finally, in the functions $\mathcal{K}^*_{\mathcal{T}(\mathfrak{n}_0)}$ and $\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$ for the leaf pair $\{\mathfrak{m},\mathfrak{m}'\}$ containing $\mathfrak{p}_0$, the input variable $t_{\mathfrak{n}_0}$ should be replaced by $t_{\mathfrak{n}_0} + \sigma$.

Remark 5.5. In the small gap case, due to the absence of $\epsilon_{k_{\mathfrak{n}_{01}}k_{\mathfrak{n}_{02}}k_{\mathfrak{n}_{03}}}$ in $\widetilde{\epsilon}_{\mathscr{I}_{\mathrm{sk}}^<}$, in the summation in (5.10), the decoration $(k_{\mathfrak{n}})$ may be resonant at the node $\mathfrak{n}_0$ (i.e. $(k_{\mathfrak{n}_{01}},k_{\mathfrak{n}_{02}},k_{\mathfrak{n}_{03}}) \notin \mathfrak{S}$, see (2.6)), but it must not be resonant at any other branching node. This resonance may lead to an (at most) L^{4d} loss in the counting estimates in Proposition 6.8, but this can always be covered by the L^{-40d} gain from $\mathcal{P}$ in (5.8). See Remark 6.9 for further explanation.

5.4. Summary

Now we may repeat the reduction described above for every irregular chain $\mathcal{H}^\circ$ in $\mathscr{G}_{\mathrm{sk}}$, noticing that these irregular chains do not affect each other, in the same way as in [12, Section 8.4]. Let $\mathscr{G}_{\mathrm{sk}}^\#$ be the marked garden obtained by removing all the irregular chains $\mathcal{H}^\circ$ from $\mathscr{G}_{\mathrm{sk}}$ as described above in Section 5.3. This does not depend on the choice of $\mathscr{G}_{\mathrm{sk}}$ in

the fixed congruence class, nor on the choice of $\mathcal{G} \in \mathcal{F}$. We then have

$$(5.4) = (C^+\delta)^{m_1} \left(\frac{\delta}{2L^{d-1}} \right)^{m'_0} \zeta^*(\mathcal{G}^{\#}_{\mathrm{sk}})$$

$$\times \int_{\mathbb{R}^{\Xi}} G(\boldsymbol{\lambda}) \, \mathrm{d}\boldsymbol{\lambda} \int_{[0,1]^{\Xi}} \mathrm{d}\boldsymbol{\sigma} \sum_{\mathscr{I}^{\#}_{\mathrm{sk}}} \int_{\tilde{\mathcal{J}}^{\#}_{\mathrm{sk}}} \epsilon_{\mathscr{I}^{\#}_{\mathrm{sk}}} \mathcal{P}(\boldsymbol{\lambda}, \boldsymbol{\sigma}, k[\mathcal{G}^{\#}_{\mathrm{sk}}]) \prod_{\mathfrak{n} \in \Xi} e^{\pi i \lambda_{\mathfrak{n}} t_{\mathfrak{n}}}$$

$$\times \prod_{\mathfrak{n} \in (\mathcal{N}^{\#}_{\mathrm{sk}})^*} e^{\zeta_{\mathfrak{n}} \pi i \cdot \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \, \mathrm{d}t_{\mathfrak{n}} \prod_{\mathfrak{m} \in (\mathcal{L}^{\#}_{\mathrm{sk}})^*}^{(+)} \mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}(t_{\mathfrak{m}^p}, t_{(\mathfrak{m}')^p}, k_{\mathfrak{m}})$$

$$\times \prod_{\mathfrak{m} \in (\mathcal{N}^{\#}_{\mathrm{sk}})^*} \mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}(t_{\mathfrak{m}^p}, t_{\mathfrak{m}}, k_{\mathfrak{m}}). \tag{5.11}$$

Here m'_0 is the scale of $\mathcal{G}^{\#}_{\mathrm{sk}}$ and m_1 is the sum of all the m_{tot} and q in (5.10), the summation is taken over all k-decorations $\mathscr{I}^{\#}_{\mathrm{sk}}$ of $\mathcal{G}^{\#}_{\mathrm{sk}}$, and the other notations are all associated with $\mathcal{G}^{\#}_{\mathrm{sk}}$, except $\tilde{\mathcal{J}}^{\#}_{\mathrm{sk}}$; instead, for $\tilde{\mathcal{J}}^{\#}_{\mathrm{sk}}$ we add the extra conditions $t_{\mathfrak{n}^p} > t_{\mathfrak{n}} + \sigma_{\mathfrak{n}}$ (where $\mathfrak{n}^p$ is the parent of $\mathfrak{n}$) to the original definition (4.2), for $\mathfrak{n} \in \Xi$, where Ξ is a subset of the set $(\mathcal{N}^{\#}_{\mathrm{sk}})^*$ of branching nodes. The vector parameters are $\boldsymbol{\lambda} = \lambda[\Xi] \in \mathbb{R}^{\Xi}$ and $\boldsymbol{\sigma} = \sigma[\Xi] \in [0,1]^{\Xi}$ respectively, and $k[\mathcal{G}^{\#}_{\mathrm{sk}}]$ is the vector of all the $k_{\mathfrak{n}}$'s. The functions $G(\boldsymbol{\lambda})$ and $\mathcal{P}(\boldsymbol{\lambda}, \boldsymbol{\sigma}, k[\mathcal{G}^{\#}_{\mathrm{sk}}])$ satisfy the bounds

$$\left\| \prod_{\mathfrak{n} \in \Xi} \langle \lambda_{\mathfrak{n}} \rangle^{1/18} G \right\|_{L^1} \lesssim (C^+)^m, \qquad \sup_{\boldsymbol{\lambda}, k[\mathcal{Q}^{\#}_{\mathrm{sk}}]} \int_{[0,1]^{\Xi}} |\mathcal{P}(\boldsymbol{\lambda}, \boldsymbol{\sigma}, k[\mathcal{Q}^{\#}_{\mathrm{sk}}])| \, \mathrm{d}\boldsymbol{\sigma} \lesssim 1. \tag{5.12}$$

We also insert various small gap or large gap cutoff functions, and some input variables in some of the $\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$ or $\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}$ functions may be translated by some $\sigma_{\mathfrak{n}}$, in the same way as in (5.10). Finally, the function $\epsilon_{\mathscr{I}^{\#}_{\mathrm{sk}}}$ may miss a few $\epsilon_{k_{\mathfrak{n}} k_{\mathfrak{n}_1} k_{\mathfrak{n}_2} k_{\mathfrak{n}_3}}$ factors compared to the original definition (3.1), but for each such missing factor we can gain a power L^{-40d} on the right hand side in the second inequality in (5.12).

At this point, we may expand the functions $\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$ and $\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}$ (or their leading or remainder contributions) using their Fourier L^1 (or $X^{\kappa}_{\mathrm{loc}}$) bounds, and combine the $\mathcal{K}$ factors and the $\mathcal{P}$ factor in (5.11) to further reduce to the expression

$$(5.4) = (C^+\delta)^{\frac{m-m'_0}{2}} \left(\frac{\delta}{2L^{d-1}} \right)^{m'_0} \zeta^*(\mathcal{G}^{\#}_{\mathrm{sk}})$$

$$\times \int_{\mathbb{R}^{\Lambda} \times \mathbb{R}^2} G(\boldsymbol{\lambda}) \, \mathrm{d}\boldsymbol{\lambda} \cdot e^{\pi i (\lambda t + \mu s)} \int_{[0,1]^{\Xi}} \mathrm{d}\boldsymbol{\sigma} \sum_{\mathscr{I}^{\#}_{\mathrm{sk}}} \int_{\tilde{\mathcal{J}}^{\#}_{\mathrm{sk}}} \epsilon_{\mathscr{I}^{\#}_{\mathrm{sk}}}$$

$$\times \prod_{\mathfrak{n} \in \Lambda} e^{\zeta_{\mathfrak{n}} \pi i \cdot \delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}} \prod_{\mathfrak{n} \in \Lambda} e^{\pi i \lambda_{\mathfrak{n}} t_{\mathfrak{n}}} \, \mathrm{d}t_{\mathfrak{n}} \cdot \mathcal{X}_{\mathrm{tot}}(\boldsymbol{\lambda}, \boldsymbol{\sigma}, k[\mathcal{G}^{\#}_{\mathrm{sk}}]). \tag{5.13}$$

Here the set $\Lambda = (\mathcal{N}^{\#}_{\mathrm{sk}})^*$ and $\boldsymbol{\lambda} = (\lambda[\Lambda], \lambda, \mu) \in \mathbb{R}^{\Lambda} \times \mathbb{R}^2$, the function G is different from the one in (5.11), but still satisfies the same first inequality in (5.12) with weights in λ and μ also included. Using the second inequality in (5.12), the $X^{\kappa}_{\mathrm{loc}}$ bounds for $\mathcal{K}_{\mathcal{Q}(\mathfrak{m},\mathfrak{m}')}$

and $\mathcal{K}^*_{\mathcal{T}(\mathfrak{m})}$ and their components, and the definition of markings $\mathfrak{L}$ and $\mathfrak{R}$, we deduce that the function $\mathcal{X}_{\text{tot}}$ satisfies

$$\int_{[0,1]^\Xi} |\mathcal{X}_{\text{tot}}(\boldsymbol{\lambda}, \boldsymbol{\sigma}, k[\mathcal{G}^\#_{\text{sk}}])| \, d\boldsymbol{\sigma} \lesssim \prod_{\mathfrak{l} \in (\mathscr{L}^\#_{\text{sk}})^*}^{(+)} \langle k_\mathfrak{l} \rangle^{-40d} \cdot L^{-2\nu r_0} \tag{5.14}$$

uniformly in $\boldsymbol{\lambda}$, where r_0 is the total number of branching nodes and leaf pairs that are marked $\mathfrak{R}$ in the marked garden $\mathcal{G}^\#_{\text{sk}}$. In (5.14) we can also gain a power L^{-40d} per missing factor $\epsilon_{k_{\mathfrak{n}_1} k_{\mathfrak{n}_2} k_{\mathfrak{n}_3}}$ in $\epsilon_{\mathscr{I}^\#_{\text{sk}}}$, as described above.

Note that the garden $\mathcal{G}^\#_{\text{sk}}$ is still mixed, and prime. Moreover, by definition, it *does not contain an irregular chain of length* $> 10^3 d$ *with all branching nodes and leaf pairs marked* $\mathfrak{L}$. In particular, if r_0 is the number of branching nodes and leaf pairs that are marked $\mathfrak{R}$, r_{irr} is the number of maximal irregular chains, and Q is the total length of these irregular chains, then we have

$$Q \leq C(r_0 + r_{\text{irr}}). \tag{5.15}$$

Based on this information, as well as the first inequality in (5.12) and (5.14), we will establish an *absolute upper bound* for the expression (5.13). This will be done in the following two sections.

6. Gardens and molecules

Definition 6.1 ([12, Definition 9.1]). A *molecule* $\mathbb{M}$ is a directed graph, formed by vertices (called *atoms*) and edges (called *bonds*), where multiple and self-connecting bonds are allowed. We will write $v \in \mathbb{M}$ and $\ell \in \mathbb{M}$ for atoms v and bonds ℓ in $\mathbb{M}$; we also write $\ell \sim v$ if v is an endpoint of ℓ. We further require that (i) each atom has at most two outgoing bonds and at most two incoming bonds (a self-connecting bond counts as outgoing once and incoming once), and (ii) there is no *saturated* (connected) component, where connectedness is always understood in terms of undirected graphs, and a component is saturated if it contains only degree 4 atoms. For a molecule $\mathbb{M}$ we define V to be the number of atoms, E the number of bonds and F the number of components. Define $\chi := E - V + F$.

Definition 6.2 ([12, Definition 9.3]). Given a garden $\mathcal{G}$, define the molecule $\mathbb{M}$ associated with $\mathcal{G}$, as follows. The atoms of $\mathcal{G}$ are all the 4-element subsets formed by a branching node in $\mathfrak{n} \in \mathcal{N}^*$ and its three children nodes. For any two atoms, we connect them by a bond if either (i) a branching node is the parent in one atom and a child in the other, or (ii) two leaves from these two atoms are paired with each other. We call this bond a *PC* (parent-child) *bond* in case (i) and an *LP* (leaf-pair) *bond* in case (ii). Note that multiple bonds are possible, and a self-connecting bond occurs when two sibling leaves are paired.

We fix a direction of each bond as follows. If a bond corresponds to a leaf pair, then it goes from the atom containing the leaf with a $-$ sign to the atom containing the leaf with a $+$ sign. If a bond corresponds to a branching node $\mathfrak{n}$ that is not a root, suppose $\mathfrak{n}$ is the
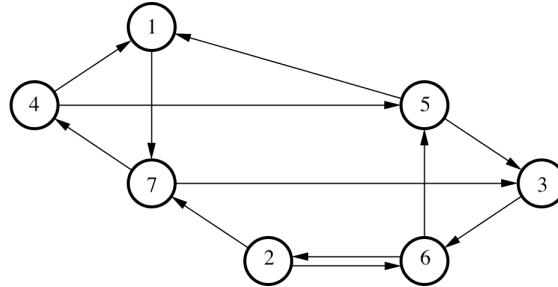
Fig. 6. The molecule associated with the garden $\mathcal{G}$ in Figure 4. Here atoms 1–4 correspond to branching nodes k_1–k_4, and atoms 5–7 correspond to branching nodes ℓ_1–ℓ_3 in $\mathcal{G}$.

parent in the atom v_1 and is a child in the atom v_2, then the bond goes from v_1 to v_2 if $\mathfrak{n}$ has a $+$ sign, and go from v_2 to v_1 otherwise. See Figure 6 for an example.

Proposition 6.3. *Let $\mathcal{G}$ be a* mixed *garden of width $2R$ and scale m. Then for the molecule $\mathbb{M}$ associated with $\mathcal{G}$ as in Definition 6.2, we have $\chi \leq m - R/2$.*

Proof. Let $\mathcal{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R})$. By definition of mixed gardens we know that no $\mathcal{T}_i$ and $\mathcal{T}_j$ have their leaves completely paired. For the molecule $\mathbb{M}$, clearly the number V of atoms is m, since each atom in $\mathbb{M}$ corresponds to a unique branching node in $\mathcal{G}$. Moreover, the number E of bonds is $2m - R$. This is because each bond corresponds to either a unique nonroot leaf pair or a nonroot branching node. The total number of leaf pairs and branching nodes (including roots) is $(m + R) + m = 2m + R$, but each root should be subtracted once (it should be excluded from the set of branching nodes if it is a branching node, and should be excluded from the set of leaf pairs if it is a leaf and is paired to another leaf), and only once (because there do not exist two roots that are both leaves and are paired to each other). This implies $E = 2m - R$ as there are $2R$ roots.

Finally, for any $\mathcal{T}_j$ $(1 \leq j \leq 2R)$, let S_j be the set of atoms corresponding to branching nodes in $\mathcal{T}_j$. Then $\mathbb{M}$ is the union of all S_j $(1 \leq j \leq 2R)$. By definition all atoms in S_j are connected to each other. Moreover, if some leaf in $\mathcal{T}_j$ is paired to some leaf in $\mathcal{T}_{j'}$ then S_j and $S_{j'}$ are also connected to each other. Since the leaves in the union of any *odd* number of trees $\mathcal{T}_j$ cannot all be paired with each other (since each $\mathcal{T}_j$ has an odd number of leaves), and also that the garden does not contain *two* trees $\mathcal{T}_i$ and $\mathcal{T}_j$ with their leaves completely paired, we know that any connected component in $\mathbb{M}$ must be the union of at least *four* S_j, in particular $F \leq R/2$. This implies that $\chi = E - V + F \leq m - R/2$. $\blacksquare$

Proposition 6.4. *Fix m and R. Given any molecule $\mathbb{M}$ of m atoms, the number of gardens $\mathcal{G}$ of width $2R$ and scale m that corresponds to $\mathbb{M}$ in the sense of Definition 6.2 is at most $(CR)!C^m$.*

Proof. This is basically the same as [12, Proposition 9.6]. For each atom $v \in \mathbb{M}$, each bond $\ell \sim v$ corresponds to a unique node $\mathfrak{n}$ in the 4-node subset corresponding to v. We

may assign a code to this pair (v, ℓ) indicating the relative position of $\mathfrak{n}$ in this subset (say code 0 if $\mathfrak{n}$ is the parent node, and codes 1, 2 or 3 if $\mathfrak{n}$ is the left, mid or right child node). In this way we get an encoded molecule which has a code assigned to each pair (v, ℓ) where $\ell \sim v$. Clearly if $\mathbb{M}$ is fixed then the corresponding encoded molecule has at most C^{m+R} possibilities, so it suffices to reconstruct $\mathcal{G}$ from the encoded molecule.

In fact, if the encoded molecule is fixed, then the branching nodes of $\mathcal{G}$ uniquely correspond to the atoms of $\mathbb{M}$. Moreover, the branching node corresponding to v_2 is the α-th child of the branching node corresponding to v_1, if and only if v_1 and v_2 are connected by a bond ℓ such that the codes of (v_1, ℓ) and (v_2, ℓ) are α and 0 respectively. Next, we can determine the leaves of $\mathcal{G}$ by putting a leaf as the α-th child for each branching node and each α, as long as this position is not occupied by another branching node; moreover, the α-th child of the branching node corresponding to v_1 and the β-th child of the branching node corresponding to v_2 are paired if and only if v_1 and v_2 are connected by a bond ℓ such that the codes of (v_1, ℓ) and (v_2, ℓ) are α and β respectively.

Finally, note that a node $\mathfrak{n}$ is a root if and only if it is not a child of any other node, so we can uniquely identify the roots of the trees. Permuting these $2R$ roots leads to at most $(CR)!$ choices, and once a permutation is fixed, the garden $\mathcal{G}$ will also be fixed as the structure of each tree, as well as the leaf pairing structure, has been fixed as above. This gives at most $(CR)!C^m$ possible choices for $\mathcal{G}$. Note that if one of the trees in $\mathcal{G}$ is trivial, then the reconstruction will be slightly different, but this does affect the result. ∎

Definition 6.5 ([12, Definition 9.7]). We define the *type I* and *type II (molecular) chains* in a molecule $\mathbb{M}$, as in Figure 7. Note that type I chains are formed by double bonds, and type II chains are formed by double bonds and pairs of single bonds. For type I chains, we require that the two bonds in any double bond have opposite directions. For type II chains, we require that any pair of single bonds have opposite directions; see Figure 7.
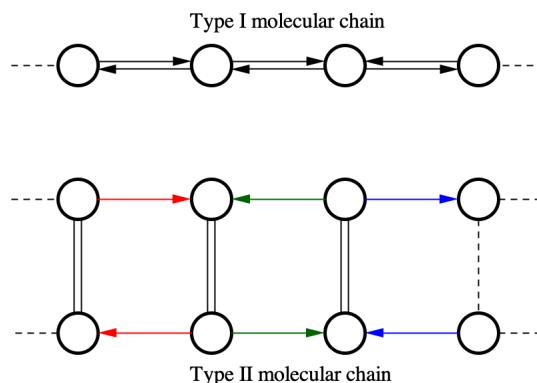

Fig. 7. The two types of molecular chains. For type II, the single bonds of the same color are paired single bonds, and must have opposite directions.

Given a molecule $\mathbb{M}$, the main subject of this section is the following counting problem associated with $\mathbb{M}$, similar to [12].

Definition 6.6 ([12, Definition 9.8]). Given a molecule $\mathbb{M}$ and a set S of atoms. Suppose we fix (i) $a_\ell \in \mathbb{Z}_L^d$ for each bond $\ell \in \mathbb{M}$, (ii) $c_v \in \mathbb{Z}_L^d$ for each (nonisolated, same below) atom $v \in \mathbb{M}$, assuming $c_v = 0$ if v has degree 4, (iii) $\Gamma_v \in \mathbb{R}$ for each atom v, and (iv) $f_v \in \mathbb{Z}_L^d$ for each $v \in S$ with $d(v) < 4$. Define $\mathfrak{D}(\mathbb{M})$ to be the set of vectors $k[\mathbb{M}] := (k_\ell)_{\ell \in \mathbb{M}}$ such that each k_ℓ is in $\mathbb{Z}_L^d$ and $|k_\ell - a_\ell| \leq 1$, and

$$\sum_{\ell \sim v} \zeta_{v,\ell} k_\ell = c_v, \quad \left| \sum_{\ell \sim v} \zeta_{v,\ell} |k_\ell|_\beta^2 - \Gamma_v \right| \leq \delta^{-1} L^{-2} \tag{6.1}$$

for each atom v. Here the sum is taken over all bonds $\ell \sim v$, and $\zeta_{v,\ell}$ equals 1 if ℓ is outgoing from v, and equals -1 otherwise. We also require that (a) the values of k_ℓ for different $\ell \sim v$ are all equal given each $v \in S$, and this value equals f_v if also $d(v) < 4$, and (b) for any $v \notin S$ and any bonds $\ell_1, \ell_2 \sim v$ of opposite directions (viewed from v), we have $k_{\ell_1} \neq k_{\ell_2}$. Note that this actually makes $\mathfrak{D}$ depending on S, but we will omit this dependence for simplicity. We say an atom v is *degenerate* if $v \in S$, and is *tame* if moreover $d(v) < 4$.

In addition, we may add some extra conditions to the definition of $\mathfrak{D}(\mathbb{M})$. These conditions are independent of the parameters, and have the form of (combinations of) $(k_{\ell_1} - k_{\ell_2} \in E)$ for some bonds $\ell_1, \ell_2 \in \mathbb{M}$ and fixed subsets $E \subset \mathbb{Z}_L^d$. Let Ext be the set of these extra conditions, and denote the corresponding set of vectors $k[\mathbb{M}]$ by $\mathfrak{D}(\mathbb{M}, \text{Ext})$. We are interested in the quantities $\sup \# \mathfrak{D}(\mathbb{M}, \text{Ext})$, where the supremum is taken over all possible choices of parameters $(a_\ell, c_v, \Gamma_v, f_v)$.

Remark 6.7. The vectors $k[\mathbb{M}]$ will come from decorations of the garden $\mathcal{G}$ from which $\mathbb{M}$ is obtained. In fact, if $k[\mathcal{G}]$ is a $(k_1, \ldots, k_{2R})$-decoration of $\mathcal{G}$, then it uniquely corresponds to a vector $k[\mathbb{M}]$. Let $v \in \mathbb{M}$ be an atom corresponding to a branching node $\mathfrak{n} \in \mathcal{G}$. Then $d(v) = 4$ unless $\mathfrak{n}$ is the root of some $\mathcal{T}_j$, or some other $\mathcal{T}_i$ is a trivial tree paired with a child of $\mathfrak{n}$ (there may be more than one such i).

It is easy to check, using Definitions 3.3 and 6.6, that the following holds. If $d(v) = 4$ then $\sum_{\ell \sim v} \zeta_{v,\ell} k_\ell = 0$, and $\sum_{\ell \sim v} \zeta_{v,\ell} |k_\ell|_\beta^2 = -\zeta_\mathfrak{n} \Omega_\mathfrak{n}$. If $d(v) < 4$, then the right hand sides of the above equations should be corrected by suitable algebraic sums of k_j and (or) k_i, and $|k_j|_\beta^2$ and (or) $|k_i|_\beta^2$, where j and i are associated with $\mathfrak{n}$ as stated above. Note that all these k_j and k_i are fixed when considering the decoration $k[\mathcal{G}]$. Moreover, if $(k_{\mathfrak{n}_1}, k_{\mathfrak{n}_2}, k_{\mathfrak{n}_3}) \in \mathfrak{S}$, then either the values of k_ℓ for different $\ell \sim v$ are all equal (and this value equals k_j if $d(v) < 4$ where j is as above), or for any bonds $\ell_1, \ell_2 \sim v$ of opposite directions we have $k_{\ell_1} \neq k_{\ell_2}$. Note that a degenerate atom corresponds exactly to a branching node $\mathfrak{n}$ for which $\epsilon_{k_{\mathfrak{n}_1} k_{\mathfrak{n}_2} k_{\mathfrak{n}_3}} = -1$.

Proposition 6.8. *Let $\mathbb{M}$ be the molecule associated with a mixed garden $\mathcal{G}$ of width $2R$ and scale m, where $R, m \leq (\log L)^{20}$. Suppose also that $\mathbb{M}$ does not contain any triple bond. Then $\mathfrak{D}(\mathbb{M})$ is the union of at most C^m subsets. Each subset has the form $\mathfrak{D}(\mathbb{M}, \text{Ext})$, and there exist $0 \leq r \leq m$ and a collection of at most $C(r + R)$ molecular*

chains of either type I or type II in $\mathbb{M}$ *such that* (i) *the number of atoms not in one of these chains is at most* $C(r + R)$, *and* (ii) *for any type II chain in the collection and any two paired single bonds* (ℓ_1, ℓ_2) *in this chain* (*see Figure* 7), *the set* Ext *includes the condition* $(k_{\ell_1} = k_{\ell_2})$. *Moreover,*

$$\sup \# \mathfrak{D}(\mathbb{M}, \text{Ext}) \leq (C^+)^m \delta^{-\frac{m+m_1}{2}} L^{(d-1)(m-R/2)-2vr}, \tag{6.2}$$

where m_1 *is the number of atoms in the union of type I chains.*

Remark 6.9. In view of Remark 5.5, in Definition 6.6 we may also fix some set S^* of atoms such that neither (a) nor (b) is required for $v \in S^*$, but we are allowed to multiply the left hand side of (6.2) by $L^{-40d \cdot |S^*|}$. In this way we can restate Proposition 6.8 appropriately, and the new result can be easily proved with little difference in the arguments, due to the large power gains. For simplicity we will not include this in the proof below.

Proof of Proposition 6.8. The proof is basically the same as the proof of [12, Proposition 9.10]. We define the same steps as in [12, Section 9.3], including the good and normal steps, and apply the same algorithm as in [12, Section 9.4]. Let the total number of good steps in the process be $r \geq 0$ (we may assume $r \leq m$ up to a constant because the total number of steps is at most $O(m)$), then we may repeat the proofs in [12, Section 9.5]. The only difference here is the initial state of the molecule (as $\mathbb{M}$ is obtained from a mixed garden rather than a couple), but in the current case we still have $V_{<4} + F = O(R)$, where $V_{<4}$ and F are the number of atoms with degree < 4 and the number of connected components and the constant in O depends only on d.

Note that in [12, proof of Proposition 9.10], the quantities that are monitored include V, E, F, V_j for $0 \leq j \leq 4$ (the number of atoms with degree j), V_2^* (the number of degree 2 atoms with two single bonds), and ξ (the number of "special bonds" connecting two degree 3 atoms that have a special form [12, Definition 9.12]). Since $V_{<4} + F = O(R)$, it is clear that at the beginning, the value of each of these quantities in the current case is the same as in [12], up to errors of size $O(R)$. Thus, the same proof as in [12] shows that $\mathbb{M}$ contains at most $C(R + r)$ type I or II molecular chains, such that the number of atoms not in one of these chains is at most $C(R + r)$. Moreover,

$$\sup \# \mathfrak{D}(\mathbb{M}, \text{Ext}) \leq (C^+)^m \delta^{-\kappa} L^{(d-1)\gamma},$$

where κ and γ are calculated retrospectively from the algorithm, as described in [12, Section 9.2]. The calculation for κ is the same up to $O(r)$ errors, so we have $\kappa = \frac{m+m_1}{2}$ up to errors $O(R + r)$ which are acceptable. To calculate γ, note that in [12] we are actually calculating $\gamma - \chi$, and the same proof yields $(d - 1)(\gamma - \chi) \leq -2vr$ for the initial molecule. Now by Proposition 6.3 we know $\chi \leq m - R/2$, hence

$$\sup \# \mathfrak{D}(\mathbb{M}, \text{Ext}) \leq (C^+)^m \delta^{-\frac{m+m_1}{2}} L^{(d-1)(m-R/2)-2vr},$$

as desired.   ∎

7. L^1 coefficient bounds

We now return to the study of the expression (5.13). Let $\mathscr{G}_{sk}^{\#}$ and (r_0, r_{irr}) be as in Section 5.4. For simplicity, in this section we will write $\mathscr{G}_{sk}^{\#}$ simply as $\mathscr{G}$, and the associated sets $(\mathscr{N}_{sk}^{\#})^*$ as $\mathscr{N}^*$ etc. Recall, by (5.15), that the total length of the irregular chains in $\mathscr{G}$ is at most $C(r_0 + r_{irr})$. Let Ξ be a subset of $\mathscr{N}^*$. We may define, as in (5.13), the function

$$\mathcal{U}_{\mathscr{G}}(t, s, \boldsymbol{\sigma}, \alpha[\mathscr{N}^*]) = \int_{\widetilde{\mathscr{I}}} \prod_{\mathfrak{n} \in \mathscr{N}^*} e^{\pi i \alpha_{\mathfrak{n}} t_{\mathfrak{n}}} \, dt_{\mathfrak{n}}, \tag{7.1}$$

where $\boldsymbol{\sigma} = \sigma[\Xi] \in [0, 1]^{\Xi}$, and the domain $\widetilde{\mathscr{I}}$ is defined as in (4.2), but with the extra conditions $t_{\mathfrak{n}^p} > t_{\mathfrak{n}} + \sigma_{\mathfrak{n}}$ for $\mathfrak{n} \in \Xi$, where $\mathfrak{n}^p$ is the parent of $\mathfrak{n}$. Let m_0' be the scale of $\mathscr{G}$. We can write

$$(5.13) = (C^+ \delta)^{\frac{m - m_0'}{2}} \left(\frac{\delta}{2L^{d-1}} \right)^{m_0'} \zeta^*(\mathscr{G}) \int_{\mathbb{R}^{\mathscr{N}^*} \times \mathbb{R}^2} G(\lambda) \cdot e^{\pi i (\lambda t + \mu s)} \, d\lambda \int_{[0,1]^{\Xi}} d\sigma$$

$$\times \sum_{\mathscr{I}} \epsilon_{\mathscr{I}} \mathcal{U}_{\mathscr{G}} \left(t, s, \boldsymbol{\sigma}, (\delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}})_{\mathfrak{n} \in \mathscr{N}^*} \right) \cdot \mathcal{X}_{tot}(\lambda, \boldsymbol{\sigma}, k[\mathscr{G}]). \tag{7.2}$$

Let $\mathbb{M}$ be the molecule associated with $\mathscr{G}$ as in Definition 6.2. It is easy to see that $\mathbb{M}$ contains no triple bond, as triple bonds in $\mathbb{M}$ can only come from $(1, 1)$-mini couples and mini trees (as in Definition 3.4) in $\mathscr{G}$. By the proofs in Section 6, we can introduce at most $C^{m_0'}$ sets Ext of extra conditions such that the summation in $\mathscr{I} = k[\mathscr{G}]$ in (7.2) can be decomposed into the summations with each of these sets of extra conditions imposed on $k[\mathscr{G}]$. Moreover, for each choice of Ext there is $1 \leq r_1 \leq n_0'$ such that the conclusion of Proposition 6.8, including (6.2), holds true (with r replaced by r_1).

Notice that a type I chain in $\mathbb{M}$ can *only* be obtained from either one irregular chain, or the union of two irregular chains in $\mathscr{G}$; for couples this can be proved in the same way as in [12, Section 10.1.2] (which deals with type II chains), and the same proof works also for gardens. Therefore, the total length p of type I chains in $\mathbb{M}$ is bounded by the total length of irregular chains in $\mathscr{G}$, which is at most $C(r_0 + r_{irr})$. However, each irregular chain in $\mathscr{G}$ also corresponds to a type I chain in the base molecule, so $r_{irr} \leq C(r_1 + R)$, hence $p \leq C(r + R)$, where $r = r_0 + r_1$. This means the number of atoms in $\mathbb{M}$ that are not in one of the (at most $C(r + R)$) type II chains is at most $C(r + R)$.

Now, suppose $\mathfrak{n}$ and $\mathfrak{n}'$ are two branching nodes in $\mathscr{G}$ which correspond to two atoms in $\mathbb{M}$ *that are connected by a double bond in a type II chain*. Then we must have $\zeta_{\mathfrak{n}'} \Omega_{\mathfrak{n}'} = -\zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}}$ under the extra conditions in Ext (see Remark 6.7). In fact, we will restrict $\{\mathfrak{n}, \mathfrak{n}'\}$ to the *interior* of this type II chain by omitting five pairs of atoms at both ends of the chain, in the same way as in Definition 5.4. Then, we make such $\{\mathfrak{n}, \mathfrak{n}'\}$ a pair, and choose one node from each such pair to form a set $\widetilde{\mathscr{N}}^{ch}$. If it happens that one of $\{\mathfrak{n}, \mathfrak{n}'\}$ is a parent of the other, we assume the parent belongs to $\widetilde{\mathscr{N}}^{ch}$. Let $\mathscr{N}^{rm}$ be the set of branching nodes not in these pairs, and define $\widetilde{\mathscr{N}} = \widetilde{\mathscr{N}}^{ch} \cup \mathscr{N}^{rm}$.

We will be interested in estimates of the function $\mathcal{U}_{\mathscr{G}}$ in (7.1) where $\alpha_{\mathfrak{n}} = \delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}}$, which means that $\alpha_{\mathfrak{n}} + \alpha_{\mathfrak{n}'} = \mu_{\mathfrak{n}}$ for each $\mathfrak{n} \in \widetilde{\mathscr{N}}^{ch}$, where $\mathfrak{n}'$ is the

node paired to $\mathfrak{n}$ and $\mu_{\mathfrak{n}} = \lambda_{\mathfrak{n}} + \lambda_{\mathfrak{n}'}$ is a parameter depending on λ. Under this assumption on $\alpha_{\mathfrak{n}}$, we can write

$$\mathcal{U}_{\mathscr{G}}(t, s, \sigma, \alpha[\mathcal{N}^*]) = \mathcal{V}_{\mathscr{G}}(t, s, \sigma, \alpha[\widetilde{\mathcal{N}}]) \tag{7.3}$$

for some function $\mathcal{V}_{\mathscr{G}}$. This function actually also depends on the parameters $\mu_{\mathfrak{n}}$ for $\mathfrak{n} \in \widetilde{\mathcal{N}}^{\mathrm{ch}}$, but we will omit this for notational convenience. We then have to prove the following.

Proposition 7.1. *Suppose $\mathscr{G}$ has scale m_0'. For each $\mathfrak{n} \in \widetilde{\mathcal{N}}$, suppose that $S_{\mathfrak{n}} \subset \mathbb{Z}$ and $\#S_{\mathfrak{n}} \leq L^{10d}$. Then, uniformly in (t, s), in the choices of $(S_{\mathfrak{n}})_{\mathfrak{n} \in \widetilde{\mathcal{N}}}$, and in the parameters $(\mu_{\mathfrak{n}})_{\mathfrak{n} \in \widetilde{\mathcal{N}}^{\mathrm{ch}}}$, we have*

$$\delta^{m_0'/4} \cdot \sum_{(m_{\mathfrak{n}}): m_{\mathfrak{n}} \in S_{\mathfrak{n}}} \sup_{(\alpha_{\mathfrak{n}}): |\alpha_{\mathfrak{n}} - m_{\mathfrak{n}}| \leq 1} \sup_{\sigma} |\mathcal{V}_{\mathscr{G}}(t, s, \sigma, \alpha[\widetilde{\mathcal{N}}])|$$

$$\leq (C^+)^{m_0'} L^{C(r+R)\sqrt{\delta}} (\log L)^{C(r+R)}, \tag{7.4}$$

where $r = r_0 + r_1$.

Proof. The proof is exactly the same as in [12, proof of Proposition 10.1], except that all the "couples" should be replaced by "gardens". The reason is that the proof in [12] goes by induction; moreover, in each inductive step we remove either two branching nodes or one chain containing modules A and B (see [12, Sections 10.1, 10.2]). In either case this step involves at most two trees and the other trees are not affected, so the proof is the same for couples and for general gardens.

In the proof in [12] we have also introduced the simpler structures (for the purpose of induction) of unsigned couples and double trees, which are naturally replaced by unsigned gardens and multi-trees (the collection of $2R$ trees with some branching nodes paired, compared to two trees in [12]). The rest of the proof is exactly the same. Note also that the exponents Cr in [12, Proposition 10.1] are replaced by $C(r + R)$ because the number of type II chains in $\mathbb{M}$, as well as the number of atoms not in one of these chains, is now $C(r + R)$ instead of Cr due to Proposition 6.8. ∎

8. Proof of Theorem 1.3

In this section we prove Propositions 4.7 and 4.8, which completes the proof of Theorem 1.3.

8.1. *Proof of Proposition 4.7*

Note that if $\mathscr{G}$ and $\mathscr{G}'$ are congruent in the sense of Definition 5.4, then they have the same width, the same signature, and the same scale for each of their component trees; moreover, $\mathscr{G}'$ is mixed if and only if $\mathscr{G}$ is mixed. Thus, the sum in (4.4) can be decomposed into different terms, where each term has the form (5.4) for one congruence class $\mathscr{F}$.

For any fixed $\mathscr{F}$, consider (5.4), which then equals (5.13) and (7.2). Note that in (7.2), $\mathscr{G}$ actually means $\mathscr{G}_{sk}^{\#}$ in our notation. Using the decay factors in (5.14) we can restrict to the subset where $|k_I - a_I| \leq 1$ ($\forall I \in (\mathscr{L}_{sk}^{\#})^{*}$) for some fixed parameters (a_I), with summability in (a_I) guaranteed. Using the first inequality in (5.12), we may also fix the value of $\boldsymbol{\lambda}$ (and hence $\mu_{\mathfrak{n}}$).

As in Section 6, by decomposing into at most $C^{m_0'}$ terms (where m_0' is the scale of $\mathscr{G}_{sk}^{\#}$), we can add the set $\mathtt{Ext}$ of extra conditions, which also defines the sets $\widetilde{\mathcal{N}}$ (as in Section 7), etc., and the value $r_1 \geq 1$. Let $r = r_0 + r_1$ as above. Then thanks to $\mathtt{Ext}$, we can use (7.3) to reduce $\mathcal{U}_{\mathscr{G}_{sk}^{\#}}$ to $\mathcal{V}_{\mathscr{G}_{sk}^{\#}}$. Moreover, for each $\mathfrak{n} \in \widetilde{\mathcal{N}}$, the value $\delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}}$ belongs to some subset of $\mathbb{R}$ of cardinality at most L^{3d}, as $k[\mathscr{G}_{sk}^{\#}]$ varies (because each $k_{\mathfrak{n}}$ belongs to a ball of radius at most $n \leq (\log L)^{20}$ under our assumptions). In particular, the value $m_{\mathfrak{n}} = \lfloor \delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}} \rfloor$ belongs to a set $S_{\mathfrak{n}} \subset \mathbb{Z}$ of cardinality at most L^{3d}, for all possible choices of $k[\mathscr{G}_{sk}^{\#}]$.

To estimate (7.2) with $\boldsymbol{\lambda}$ fixed, we first integrate in $\boldsymbol{\sigma}$. Using (5.14), we can estimate (7.2) using

$$\sum_{\mathscr{I}_{sk}^{\#}} |\epsilon_{\mathscr{I}_{sk}^{\#}}| \cdot \sup_{\boldsymbol{\sigma}} \left| \mathcal{V}_{\mathscr{G}_{sk}^{\#}}\left(t, s, \boldsymbol{\sigma}, (\delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}})_{\mathfrak{n} \in \widetilde{\mathcal{N}}}\right) \right|, \tag{8.1}$$

where $\mathscr{I}_{sk}^{\#} = k[\mathscr{G}_{sk}^{\#}]$ is a k-decoration of $\mathscr{G}_{sk}^{\#}$ (we also have additional factors that will be collected at the end). We next fix the values of $m_{\mathfrak{n}} \in S_{\mathfrak{n}}$ for each $\mathfrak{n}$; note that then

$$\sup_{\boldsymbol{\sigma}} \left| \mathcal{V}_{\mathscr{G}_{sk}^{\#}}\left(t, s, \boldsymbol{\sigma}, (\delta L^2 \zeta_{\mathfrak{n}} \Omega_{\mathfrak{n}} + \lambda_{\mathfrak{n}})_{\mathfrak{n} \in \widetilde{\mathcal{N}}}\right) \right| \leq \sup_{(\alpha_{\mathfrak{n}}): |\alpha_{\mathfrak{n}} - m_{\mathfrak{n}}| \leq 1} \sup_{\boldsymbol{\sigma}} |\mathcal{V}_{\mathscr{G}_{sk}^{\#}}(t, s, \boldsymbol{\sigma}, \alpha[\widetilde{\mathcal{N}}])|$$

by definition, so if we use (7.4) to sum over $(m_{\mathfrak{n}})$ in the end, we can further estimate (8.1) using

$$\sum_{\mathscr{I}_{sk}^{\#}} |\epsilon_{\mathscr{I}_{sk}^{\#}}| \cdot \prod_{I} \mathbf{1}_{|k_I - a_I| \leq 1} \prod_{\mathfrak{n}} \mathbf{1}_{|\Omega_{\mathfrak{n}} - b_{\mathfrak{n}}| \leq \delta^{-1} L^{-2}}, \tag{8.2}$$

where a_I and $b_{\mathfrak{n}}$ are constants, and we also include the conditions in $\mathtt{Ext}$. Now (8.2) is almost exactly the counting problem $\mathfrak{D}(\mathbb{M}, \mathtt{Ext})$ stated in Definition 6.6, due to Remark 6.7, except that we only assume $|k_I - a_I| \leq 1$ for *leaves* I. However, for any branching node $\mathfrak{n}$ there exists a child $\mathfrak{n}'$ of $\mathfrak{n}$ such that $k_{\mathfrak{n}} \pm k_{\mathfrak{n}'}$ belongs to a fixed ball of radius $\mu_{\mathfrak{n}}^{\circ}$ (with $\mu_{\mathfrak{n}}^{\circ}$ defined Lemma 3.6), so by using (3.3), one can reduce (8.2) to at most $C^{m_0'}$ counting problems, each of exactly the same form as $\mathfrak{D}(\mathbb{M}, \mathtt{Ext})$ in Definition 6.6. Therefore, (8.2) can be bounded using Proposition 6.8 (and Remark 6.9 if necessary). Collecting all the factors appearing in the above estimates, we get

$$|(5.4)| \leq (C^+)^m \delta^{(m-m_0')/2} \delta^{3m_0'/4} L^{-(d-1)m_0'} L^{-2\nu r_0}$$
$$\times L^{C(r+R)\sqrt{\delta}} (\log L)^{C(r+R)} \delta^{-(m_0'+q)/2} L^{(d-1)(m_0'-R/2)-2\nu r_1}, \tag{8.3}$$

which is then bounded by $(C^+ \delta^{1/4})^m L^{-3\nu(r+R)/2} \delta^{-q/2}$, where q is the total length of type I chains in the molecule obtained from $\mathbb{Q}_{sk}^{\#}$. We know $q \leq C(r + R)$ so $\delta^{-q/2} \leq L^{\nu(r+R)/2}$, which implies that

$$|(5.4)| \leq (C^+ \delta^{1/4})^m L^{-\nu(r+R)}. \tag{8.4}$$

Finally, suppose we fix r. Then the molecule associated with $\mathscr{G}_{\mathrm{sk}}^{\#}$ (see Definition 6.2) is, up to at most $C(r + R)$ remaining atoms, a union of at most $C(r + R)$ type II chains with total length at most m_0'. This clearly has at most $(C(r + R))! C^m$ possibilities. By Proposition 6.4, the number of choices for $\mathscr{G}_{\mathrm{sk}}^{\#}$ is also at most $(C(r + R))! C^m$. To form $\mathscr{G}_{\mathrm{sk}}$ from $\mathscr{G}_{\mathrm{sk}}^{\#}$ one needs to insert at most $C(r + R)$ irregular chains with total length at most m, which also has at most C^m possibilities. Finally, using Corollary 4.6, we see that $\mathscr{G}$ has at most $(C(r + R))! C^m$ choices. The number of choices for markings, as well as Ext, are also at most C^m and can be accommodated. This means that if we decompose (4.4) into terms of the form (5.4), then further decompose by markings and/or Ext etc., then each resulting term has an index $r \geq 0$ such that each term of index r is bounded by $(C^+ \delta^{1/4})^m L^{-\nu(r+R)}$ (see (8.4)), and the number of terms with index r is at most $(C(r + R))! C^m$. Therefore

$$(4.4) \leq \sum_{r=0}^{m} (C^+ \delta^{1/4})^m L^{-\nu(r+R)} \cdot (C(r + R))! C^m \leq (C^+ \delta^{1/4})^m L^{-\nu R}, \qquad (8.5)$$

because in any case $r + R$ is bounded by a power of $\log L$, which is $\ll L^\nu$. This completes the proof of Proposition 4.7.

8.2. Proof of Proposition 4.8

The proof is almost identical with the corresponding proofs in [12], which we briefly present here.

First, by Chebyshev's inequality, to prove (4.6) it suffices to show that

$$\mathbb{E} \left| \sup_{k,t} \langle k \rangle^{9d} (\mathscr{J}_n)_k(t) \right|^2 \leq (C^+ \sqrt{\delta})^n L^{100d} \qquad (8.6)$$

and

$$\mathbb{E} \left| \sup_{k,t} \langle k \rangle^{9d} \mathscr{R}_k(t) \right|^2 \leq (C^+ \sqrt{\delta})^N L^{100d}. \qquad (8.7)$$

Note that due to our choice $N = \lfloor (\log L)^4 \rfloor$ instead of $N = \lfloor \log L \rfloor$, the proof of (4.6) is conceptually easier than in [12] as we do not need the hypercontractivity property [12, Lemma A.3] or the higher moment estimates.

Now, to prove (8.6), we argue as in [12, proof of Proposition 12.1] (but with p replaced by 2), and apply the Gagliardo–Nirenberg inequality to bound the left hand side of (8.6), up to a multiple L^{10d}, by

$$\sup_{k,t} \langle k \rangle^{20d} \left(\mathbb{E} |(\mathscr{J}_n)_k(t)|^2 + \mathbb{E} |\partial_t (\mathscr{J}_n)_k(t)|^2 \right),$$

which is then bounded by $(C^+ \sqrt{\delta})^n L^{40d}$ in the same way as in [12]. In fact, the bound for $\mathbb{E} |(\mathscr{J}_n)_k(t)|^2$ is as in [12, Proposition 2.5] (again our choice $N = \lfloor (\log L)^4 \rfloor$ here does not affect the proof), while the bound for $\mathbb{E} |\partial_t (\mathscr{J}_n)_k(t)|^2$ is as in [12, (12.4)], which is proved by similar arguments. This settles (8.6). The proof of (8.7) is the same, except that $\mathscr{J}_n$ is replaced by $\mathscr{R}$ and n is replaced by N.

Finally, to prove (4.7), again by Chebyshev's inequality, it suffices to show that the kernel $(\mathscr{L}^n)_{k\ell}^{\zeta}(t,s)$ of the $\mathbb{R}$-linear operator $\mathscr{L}^n$ (with $\zeta \in \{\pm\}$ indicating the linear and conjugate linear parts) can be decomposed as

$$(\mathscr{L}^n)_{k\ell}^{\zeta}(t,s) = \sum_{n \leq m \leq N^3} (\mathscr{L}^n)_{k\ell}^{m,\zeta}(t,s)$$

and that for each $n \leq m \leq N^3$, the kernel $(\mathscr{L}^n)_{k\ell}^{m,\zeta}(t,s)$ satisfies

$$\mathbb{E}\left|\sup_{k,\ell} \sup_{0 \leq s < t \leq 1} \langle k-\ell \rangle^{9d} (\mathscr{L}^n)_{k\ell}^{m,\zeta}(t,s)\right|^2 \leq (C^+\sqrt{\delta})^m L^{100d}. \tag{8.8}$$

Now the decomposition is provided as in [12, Proposition 11.2], and (8.8) is proved as in [12, proof of Proposition 12.2] (in particular, this proof does require the hypercontractivity property). Note that in that proof, we actually further decompose $(\mathscr{L}^n)_{k\ell}^{m,\zeta}(t,s)$ into $(\mathscr{L}^n)_{M,k\ell}^{m,\zeta}(t,s)$ for dyadic M, and proves (8.8) for $(\mathscr{L}^n)_{M,k\ell}^{m,\zeta}$ with the right hand side involving a negative power of M (see [12, (12.10)]). Both proofs carry over to the current case with our choice $N = \lfloor (\log L)^4 \rfloor$ without any change, which then proves (4.7) and completes the proof of Proposition 4.8.

9. The non-Gaussian case

In this section we briefly discuss the non-Gaussian case, i.e. Theorem 1.4, which we prove in Sections 9.1 and 9.2 (Theorem 1.5 basically follows from it and is proved separately in Section 9.3). Since much of the proof will be identical with Theorem 1.3, we will only elaborate on the parts where the proofs are different.

First, in the Gaussian case our proof yields uniform estimates as long as $R \leq \log L$ (or $R \leq 2\log L$); here we will make slightly stronger assumptions

$$R \leq \frac{\log L}{(\log\log L)^2}.$$

Again we may consider $2R$ at some places, but it does not affect the result.

Next, using the expansion (3.6), we can reduce the proof of Theorem 1.4 to analyzing the correlations

$$\mathbb{E}\left(\prod_{j=1}^{2R} (\mathscr{J}_{\mathcal{T}_j})_{k_j}^{\zeta_j}(t)\right). \tag{9.1}$$

The rest of the proof, including the treatment of the remainder term b, can be done in the same way as with these correlations (see Section 8.2). Specifically, the Gaussian hypercontractivity inequality, which is used in Section 8.2, can be substituted by similar inequalities for the current density function thanks to the assumed bound on $\mu_r \leq (Cr)!$; an instance of such argument can be found in [11, Lemma 3.1] which treats the particular case of the uniform distribution on the unit circle, but the general case can be treated in the same manner. Therefore, below we will focus on the study of (9.1).

9.1. A substitute for Isserlis' theorem

The obvious difference in the study of (9.1) in the non-Gaussian case is that Isserlis' theorem is not available. Instead we have the following substitute:

Lemma 9.1. *Recall all the random variables η_k are i.i.d. with radial law, and $\mathbb{E}\,|\eta_k|^{2r} = \mu_r$ with $\mu_1 = 1$ and $\mu_r \leq (Cr)!$ for a positive integer C. Then, for any $k_j \in \mathbb{Z}_L^d$ $(1 \leq j \leq n)$ and $\zeta_j \in \{\pm\}$ $(1 \leq j \leq n)$ we have*

$$\mathbb{E}\Big(\prod_{j=1}^n \eta_{k_j}^{\zeta_j}\Big) = \sum_{\mathcal{O}} \lambda(\mathcal{O}) \cdot \sum_{\mathscr{O} \models \mathcal{O}} \prod_{A \in \mathscr{O}} \prod_{j,j' \in A} \mathbf{1}_{k_j = k_{j'}}. \tag{9.2}$$

Here $\mathcal{O}$ runs over all partitions of n into even *positive integers (in particular, if n is odd then the right hand side of (9.2) is zero). For fixed $\mathcal{O}$, the $\mathscr{O}$ runs over all* over-pairings *of the set $\{1, \ldots, n\}$ subordinate to $\mathcal{O}$ (we use the notation $\mathscr{O} \models \mathcal{O}$), which are partitions of $\{1, \ldots, n\}$ such that the cardinalities of the subsets exactly form the partition $\mathcal{O}$, and for each subset A exactly half of the signs ζ_j $(j \in A)$ are $+$ and half are $-$.*

The coefficients $\lambda(\mathcal{O})$ are constants depending only on $\mathcal{O}$ (and n), and $\lambda(2, \ldots, 2) = 1$; in general we have $\lambda(2, \ldots, 2, 2a_1, \ldots, 2a_r) = \lambda(2a_1, \ldots, 2a_r)$. Moreover, let q be the sum of the elements in $\mathcal{O}$ that are at least 4. Then $|\lambda(\mathcal{O})| \leq C_1^n n^{C_1 q}$ for some constant $C_1 \gg C$.

Proof. We may assume n is even and half the signs ζ_j $(1 \leq j \leq n)$ are $+$ and half are $-$, since otherwise both sides of (9.2) are zero. Denote by $|\mathcal{O}|$ the number of elements in $\mathcal{O}$ (counted with multiplicity). For two partitions $\mathcal{O}$ and $\mathcal{O}'$, we write $\mathcal{O}' \preceq \mathcal{O}$ if $\mathcal{O}'$ can be formed by further partitioning some elements in $\mathcal{O}$ into even integers (also define $\succeq$ and $\prec$ etc. accordingly). Similarly for set partitions $\mathscr{O}$ and $\mathscr{O}'$, we write $\mathscr{O}' \preceq \mathscr{O}$ if $\mathscr{O}'$ can be formed by further partitioning some subsets in $\mathscr{O}$ (still keeping half of the signs $+$ and half $-$ in each subset).

Now, for $\mathcal{O}' \preceq \mathcal{O}$, we define $\xi_{\mathcal{O},\mathcal{O}'}$ as follows: given a partition $\mathscr{O} \models \mathcal{O}$, consider the number of partitions $\mathscr{O}' \preceq \mathscr{O}$ such that $\mathscr{O}' \models \mathcal{O}'$. The number of choices for $\mathscr{O}'$ is independent of the choice of $\mathscr{O}$, and we define it to be $\xi_{\mathcal{O},\mathcal{O}'}$. Obviously $\xi_{\mathcal{O},\mathcal{O}} = 1$. We define the coefficients $\lambda(\mathcal{O})$ for each $\mathcal{O}$ so that they satisfy the following recurrence relation: first $\lambda(2, \ldots, 2) = 1$, and for each $\mathcal{O}$, we have

$$\prod_{2b \in \mathcal{O}} \mu_b = \sum_{\mathcal{O}' \preceq \mathcal{O}} \xi_{\mathcal{O},\mathcal{O}'} \cdot \lambda(\mathcal{O}'). \tag{9.3}$$

Here the product is taken over all elements $2b$ appearing in $\mathcal{O}$, counted with multiplicity. Clearly the values of of $\lambda(\mathcal{O})$ for each $\mathcal{O}$ are uniquely determined by (9.3). To prove $\lambda(2, \ldots, 2, 2a_1, \ldots, 2a_r) = \lambda(2a_1, \ldots, 2a_r)$, we simply notice $\mu_1 = 1$ and if $\mathcal{O}$ contains a certain number of terms 2, then any $\mathcal{O}' \preceq \mathcal{O}$ must contain at least the same number of 2's. Then we may proceed inductively using (9.3).

Next we prove (9.2) with $\lambda(\mathcal{O})$ defined by (9.3). Assume all different values of these k_j are m_i $(1 \leq i \leq r)$, where for each j there are a_i copies of m_i with corresponding

$\zeta_j = +$, and b_j copies with $\zeta_j = -$. We may assume $a_i = b_i$ (otherwise it is easy to check that both sides of (9.2) are zero), hence the left hand side of (9.2) is reduced to

$$\mathbb{E}\Big(\prod_{i=1}^{r} |\eta_{m_i}|^{2a_i}\Big) = \prod_{j=1}^{r} \mu_{a_i}.$$

Note also that $\mathcal{O}_* = (2a_1, \ldots, 2a_r)$ is a partition of n and the sets $\{j : k_j = m_i\}$ form a partition $\mathcal{O}_* \models \mathcal{O}_*$. Moreover, on the right hand side of (9.2), in order for the product $\prod_{A \in \mathcal{O}} \prod_{j,j' \in A} \mathbf{1}_{k_j = k_{j'}}$ to be nonzero, one must have $\mathcal{O} \preceq \mathcal{O}_*$ and $\mathcal{O} \preceq \mathcal{O}_*$, and in this case this product equals 1. Thus the right hand side of (9.2) equals

$$\sum_{\mathcal{O} \preceq \mathcal{O}_*} \lambda(\mathcal{O}) \sum_{\mathcal{O} \preceq \mathcal{O}_*,\, \mathcal{O} \not\models \mathcal{O}} 1 = \sum_{\mathcal{O} \preceq \mathcal{O}_*} \xi_{\mathcal{O}_*,\mathcal{O}} \cdot \lambda(\mathcal{O}) = \prod_{i=1}^{r} \mu_{a_i}$$

using the definition of $\xi_{\mathcal{O}_*,\mathcal{O}}$ and (9.3), as desired.

Next we prove that

$$|\lambda(\mathcal{O})| \leq \frac{(n/2)!}{|\mathcal{O}|!} \prod_{2b \in \mathcal{O}} (C_0 b)! \tag{9.4}$$

for $C_0 = C + 40$. The base case $\mathcal{O} = (2, \ldots, 2)$ is clear. By induction, and using $\mu_r \leq (Cr)!$, we only need to prove that for any $\mathcal{O}$,

$$\sum_{\mathcal{O}' \prec \mathcal{O}} \xi_{\mathcal{O},\mathcal{O}'} \frac{|\mathcal{O}|!}{|\mathcal{O}'|!} \frac{\prod_{2b \in \mathcal{O}'} (C_0 b)!}{\prod_{2b \in \mathcal{O}} (C_0 b)!} \leq \frac{1}{2}. \tag{9.5}$$

Now, fix a partition $\mathcal{O}$ of $\{1, \ldots, n\}$ subordinate to $\mathcal{O}$. To construct $\mathcal{O}'$, we first fix a partition of each element of $\mathcal{O}$ into even positive integers such that the terms in these partitions exactly constitute $\mathcal{O}'$. Let the number of choices for these partitions be $\eta_{\mathcal{O},\mathcal{O}'}$. Once these partitions are fixed, we have

$$\text{number of choices for } \mathcal{O}' \text{ is } \leq \Big(\frac{\prod_{2b \in \mathcal{O}} b!}{\prod_{2b \in \mathcal{O}'} b!}\Big)^2, \quad \text{so} \quad \xi_{\mathcal{O},\mathcal{O}'} \leq \Big(\frac{\prod_{2b \in \mathcal{O}} b!}{\prod_{2b \in \mathcal{O}'} b!}\Big)^2 \cdot \eta_{\mathcal{O},\mathcal{O}'}. \tag{9.6}$$

In fact, consider any subset $A \in \mathcal{O}$, say $|A| = 2a$, with $a = b_1 + \cdots + b_q$ as described above. To divide A into subsets of cardinalities $2b_j$ ($1 \leq j \leq q$) to form part of $\mathcal{O}'$ (we may call this part $\mathcal{O}_A$), we need to divide the set of elements with a $+$ sign and the set of elements with a $-$ sign separately, leading to at most $(a!)^2/((b_1)! \cdots (b_q)!)^2$ choices, considering also that there may be repetitions due to symmetry. Applying this for each A, we get the upper bound (9.6).

We write $\mathcal{O} = (2a_1, \ldots, 2a_r)$, and define $G(a) = (C_0 a)!/(a!)^2$. Using (9.6), we can bound the left hand side of (9.5) by

$$\sum_{\mathcal{O}' \prec \mathcal{O}} \eta_{\mathcal{O},\mathcal{O}'} \frac{r!}{|\mathcal{O}'|!} \frac{\prod_{2b \in \mathcal{O}'} G(b)}{G(a_1) \cdots G(a_r)}.$$

Using the definition of $\eta_{\mathcal{O},\mathcal{O}'}$, we can further reduce this to

$$\sum_{(\mathcal{P}_1,\ldots,\mathcal{P}_r)} \frac{r!}{(|\mathcal{P}_1|+\cdots+|\mathcal{P}_r|)!} \prod_{i=1}^{r} \frac{1}{G(a_i)} \prod_{2b\in\mathcal{P}_i} G(b)$$

where each $\mathcal{P}_i$ is a partition of $2a_i$ into even positive integers, and at least one $\mathcal{P}_i$ is non-trivial (i.e. contains at least two elements). Suppose q of these r partitions are nontrivial. Then $|\mathcal{P}_1|+\cdots+|\mathcal{P}_r| \geq r+q$. Thus we get an upper bound

$$\sum_{q=1}^{r} \binom{r}{q} \frac{r!}{(r+q)!} \left(\sup_a \sum_{\mathcal{P}} \frac{1}{G(a)} \prod_{2b\in\mathcal{P}} G(b)\right)^q, \tag{9.7}$$

where $\mathcal{P}$ runs over all nontrivial even partitions of $2a$. As $\binom{r}{q}\frac{r!}{(r+q)!} \leq 1/q!$, it suffices to show that

$$\sum_{\mathcal{P}} \frac{1}{G(a)} \prod_{2b\in\mathcal{P}} G(b) \leq \frac{1}{4}, \tag{9.8}$$

which would then imply that (9.7) $\leq 1/2$ and thus complete the induction.

The proof of (9.8) is easy. Note that $\log G$ is convex, so if $|\mathcal{P}| = s$, then

$$\frac{1}{G(a)} \prod_{2b\in\mathcal{P}} G(b) \leq \frac{((C_0)!)^{s-1} G(a-s+1)}{G(a)}$$

$$\leq \min\left(\frac{(C_0!)^{s-1}}{(C_0(s-1))!}, \frac{(C_0!)^{s-1}}{(C_0(a-s+1))^{C_0(s-1)}}\right) a^{2(s-1)}.$$

Using also $(m/3)^m \leq m! \leq m^m$, we can further bound this by

$$\frac{1}{G(a)} \prod_{2b\in\mathcal{P}} G(b) \leq \frac{C_0^{C_0(s-1)}}{(C_0\max((s-1)/3, a-s+1))^{C_0(s-1)}} \leq (a/4)^{-C_0(s-1)} a^{2(s-1)}.$$

The number of choices of $\mathcal{P}$ is at most a^{s-1}, so the left hand side of (9.8) is bounded by

$$\sum_{s=2}^{a} a^{s-1}(a/4)^{-C_0(s-1)} a^{2(s-1)} \leq \frac{1}{4}$$

provided C_0 is large enough (we may assume $a \geq 5$, since the cases $a \leq 4$ are easily verified). This proves (9.8), and finishes the inductive proof of (9.4).

Finally, suppose the sum of the elements in $\mathcal{O}$ that are at least 4 is q. Then $|\mathcal{O}| \geq (n-q)/2$. Thus by (9.4) we have

$$|\lambda(\mathcal{O})| \leq \frac{(n/2)!}{((n-q)/2)!}(C_0!)^{(n-q)/2}(C_0q/2)! \leq (C_0!)^{n/2}(n/2)^{q/2}(C_0q/2)^{C_0q/2}$$
$$\leq C_1^n n^{C_1 q},$$

which completes the proof. ∎

Remark 9.2. Note that $\xi_{\mathcal{O},(2,\ldots,2)} = \prod_{2b\in\mathcal{O}} b!$. In the Gaussian case, where $\mu_b = b!$, (9.3) implies that $\lambda(\mathcal{O}) = 0$ for *any* $\mathcal{O} \neq (2,\ldots,2)$, hence we reduce to Isserlis' theorem.

9.2. Over-gardens

Using Lemma 9.2 instead of Isserlis' theorem, we can replace (4.3) by

$$\mathbb{E}\left(\prod_{j=1}^{2R}(\mathcal{J}_{\mathcal{T}_j})_{k_j}^{\zeta_j}(t)\right) = \sum_{\mathcal{O}}\lambda(\mathcal{O})\mathcal{M}_{\mathcal{O}\mathcal{G}}(t, k_1, \ldots, k_{2R}). \tag{9.9}$$

Here $\mathcal{O}$ is a set of *over-pairings* of the leaves of the trees $\mathcal{T}_j$ $(1 \leq j \leq 2R)$, which is a partition of these leaves into subsets such that the number of leaves with sign $+$ in each subset is equal to the number of leaves with sign $-$. The total number of these leaves is $2(m + R)$ where m is the sum of scales of $\mathcal{T}_j$ $(1 \leq j \leq 2R)$, and $\mathcal{O}$ induces an even partition of $2(m + R)$, which we denote by $\mathcal{O}$, such that $\mathcal{O} \models \mathcal{O}$ in the sense of Lemma 9.1. The coefficient $\lambda(\mathcal{O})$ is as in Lemma 9.1, and $\mathcal{O}\mathcal{G}$ is the set of these $2R$ trees together with the set of over-pairings, which we refer to as an *over-garden*. Note that we may also write $\mathcal{O}\mathcal{G} \models \mathcal{O}$ instead of $\mathcal{O} \models \mathcal{O}$. Finally, as in (4.1) we have

$$\mathcal{M}_{\mathcal{O}\mathcal{G}}(t, k_1, \ldots, k_{2R})$$
$$= \left(\frac{\delta}{2L^{d-1}}\right)^m \zeta^*(\mathcal{O}\mathcal{G}) \sum_{\mathscr{I}} \epsilon_{\mathscr{I}} \int_{\mathcal{J}} \prod_{\mathfrak{n}\in\mathcal{N}^*} e^{\pi i \zeta_{\mathfrak{n}}\cdot\delta L^2 \Omega_{\mathfrak{n}} t_{\mathfrak{n}}}\, dt_{\mathfrak{n}} \cdot \prod_{\mathfrak{l}\in\mathcal{L}^*}^{(+)} n_{\text{in}}(k_{\mathfrak{l}}), \tag{9.10}$$

where all the objects are defined as in (4.1), except that for the decoration $\mathscr{I}$ we require that $k_{\mathfrak{n}}$ $(\mathfrak{n} \in A)$ are all equal for each $A \in \mathcal{O}$.

Clearly an over-garden $\mathcal{O}\mathcal{G}$ can be turned into a garden $\mathcal{G}$ by dividing each over-pairing $A \in \mathcal{O}$ into leaf pairs; below we will write $\mathcal{O}\mathcal{G} \sim \mathcal{G}$ for this. In this case $\mathcal{M}_{\mathcal{O}\mathcal{G}}$ is the same as $\mathcal{M}_{\mathcal{G}}$, except for the finitely many additional conditions of the form $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$ associated with the over-pairings structure of $\mathcal{O}\mathcal{G}$, which are added to the decoration $\mathscr{I}$ in the summation (9.10). Now by (9.9) we have

$$\mathbb{E}\left(\prod_{j=1}^{2R}(\mathcal{J}_{m_j})_{k_j}^{\zeta_j}(t)\right) = \sum_{\mathcal{O}}\lambda(\mathcal{O}) \sum_{\mathcal{O}\mathcal{G}\models\mathcal{O}}\mathcal{M}_{\mathcal{O}\mathcal{G}}(t, k_1, \ldots, k_{2R}), \tag{9.11}$$

where $\mathcal{O}$ is an even partition of $2(m + R)$ with $m = m_1 + \cdots + m_{2R}$, and $\mathcal{O}\mathcal{G} = (\mathcal{T}_1, \ldots, \mathcal{T}_{2R}) \models \mathcal{O}$ is the over-garden of width $2R$ and signature $(\zeta_1, \ldots, \zeta_{2R})$ such that the scale of $\mathcal{T}_j$ is m_j for $1 \leq j \leq 2R$. Note that for fixed $\mathcal{G}$, the number of $\mathcal{O}\mathcal{G}$ such that $\mathcal{O}\mathcal{G} \models \mathcal{O}$ and $\mathcal{O}\mathcal{G} \sim \mathcal{G}$ depends only on $\mathcal{O}$; similarly, for fixed $\mathcal{O}\mathcal{G} \models \mathcal{O}$, the number of $\mathcal{G}$ such that $\mathcal{O}\mathcal{G} \sim \mathcal{G}$ also depends only on $\mathcal{O}$. We denote them by $\sigma_1(\mathcal{O})$ and $\sigma_2(\mathcal{O})$ respectively. Thus (9.11) can be rewritten as

$$\mathbb{E}\left(\prod_{j=1}^{2R}(\mathcal{J}_{m_j})_{k_j}^{\zeta_j}(t)\right) = \sum_{\mathcal{O}}\frac{\lambda(\mathcal{O})}{\sigma_2(\mathcal{O})} \sum_{\mathcal{G}} \sum_{\mathcal{O}\mathcal{G}\sim\mathcal{G},\,\mathcal{O}\mathcal{G}\models\mathcal{O}}\mathcal{M}_{\mathcal{O}\mathcal{G}}(t, k_1, \ldots, k_{2R}), \tag{9.12}$$

where $\mathcal{G}$ satisfies the same condition as $\mathcal{O}\mathcal{G}$ but it runs over gardens instead of over-gardens.

Now the study of (9.1) reduces to the study of the quantities $\mathcal{M}_{\mathcal{O}\mathscr{G}}$ for over-gardens $\mathcal{O}\mathscr{G}$. To this end we introduce the notion of regular over-gardens, and one simple linear algebra lemma.

Definition 9.3. Define an over-garden $\mathcal{O}\mathscr{G}$ to be a *regular* over-garden if there exists $\mathscr{G}$ such that $\mathcal{O}\mathscr{G} \sim \mathscr{G}$, and (i) $\mathscr{G}$ is a regular multi-couple (Definition 4.1), and (ii) for each leaf $\mathfrak{l}$ in each over-pairing $A \in \mathcal{O}$ with $|A| \geq 4$, the tree $\mathcal{T}_j$ containing $\mathfrak{l}$ *must* be a regular tree and $\mathfrak{l}$ must be its lone leaf.

Lemma 9.4. *Let $A \subset B$ be two sets of affine linear equations posed on $\mathbb{R}^n$, in terms of the coordinates $(x_1, \ldots, x_n)$. Let $\mathcal{A} \supset \mathcal{B}$ be the affine submanifold of $\mathbb{R}^n$ determined by equations in A and B respectively, assume $\mathcal{B} \neq \varnothing$, and denote $p = \operatorname{codim}_{\mathcal{A}}(\mathcal{B})$. Let $1 \leq r \leq n - 1$.*

For any fixed $x = (x_1, \ldots, x_r) \in \mathbb{R}^r$, consider the affine submanifold $\mathcal{A}_x \supset \mathcal{B}_x$ of $(x_{r+1}, \ldots, x_n) \in \mathbb{R}^{n-r}$ determined by equations in A and B for $(x, x_{r+1}, \ldots, x_n)$ respectively. Let $\mathcal{A}^\circ \supset \mathcal{B}^\circ$ be the affine manifold for x such that $\mathcal{A}_x \neq \varnothing$ and $\mathcal{B}_x \neq \varnothing$ respectively. If $\operatorname{codim}_{\mathcal{A}^\circ}(\mathcal{B}^\circ) = p^\circ$, then $\operatorname{codim}_{\mathcal{A}_x}(\mathcal{B}_x) = p - p^\circ$ for any $x \in \mathcal{B}^\circ$.

Proof. We omit the proof as it is elementary. ∎

Now we can state the main estimate for the non-Gaussian case.

Proposition 9.5. *Fix R and $(\zeta_1, \ldots, \zeta_{2R})$ and $(k_1, \ldots, k_{2R})$ and $(m_1, \ldots, m_{2R})$ as in Proposition 4.7. Assume $R \leq 2 \log L / (\log \log L)^2$ and $m_j \leq N$ $(1 \leq j \leq 2R)$, and set $m := m_1 + \cdots + m_{2R}$. Consider the sum $\mathscr{I}'$ as in (9.12), but we restrict to* nonregular *over-gardens $\mathcal{O}\mathscr{G}$ in the summation. Then*

$$|\mathscr{I}'| \leq (C^+ \delta^{1/4})^m L^{-\nu} \tag{9.13}$$

uniformly in t and $(k_1, \ldots, k_{2R})$.

Proof. Let $\mathcal{O}\mathscr{G}$ be an over-garden and $\mathcal{O}\mathscr{G} \sim \mathscr{G}$. Since the expression $\mathcal{M}_{\mathcal{O}\mathscr{G}}$ in (9.10) is just the expression $\mathcal{M}_{\mathscr{G}}$ in (4.1) with finitely many extra requirements of the form $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$ in the summation, it is clear that $\mathcal{M}_{\mathcal{O}\mathscr{G}}$ can at least be estimated in the same way as $\mathcal{M}_{\mathscr{G}}$ with no power loss. Also the number of $\mathcal{O}$ is at most C^m, and for fixed $\mathscr{G}$ and $\mathcal{O}$, the third sum in (9.12) contains $\sigma_1(\mathcal{O})$ terms, where $\sigma_1(\mathcal{O})$ has the same upper bound as $\lambda(\mathcal{O})$ in Lemma 9.2.

Therefore, compared to the bounds for $\mathcal{M}_{\mathscr{G}}$ that we already know, we only have two tasks in proving the desired result: first, gain the extra power $L^{-\nu}$, and second, gain enough extra powers to cancel the factor $|\lambda(\mathcal{O})|$ in case it is too large. Note also that if $\mathcal{O}$ is given and $\mathscr{G}$ and $\mathscr{G}'$ are congruent in the sense of Definition 5.4, then the over-gardens $\mathcal{O}\mathscr{G} \models \mathcal{O}$, $\mathcal{O}\mathscr{G} \sim \mathscr{G}$ are in one-to-one correspondence with the over-gardens $\mathcal{O}\mathscr{G}' \models \mathcal{O}$, $\mathcal{O}\mathscr{G}' \sim \mathscr{G}'$, and cancellations between the terms $\mathcal{M}_{\mathcal{O}\mathscr{G}}$ and $\mathcal{M}_{\mathcal{O}\mathscr{G}'}$ are the same as the cancellations between $\mathcal{M}_{\mathscr{G}}$ and $\mathcal{M}_{\mathscr{G}'}$ in Section 5, up to minor modifications. Hence, we can exploit the same cancellation for irregular chains in $\mathscr{G}$ as in Section 5 and [12].

Now let us go over the process of studying $\mathcal{M}_\mathscr{G}$, and see what the extra conditions $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$ may do at each step of this process. Below let q_0 be the number of independent extra equations $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$. Then $q_0 \sim q$, where q is the sum of the elements in $\mathcal{O}$ that are at least 4, as in Lemma 9.1. In fact, we have

$$q = \sum_{4 \leq a \in \mathcal{O}} a, \quad q_0 = \sum_{4 \leq a \in \mathcal{O}} (a/2 - 1),$$

hence $2q_0 \leq q \leq 4q_0$. We will keep track of the codimension introduced by these q_0 extra equations to the affine manifold of all possible decorations $(k_{\mathfrak{l}})$ using Lemma 9.4.

(1) Assume that $\mathscr{G}$ is a regular multi-couple. In this case, we shall estimate the summation (together with the integration) in (9.10) using the method of [12, Section 6] (note that here we have to treat all the regular couples in $\mathscr{G}$ together—instead of one at a time in [12, Section 6]—because of the extra conditions linking different regular couples together, but this will cause minor changes to the proof). In particular, we define the variables $x_\mathfrak{n}$ and $y_\mathfrak{n}$ as in [12, proof of Proposition 6.7]. Note that there are $2R$ linear equations that any decoration of the leaves of $\mathscr{G}$ must satisfy (and such a decoration of leaves uniquely determines the full decoration of $\mathscr{G}$); moreover, the set of decorations satisfying these $2R$ equations is in affine bijection with the set of free variables $(x_\mathfrak{n}, y_\mathfrak{n})$ (see [12, proof of Proposition 6.7]).

Now, with the extra conditions, the dimension of the affine manifold of all possible decorations $(k_{\mathfrak{l}})$ gets strictly lower, and the codimension r introduced satisfies $r \gtrsim \max(1, q_0 - O(R))$. In fact, we have $r > 0$ because at least one extra condition must take the form $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$ where $\mathfrak{l}$ is not the lone leaf of a regular tree by Definition 9.3, and this equation will be independent of the $2R$ original equations stated above (since the only way for this extra condition to be dependent is for the two trees containing $\mathfrak{l}$ and $\mathfrak{l}'$ to be distinct and coupled, which would easily imply that they are two regular trees with lone leaves $\mathfrak{l}$ and $\mathfrak{l}'$). The lower bound $q_0 - O(R)$ is because the number of independent extra equations is q_0, and we subtract $O(R)$ because some extra equations combined may imply some of the $2R$ original equations.

Using the affine bijection, we know that the $(x_\mathfrak{n}, y_\mathfrak{n})$ variables must satisfy r independent linear equations. Then, we proceed as in [12, proof of Proposition 6.1], and sum over the $(x_\mathfrak{n}, y_\mathfrak{n})$ variables one by one. At each step, suppose we are summing over the pair (x_j, y_j); depending on the extra equations satisfied by these variables, we have one of three possibilities: (a) there is no restriction on (x_j, y_j) and we are summing over all choices of (x_j, y_j); (b) we are summing over (x_j, y_j) that satisfy one linear equation (such as $x_j = \text{const}$ or $y_j = \text{const}$ or $x_j \pm y_j = \text{const}$); (c) we are summing over (x_j, y_j) that satisfy two linear equations, i.e. over only one point (x_j^*, y_j^*). In either case the summation can be performed in the same way as in [12], and in cases (b) and (c) we are gaining a power of L in this summation, compared to the factor L^{2d-2} in [12]. Also, by repeating Lemma 9.4, we know that case (b) or (c) must occur at least $\gtrsim r$ times during this process.

Therefore, altogether, with these extra conditions we can gain a power L^{-cr} for some small constant c with $r \gtrsim \max(1, q - O(R))$, in the estimate of $\mathcal{M}_{\mathcal{O}\mathscr{G}}$ compared to $\mathcal{M}_\mathscr{G}$.

This already provides the needed $L^{-\nu}$ gain. It also covers any possible loss due to $\lambda(\mathcal{O})$ because $|\lambda(\mathcal{O})| \leq C^m m^{Cq} \leq C^m (\log L)^{Cq_0}$ by Lemma 9.1, hence

$$|\lambda(\mathcal{O})| \leq C^m (\log L)^{O(R)} \cdot (\log L)^{C \max(1, q_0 - O(R))}$$
$$\leq C^m L^{\nu/2} \cdot (\log L)^{O(r)}$$

by our choice $R \leq \log L / (\log \log L)^2$, which can then be covered by the $L^{-\nu}$ gain; also the various losses of C^m are unimportant as they can be absorbed into $(C^+)^m$ in (9.13).

(2) Now we assume $\mathcal{G}$ is *not* a regular multi-couple. Then the sum of $\mathcal{M}_{\mathcal{G}}$ already gains the power $L^{-\nu}$ in view of Proposition 4.7 and also [12, Proposition 10.4]. It then suffices to cover the possible loss due to $\lambda(\mathcal{O})$. We perform the reduction steps as in previous sections, and analyze the extra conditions appearing in each step. As in (1), the total codimension introduced by the q_0 extra equations is $r \gtrsim \max(1, q_0 - O(R))$; we may assume $q_0 \gg R$ because otherwise the loss

$$|\lambda(\mathcal{O})| \leq C^m m^{Cq_0} \leq C^m L^{\nu/2}$$

can already be covered by the guaranteed $L^{-\nu}$ gain. Therefore we now have $r \gtrsim q_0$.

Step 1. We first remove the regular couples and regular trees to reduce $\mathcal{G}$ to its skeleton $\mathcal{G}_{\text{sk}}$ as in Proposition 4.4. In the process we are fixing all the remaining $k_{\mathfrak{n}}$ variables (which are determined by the variables $k_{\mathfrak{l}_1}$ for leaves $\mathfrak{l}_1$ of $\mathcal{G}_{\text{sk}}$) and sum over the variables $k_{\mathfrak{l}_2}$, where $\mathfrak{l}_2$ are leaves of these regular couples and regular trees, similar to (1) above. By Lemma 9.4, there exists $r_1 + r_2 = r$ such that for any fixed $(k_{\mathfrak{l}_1})$, the codimension of the submanifold formed by the $(k_{\mathfrak{l}_2})$ variables is r_2, and the codimension of the submanifold formed by the $(k_{\mathfrak{l}_1})$ variables is r_1. By repeating the argument in (1) above, we can gain a power L^{-cr_2} in summing over the $(k_{\mathfrak{l}_2})$ variables. Note that some extra equations satisfied by the $(k_{\mathfrak{l}_2})$ variables may be of the form $k_{\mathfrak{l}_2} = \text{const}$ instead of $k_{\mathfrak{l}_2} = k_{\mathfrak{l}_2'}$ as in (1), but this does not affect the proof.

Step 2. We further remove the irregular chains from the skeleton $\mathcal{G}_{\text{sk}}$ and exploit the cancellation as in Section 5. Note that if $\mathcal{G}$ is not a regular multi-couple, then *any $\mathcal{O}\mathcal{G} \sim \mathcal{G}$* must be nonregular in the sense of Definition 9.3, so for fixed $\mathcal{G}$, the summation in $\mathcal{O}\mathcal{G}$ we are studying here is still the same as the one in (9.12) even though we have made the restriction that $\mathcal{O}\mathcal{G}$ is nonregular. Thus, as said above, the cancellation for irregular chains works the same way in the current situation as in Section 5. The extra conditions again lead to a gain of powers in L. As in Step 1, we can write $r_1 = r_3 + r_4$ so that we can gain a power L^{-cr_3} in the current step, and after removing the irregular chains, the remaining decoration (of the remaining garden $\mathcal{G}_{\text{sk}}^{\#}$, see Section 5.4) still satisfies r_4 extra linear equations.

Step 3. Finally, we reduce the estimate of the remaining expression to the counting problem associated with the molecule formed from $\mathcal{G}_{\text{sk}}^{\#}$. Here we only need to show that if, in addition to the equations in the counting problem, the variables in question also satisfy r_4 additional independent linear equations, then we can improve the upper bound for the counting problem by a power L^{-cr_4} with a small constant c.

To see this, we follow the procedure described in Section 6, and in particular apply the algorithm introduced in [12, Section 9.4]. In the process, when we fix some of the variables in each step, we keep track of the codimension, or the number p of independent equations satisfied by the remaining variables. Initially we have $p \gtrsim r_4$, while in the end $p = 0$. Therefore, there must be at least $\gtrsim r_4$ steps where p strictly decreases. If this step is a good step in the sense of [12, Section 9.3], then we are gaining a constant power $L^{-\nu}$ here; even if it is a normal step, since $\Delta p < 0$, by Lemma 9.4, in doing the counting estimate for this step, we can take into account an additional independent linear equation satisfied by the variables under consideration. For example, if we perform the step (3R-1) defined in [12, Section 9.3.8], then the corresponding counting problem we solve is (say)

$$\begin{cases} a - b + c = \text{const}, \\ |a|_\beta^2 - |b|_\beta^2 + |c|_\beta^2 = \text{const} + O(L^{-2}), \\ a, b, c \in \mathbb{Z}_L^d, \ |a|, |b|, |c| \lesssim 1, \end{cases}$$

which has $O(L^{2d-2})$ solutions. However, if we add to this system another independent linear equation $\alpha a + \beta b + \gamma c = \text{const}$, where (α, β, γ) is not a multiple of $(1, -1, 1)$, then the number of solutions will be at most L^d with $d < 2d - 2$. This leads to a power gain in each such step, so in total we can gain a power L^{-cr_4} for some constant c.

After the above three steps, the total power we gain would be $L^{-c(r_2+r_3+r_4)} = L^{-cr}$, which is enough to cover the loss $C^m m^{Cq}$ from $\lambda(\mathcal{O})$ because $r \gtrsim q_0 \gtrsim q$. Therefore in any case we can cover the possible loss with an extra gain of $L^{-\nu}$, hence (9.13) holds. This completes the proof. ∎

With Proposition 9.5 we can now prove Theorem 1.4.

Proof of Theorem 1.4. We use (3.6) to expand $\mathbb{E}(\prod_{j=1}^{2R} a_{k_j}^{\zeta_j}(t))$. The estimate for the remainder term b can be done using arguments similar to Section 8.2, which we shall omit. Then, using also (9.12), we can write

$$\mathbb{E}\Big(\prod_{j=1}^{2R} a_{k_j}^{\zeta_j}(t)\Big) = \sum_{\mathscr{G},\mathcal{O}} \frac{\lambda(\mathcal{O})}{\sigma_2(\mathcal{O})} \sum_{\mathcal{O}\mathscr{G}\sim\mathscr{G},\,\mathcal{O}\mathscr{G}\models\mathcal{O}} \mathcal{M}_{\mathcal{O}\mathscr{G}}(t, k_1, \ldots, k_{2R}) + O(L^{-\nu}), \quad (9.14)$$

where $\mathscr{G}$ runs over all gardens of width $2R$ such that the scale of each tree is at most N, and $\mathcal{O}$ runs over all even partitions of $2(m + R)$ where m is the scale of $\mathscr{G}$. By Proposition 9.5 and summing over all possible m_j as in the proof of Theorem 1.3 above, we see that with R fixed and $L \to \infty$, the contribution of *nonregular* over-gardens $\mathcal{O}\mathscr{G}$ decays like $L^{-\nu}$ in the limit. Thus, we only need to consider *regular* over-gardens $\mathcal{O}\mathscr{G}$. Suppose $\mathcal{O}\mathscr{G} \sim \mathscr{G}$. Then $\mathscr{G}$ is a regular multi-couple. Therefore, unless we can divide $\{1, \ldots, 2R\}$ into pairs such that for each pair $\{i, j\}$ we have $k_i = k_j$ and $\zeta_i = -\zeta_j$, the contribution of regular over-gardens must vanish, in particular (1.2) is true. Now we only need to prove (1.6).

For any regular couple $\mathcal{Q} = (\mathcal{T}_1, \mathcal{T}_2)$, the tree $\mathcal{T}_1$ is a regular tree if and only if $\mathcal{T}_2$ is a regular tree (and hence the two lone leafs are paired). In this case we say $\mathcal{Q}$ is *tangential* (since the two trees only have one leaf-pair in common), otherwise $\mathcal{Q}$ is *nontangential*. Note that by [12, proof of Theorem 1.3] we have

$$\sum_{\mathcal{Q}} \mathcal{M}_{\mathcal{Q}}(t, t, k) = n(\delta t, k) + O(L^{-\nu}), \tag{9.15}$$

where $\mathcal{Q}$ runs over all regular couples with both trees having scale at most N. If we restrict to *tangential* couples only, then the sum should be approximated by $n_0(\delta t, k)$ where n_0 is defined in (WKE-0). The easiest way to see this is by observing that the expression $\mathcal{M}_{\mathcal{Q}}(t, t, k)$ contains a factor $n_{\mathrm{in}}(k)$ if and only if $\mathcal{Q}$ is tangential, because for any regular tree $\mathcal{T}$ with root r and lone leaf l we must have $k_l = k_r = k$. Thus, since the sum (9.15) over all couples exactly matches the Taylor expansion of $n(\delta t, k)$ (as shown in [12]), we know that the sum over tangential couples will exactly match the terms in the Taylor expansion *that contain the factor* $n_{\mathrm{in}}(k)$. Due to the form of (KIN), it is easy to see that the sum of these terms is exactly $n_0(t, k)$, hence the result. Therefore we have

$$\sum_{\mathcal{Q} \text{ tangential}} \mathcal{M}_{\mathcal{Q}}(t, t, k) = n_0(\delta t, k) + O(L^{-\nu}), \tag{9.16}$$

$$\sum_{\mathcal{Q} \text{ nontangential}} \mathcal{M}_{\mathcal{Q}}(t, t, k) = n_+(\delta t, k) + O(L^{-\nu}). \tag{9.17}$$

We now return to the sum (9.14) over regular over-gardens $\mathcal{O}\mathcal{G}$. For (1.6), we may rename $(k_1, \ldots, k_{2R})$ so that there are $2a_j$ copies of k_i^* for $1 \le i \le r$ (with half of them having sign $+$ and half having sign $-$) where the k_i^* are all different and

$$a_1 + \cdots + a_r = R.$$

For simplicity we will write k_i instead of k_i^* below. Clearly the $2a_i$ trees corresponding to the input variable k_i must form a_i couples in $\mathcal{G}$; assume b_i of these a_i couples are tangential and the rest are nontangential, where $0 \le b_i \le a_i$. Note also that for any $\mathcal{O}\mathcal{G} \sim \mathcal{G}$ we have $\mathcal{M}_{\mathcal{O}\mathcal{G}} = \mathcal{M}_{\mathcal{G}}$ because over-pairings can only happen at lone leaves of regular trees. Therefore, for fixed $(b_1, \ldots, b_r)$, we can calculate the contribution of regular over-gardens as

$$\mathscr{I}_{b_1,\ldots,b_r} = \prod_{i=1}^{r} \binom{a_i}{b_i}^2 (a_i - b_i)! \left(\sum_{\mathcal{Q} \text{ nontangential}} \mathcal{M}_{\mathcal{Q}}(t, t, k_i) \right)^{a_i - b_i}$$

$$\times \sum_{\mathcal{G}', \mathcal{O}'} \frac{\lambda(\mathcal{O}')\sigma_3(\mathcal{O}')}{\sigma_2(\mathcal{O}')} \mathcal{M}_{\mathcal{G}'}(t, k_1, \ldots, k_1, \ldots, k_r, \ldots, k_r). \tag{9.18}$$

Here $\mathcal{G}'$ runs over all gardens *of width* $2(b_1 + \cdots + b_r)$ which are multi-couples formed by *tangential couples* (and otherwise same as $\mathcal{G}$), $\mathcal{O}'$ runs over all *even partitions of* $2(b_1 + \cdots + b_r)$ such that $\mathcal{O}' \preceq (2b_1, \ldots, 2b_r)$, and $\sigma_3(\mathcal{O}')$ is the number of regular

over-gardens $\mathcal{OG}$ such that $\mathcal{OG} \sim \mathcal{G}$ and $\mathcal{OG} \models \mathcal{O}$ where $\mathcal{O} = (\mathcal{O}', 2, \ldots, 2)$. We can further reduce the inner sum in (9.18) to

$$\sum_{\mathcal{G}',\mathcal{O}'} \frac{\lambda(\mathcal{O}')\sigma_3(\mathcal{O}')}{\sigma_2(\mathcal{O}')} \mathcal{M}_{\mathcal{G}'}(t, k_1, \ldots, k_1, \ldots, k_r, \ldots, k_r) = \prod_{i=1}^{r} \left(\sum_{\mathcal{Q} \text{ tangential}} \mathcal{M}_{\mathcal{Q}}(t, t, k_i) \right)^{b_i}$$

$$\times \sum_{\mathcal{O}'} \frac{\lambda(\mathcal{O}')\sigma_3(\mathcal{O}')}{\sigma_2(\mathcal{O}')} b_1! \cdots b_r!. \quad (9.19)$$

Finally, in the above summation $\mathcal{O}'$ must be $\preceq (2b_1, \ldots, 2b_r)$ otherwise $\sigma_3(\mathcal{O}') = 0$, and by definitions and a double counting argument we can show that under this assumption we have

$$\xi_{(2b_1,\ldots,2b_r),\mathcal{O}'} = \sigma_3(\mathcal{O}')\frac{b_1! \cdots b_r!}{\sigma_2(\mathcal{O}')}.$$

In fact, fix $\mathcal{O} \models (2b_1, \ldots, 2b_r)$. To construct $\mathcal{O}'$ such that $\mathcal{O}' \preceq \mathcal{O}$ and $\mathcal{O}' \models \mathcal{O}'$, we first divide each subset in $\mathcal{O}$ into pairs (which has $b_1! \cdots b_r!$ choices), thus obtaining a garden $\mathcal{G}$, then construct $\mathcal{OG} \sim \mathcal{G}$ and $\mathcal{OG} \models \mathcal{O}$ (which has $\sigma_3(\mathcal{O}')$ choices) and form $\mathcal{O}'$ accordingly. Note that each $\mathcal{O}'$ is counted exactly $\sigma_2(\mathcal{O}')$ times (which is just the number of choices of $\mathcal{G}$ with fixed $\mathcal{OG}$), hence the result.

Now we can reduce the last sum in (9.18) to

$$\sum_{\mathcal{O}'} \frac{\lambda(\mathcal{O}')\sigma_3(\mathcal{O}')}{\sigma_2(\mathcal{O}')} b_1! \cdots b_r! = \sum_{\mathcal{O}' \preceq (2b_1,\ldots,2b_r)} \lambda(\mathcal{O}')\xi_{(2b_1,\ldots,2b_r),\mathcal{O}'} = \prod_{i=1}^{r} \mu_{b_i}$$

using (9.3). Altogether, using (9.16) and (9.17), and then summing over b_i ($1 \leq i \leq r$), we get

$$\mathbb{E}\left(\prod_{j=1}^{2R} a_{k_j}^{\xi_j}(t)\right)$$

$$= \sum_{0 \leq b_i \leq a_i \ (1 \leq i \leq r)} \prod_{i=1}^{r} \binom{a_i}{b_i}^2 (a_i - b_i)! \mu_{b_i} (n_0(\delta t, k))^{b_i} (n_+(\delta t, k))^{a_i - b_i} + O(L^{-\nu}),$$

which is just (1.6) given (1.7). This completes the proof. ∎

9.3. Evolution of density

Finally, we prove Theorem 1.5. Note that if $\mu_r \leq C^r (2r)!$, then by (1.7) for any t we also have $\mu_r(t, k) \leq C^r (2r)!$, perhaps for some different C. Thus, convergence in law will be a consequence of the following lemma:

Lemma 9.6. *Suppose $\{X_n\}$ are $\mathbb{R}^d$-valued random variables such that for any multi-index μ the limit*

$$A_\mu := \lim_{n \to \infty} \mathbb{E}(X_n^\mu)$$

exists and $|A_\mu| \leq C^{|\mu|}|\mu|!$. Then $\{X_n\}$ converges in law to a random variable X satisfying $\mathbb{E}(X^\mu) = A_\mu$ for any multi-index μ. Moreover, the law with these given moments is unique.

Proof. First the assumption implies that $\mathbb{E}\,|X_n|^2$ is uniformly bounded in n, thus the sequence of laws of X_n is tight. For any subsequence X_{n_k} we then have a subsequence $X_{n_{k_\ell}}$ that converges in law to (say) some random variable X. For any μ, since $\mathbb{E}\,|X_n|^{2|\mu|}$ is bounded in n it is easy to see that $\mathbb{E}(X^\mu) = \lim \mathbb{E}(X_n^\mu) = A_\mu$. Therefore, it suffices to prove that the law of the random variables X with $\mathbb{E}(X^\mu) = A_\mu$ is unique. This is indeed true: Since $|A_\mu| \leq C^{|\mu|}|\mu|!$, we have $\mathbb{E}(e^{\varepsilon|X|}) < \infty$ for small enough ε, hence $f(\xi) = \mathbb{E}(e^{i\xi \cdot X})$ is well-defined and holomorphic in the region $|\operatorname{Im}\xi| < \varepsilon$. The moments $\mathbb{E}(X^\mu)$ uniquely determine the Taylor expansion of $f(\xi)$ at $\xi = 0$, hence uniquely determine the value of $f(\xi)$ in a neighborhood of 0—and consequently in the whole region $|\operatorname{Im}\xi| < \varepsilon$ by analyticity. In particular, the moments uniquely determine the value of $f(\xi)$ on $\mathbb{R}$, which is the characteristic function of X. Thus the law of X is unique, as desired. ∎

Now, for any $t \in [0, \delta]$ and $k \in \mathbb{Z}_L^d$, consider the unique radial density $\rho = \rho_k(t, v)$ (where $v \in \mathbb{C}$ is also viewed as an $\mathbb{R}^2$ vector) such that

$$\int_{\mathbb{C}} (v_1^2 + v_2^2)^r \rho_k(t, v)\, dv_1\, dv_2 = \mu_r(t, k)$$

$$= \sum_{p=0}^{r} \binom{r}{p}^2 (r - p)! \mu_p \cdot (n_0(t, k))^p (n_+(t, k))^{r-p}. \quad (9.20)$$

Then by Theorem 1.4 and Lemma 9.6, convergence in law described by (1.11) and (1.12) is true. Thus to prove Theorem 1.5 it suffices to show that $\rho_k(t, v)$ solves (1.10). The initial data $\rho_k(0)$ is satisfied by definition since the right hand side of (9.20) equals $\mu_r n_{\text{in}}(k)^r$ when $t = 0$. Now, by approximating from below, we may assume $\{\mu_p\}$ is bounded, and consider

$$L_k(t, \xi) = \int_{\mathbb{C}} e^{i\xi(v_1^2 + v_2^2)} \rho_k(t, v)\, dv_1\, dv_2$$

$$= \sum_{p=0}^{r} \frac{(i\xi)^r}{p!} \binom{r}{p} \mu_p \cdot (n_0(t, k))^p (n_+(t, k))^{r-p}.$$

Taking the time derivative and calculating using (WKE), (KIN), (WKE-0) and (KIN-0) yields

$$\partial_t L_k = \sum_{p=0}^{r} \frac{(i\xi)^r}{p!} \binom{r}{p} \mu_p \cdot \big\{ p(n_0(t, k))^p (n_+(t, k))^{r-p} \gamma_k(t)$$

$$+ (r - p)(n_0(t, k))^p (n_+(t, k))^{r-p-1} (\sigma_k(t) + \gamma_k(t) n_+(t, k)) \big\}, \quad (9.21)$$

where $\sigma_k(t)$ and $\gamma_k(t)$ are defined in (1.8) and (1.9). This simplifies to

$$\partial_t L_k = \gamma_k(t) \cdot \sum_{p=0}^{r} r \cdot \frac{(i\xi)^r}{p!} \binom{r}{p} \mu_p \cdot (n_0(t,k))^p (n_+(t,k))^{r-p}$$

$$+ \sigma_k(t) \cdot \sum_{p=0}^{r} (r-p) \cdot \frac{(i\xi)^r}{p!} \binom{r}{p} \mu_p \cdot (n_0(t,k))^p (n_+(t,k))^{r-p-1}. \quad (9.22)$$

The first sum on the right hand side equals $\xi \partial_\xi L_k$, and the second sum equals

$$\sum_{p=0}^{r} (r+1) \cdot \frac{(i\xi)^{r+1}}{p!} \binom{r}{p} \mu_p \cdot (n_0(t,k))^p (n_+(t,k))^{r-p}$$

by changing r to $r+1$, which then equals $(i\xi + i\xi^2 \partial_\xi)L_k$. Therefore

$$\partial_t L_k = \gamma_k(t) \cdot \xi \partial_\xi L_k + \sigma_k(t) \cdot (i\xi + i\xi^2 \partial_\xi)L_k.$$

Finally, by taking the inverse Fourier transform and switching between Cartesian and polar coordinates, we can verify that the density $\rho_k(t,v)$ solves (1.10). This completes the proof of Theorem 1.5.

10. Proof of Theorem 1.6

In this last section we prove Theorem 1.6. The proof is of "soft" nature and has a different flavor from the other parts of this paper. The main idea, already hinted at in Section 1.4, is to represent n_r as an average of tensor products for which one can apply Theorem 1.3. This is demonstrated in the following lemma, which is a variant of the classical Hewitt–Savage theorem.

Lemma 10.1 (a variant of Hewitt–Savage). *Suppose $(n_r)_{\text{in}}$ is as in the statement of Theorem 1.6, satisfies (1.14), and is admissible in the sense of (1.15). Recall the set $\mathcal{A}$ defined in (1.16). Then there exists a unique probability measure ζ on $\mathcal{A}$ such that for any r and any distinct $(k_1, \ldots, k_r)$ we have*

$$(n_r)_{\text{in}}(k_1, \ldots, k_r) = \int_{\mathcal{A}} m(k_1) \cdots m(k_r) \, d\zeta(m). \quad (10.1)$$

Proof. For simplicity we will write $(n_r)_{\text{in}}$ just as n_r. We may assume $\mathfrak{X} = 1$ (otherwise simply replace n_r by $\mathfrak{X}^{-r} n_r$ and C_1 by $\mathfrak{X}^{-1} C_1$). Then each n_r is the density of the joint distribution of some r random variables valued in $\mathbb{R}^d$, and n_{r-1} is a marginal of n_r. By Kolmogorov's extension theorem, we can find an infinite sequence $(X_1, X_2, \ldots)$ of $\mathbb{R}^d$-valued random variables such that n_r is the density of the joint distribution of $(X_1, \ldots, X_r)$. By symmetry of n_r, these random variables X_j are exchangeable, so by the Hewitt–Savage theorem [24], there exists a unique probability measure ζ on the set $\mathcal{M}$ of *all probability measures m on $\mathbb{R}^d$ such that (10.1) is true with $\mathcal{A}$ replaced by $\mathcal{M}$.

Now it remains to prove that ζ is supported on $\mathcal{A}$. We will apply the beautiful argument of Rosenzweig–Staffilani [31, Section 5], which goes back to the work of Spohn [32]. Fix a Schwartz function φ on $\mathbb{R}^d$ and indices α, β with $|\alpha|, |\beta| \leq 40d$; for any *even* integer r, we may test both sides of (10.1) by the tensor product $(\partial_k^\beta k^\alpha \varphi)^{\otimes r}$ to get

$$\int_{\mathcal{M}} (m, \partial_k^\beta k^\alpha \varphi)^r \, \mathrm{d}\zeta(m) = \int_{(\mathbb{R}^d)^r} n_r(k_1, \ldots, k_r) \prod_{j=1}^r k_j^\alpha \partial_{k_j}^\beta \varphi(k_j) \, \mathrm{d}k_1 \cdots \mathrm{d}k_r,$$

where $(m, \partial_k^\beta k^\alpha \varphi)$ is the integral of $\partial_k^\beta k^\alpha \varphi$ against the measure m, in the usual distributional sense. Recall that $\|n_r\|_{\mathfrak{G}^{40d:r}} \leq C_1^r$ from (1.14), so the right hand side is bounded by

$$\int_{(\mathbb{R}^d)^r} n_r(k_1, \ldots, k_r) \prod_{j=1}^r \partial_{k_j}^\beta k_j^\alpha \varphi(k_j) \, \mathrm{d}k_1 \cdots \mathrm{d}k_r \leq \|n_r\|_{\mathfrak{G}^{40d:r}} \cdot \|\varphi\|_{L^2}^r \leq (C_1 \|\varphi\|_{L^2})^r.$$

For any $C_2 > C_1$, by the Chebyshev inequality, this implies that

$$\zeta(\{m : |(m, \partial_k^\beta k^\alpha \varphi)| \geq C_2 \|\varphi\|_{L^2}\}) \leq (C_2 \|\varphi\|_{L^2})^{-r} \int_{\mathcal{M}} (m, \partial_k^\beta k^\alpha \varphi)^r \, \mathrm{d}\zeta(m) \leq (C_1/C_2)^r.$$

Since r is arbitrary, we conclude that

$$|(m, \partial_k^\beta k^\alpha \varphi)| \leq C_2 \|\varphi\|_{L^2} \tag{10.2}$$

ζ-almost surely for any fixed φ, fixed (α, β), and fixed $C_2 > C_1$. By choosing φ in a countable dense subset of the Schwartz space, enumerating finitely many possible (α, β), and choosing a sequence $C_2 = (1 + \varepsilon)C_1$, we find that ζ-almost surely in m, (10.2) actually holds for *all* (φ, α, β), and with C_2 replaced by C_1. By the Riesz representation theorem, this implies that

$$\|k^\alpha \partial_k^\beta m\|_{L^2} \leq C_1, \quad \forall |\alpha|, |\beta| \leq 40d$$

ζ-almost surely. By definition this means that $m \in \mathcal{A}$ because m is a probability measure and $\mathcal{X} = 1$. Therefore $\zeta(\mathcal{M} \backslash \mathcal{A}) = 0$, and hence we can define the measure ζ on $\mathcal{A}$ so that (10.1) is true. ∎

We are now ready to prove Theorem 1.6.

Proof of Theorem 1.6. First, if $(n_r)_{\text{in}}$ is hybrid then it must be admissible. In fact, by the definition of hybrid initial data we have

$$\int_{\mathcal{A}} m(k_1) \cdots m(k_r) \, \mathrm{d}\zeta(m) = (n_r)_{\text{in}}(k_1, \ldots, k_r) \tag{10.3}$$

for any distinct $k_j \in \mathbb{Z}_L^d$ ($1 \leq j \leq r$). Since m and $(n_r)_{\text{in}}$ are all continuous functions, by taking suitable limits and letting $L \to \infty$ we know that (10.3) is actually true for all $k_j \in \mathbb{R}^d$ ($1 \leq j \leq r$). Then we simply integrate (10.3) in k_r, using the integral condition in the definition of $\mathcal{A}$, to get (1.15).

From now on we shall assume $(n_r)_{\text{in}}$ is admissible. Then by Lemma 10.1, we can find a unique measure ζ such that (10.3) holds. Consider the hybrid data u_{in} described in the statement of Theorem 1.6. We can view it as obtained by first randomly selecting $m \in \mathcal{A}$ with law given by ζ, then working in the same setting as in Theorems 1.1 and 1.3–1.5 with the particular choice $n_{\text{in}} = m$. Since $m \in \mathcal{A}$, by Theorem 1.1 we know that the conditional probability $\mathbb{P}(E|m) \geq 1 - L^{-A}$ for any $m \in \mathcal{A}$, where E is the event that (NLS) has a smooth solution up to time T. This implies that

$$\mathbb{P}(E) = \int_{\mathcal{A}} \mathbb{P}(E|m) \, \mathrm{d}\zeta(m) \geq 1 - L^{-A}.$$

Finally, we prove (1.18). As is clear from the proof of Theorem 1.4, the remainder in (1.6) is in fact $O(L^{-\nu})$ for some absolute constant $\nu > 0$, where the implicit constant in $O(\cdot)$ may depend on r, but is uniform in (t, k_j) and $n_{\text{in}} = m$, as long as $m \in \mathcal{A}$. Thus, by Theorem 1.4, for the hybrid data u_{in}, for $t \in [0, T]$ we have

$$\mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t, k_j)|^2\right) = \int_{\mathcal{A}} \mathbb{E}\left(\prod_{j=1}^{r} |\hat{u}(t, k_j)|^2 \,\Big|\, m\right) \mathrm{d}\zeta(m)$$

$$= \int_{\mathcal{A}} \prod_{j=1}^{r} \tilde{m}\left(\frac{t}{T_{\text{kin}}}, k_j\right) \mathrm{d}\zeta(m) + O(L^{-\nu}),$$

where $\tilde{m}(t, k)$ is the solution to (WKE) with initial data $\tilde{m}(0, k) = m(k)$. It remains to show that

$$n_r(t, k_1, \ldots, k_r) = \int_{\mathcal{A}} \prod_{j=1}^{r} \tilde{m}(t, k_j) \, \mathrm{d}\zeta(m).$$

In fact, for any $m \in \mathcal{A}$, by definition $\tilde{m}$ is the solution to (WKE) with initial data m, so by direct calculation we see that $(\tilde{m})^{\otimes r}$ is a solution to (WKH) with initial data $m^{\otimes r}$. Since (WKH) is linear, we know that

$$\int_{\mathcal{A}} (\tilde{m})^{\otimes r} \, \mathrm{d}\zeta(m) \tag{10.4}$$

is a solution to (WKH) with initial data

$$\int_{\mathcal{A}} m^{\otimes r} \, \mathrm{d}\zeta(m) = (n_r)_{\text{in}},$$

due to (10.3). Moreover, since $m \in \mathcal{A}$, for short time $t \in [0, \delta]$, the solution (10.4) clearly belongs to the class $C_t^0 \mathfrak{L}_{s,\epsilon}^{\infty}$ defined in [31]; since short time uniqueness is proved in [31] for solutions in this class, we know that (10.4) has to equal n_r, which is the unique solution to (WKH) in this class with initial data $(n_r)_{\text{in}}$. The admissibility condition (1.15) follows from the fact that n_r equals (10.4) which is an average of factorized solutions for which admissibility is clearly true (Here we use that the (WKE) conserves the total mass $\int n(t, \xi) \, d\xi$). This completes the proof. ∎

References

[1] Ampatzoglou, I., Collot, C., Germain, P.: Derivation of the kinetic wave equation for quadratic dispersive problems in the inhomogeneous setting. arXiv:2107.11819v3 (2024)

[2] Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S.: Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations. Ann. of Math. (2) **198**, 1047–1201 (2023) Zbl 1509.82088 MR 4660136

[3] Brout, R., Prigogine, I.: Statistical mechanics of irreversible processes. VIII. General theory of weakly coupled systems. Physica **22**, 621–636 (1956) Zbl 0073.21403 MR 0104380

[4] Buckmaster, T., Germain, P., Hani, Z., Shatah, J.: Onset of the wave turbulence description of the longtime behavior of the nonlinear Schrödinger equation. Invent. Math. **225**, 787–855 (2021) Zbl 1484.35345 MR 4296350

[5] Chibbaro, S., Dematteis, G., Josserand, C., Rondoni, L.: Wave-turbulence theory of four-wave nonlinear inter-actions. Phys. Rev. E **96**, article no. 021101, 5 pp. (2017)

[6] Chibbaro, S., Dematteis, G., Rondoni, L.: 4-wave dynamics in kinetic wave turbulence. Phys. D **362**, 24–59 (2018) Zbl 1386.82051 MR 3724926

[7] Choi, Y., Jo, S., Kwon, Y.-S., Nazarenko, S.: Nonstationary distributions of wave intensities in wave turbulence. J. Phys. A **50**, article no. 355502, 8 pp. (2017) Zbl 1469.76056 MR 3688106

[8] Choi, Y., Lvov, Y. V., Nazarenko, S.: Joint statistics of amplitudes and phases in wave turbulence. Phys. D **201**, 121–149 (2005) Zbl 1143.76451 MR 2117755

[9] Collot, C., Germain, P.: Derivation of the homogeneous kinetic wave equation: longer time scales. arXiv:2007.03508 (2020)

[10] Collot, C., Germain, P.: On the derivation of the homogeneous kinetic wave equation. arXiv:1912.10368v3 (2021)

[11] Deng, Y., Hani, Z.: On the derivation of the wave kinetic equation for NLS. Forum Math. Pi **9**, article no. e6, 37 pp. (2021) Zbl 1479.35771 MR 4291361

[12] Deng, Y., Hani, Z.: Full derivation of the wave kinetic equation. Invent. Math. **233**, 543–724 (2023) Zbl 1530.35276 MR 4607721

[13] Deng, Y., Hani, Z.: Derivation of the wave kinetic equation: full range of scaling laws. arXiv:2301.07063 (2023)

[14] Deng, Y., Nahmod, A. R., Yue, H.: Random tensors, propagation of randomness, and nonlinear dispersive equations. Invent. Math. **228**, 539–686 (2022) Zbl 1506.35208 MR 4411729

[15] Dymov, A., Kuksin, S.: Formal expansions in stochastic model for wave turbulence 1: Kinetic limit. Comm. Math. Phys. **382**, 951–1014 (2021) Zbl 1482.35211 MR 4227166

[16] Dymov, A., Kuksin, S.: Formal expansions in stochastic model for wave turbulence 2: Method of diagram decomposition. J. Statist. Phys. **190**, article no. 3, 42 pp. (2023) Zbl 1504.35474 MR 4499050

[17] Dymov, A., Kuksin, S., Maiocchi, A., Vlăduţ, S.: The large-period limit for equations of discrete turbulence. Ann. Henri Poincaré **24**, 3685–3739 (2023) Zbl 1529.35464 MR 4653398

[18] Eyink, G. L., Shi, Y.-K.: Kinetic wave turbulence. Phys. D **241**, 1487–1511 (2012) Zbl 1287.82022

[19] Faou, E.: Linearized wave turbulence convergence results for three-wave systems. Comm. Math. Phys. **378**, 807–849 (2020) Zbl 1446.76118 MR 4134935

[20] Gallagher, I., Saint-Raymond, L., Texier, B.: From Newton to Boltzmann: hard spheres and short-range potentials. Zurich Lect. Adv. Math., European Mathematical Society, Zürich (2013) Zbl 1315.82001 MR 3157048

[21] Grad, H.: Principles of the kinetic theory of gases. In: Handbuch der Physik 12, Springer, Berlin, 205–294 (1958) MR 0135535

[22] Hasselmann, K.: On the non-linear energy transfer in a gravity-wave spectrum. I. General theory. J. Fluid Mech. **12**, 481–500 (1962) Zbl 0107.21402 MR 0136205

[23] Hasselmann, K.: On the non-linear energy transfer in a gravity wave spectrum. II. Conservation theorems; wave-particle analogy; irreversibility. J. Fluid Mech. **15**, 273–281 (1963) Zbl 0116.43401 MR 0151308

[24] Hewitt, E., Savage, L. J.: Symmetric measures on Cartesian products. Trans. Amer. Math. Soc. **80**, 470–501 (1955) Zbl 0066.29604 MR 0076206

[25] Jakobsen, P., Newell, A. C.: Invariant measures and entropy production in wave turbulence. J. Statist. Mech. **2004**, article no. L10002, 11 pp.

[26] Lanford, O. E., III: Time evolution of large classical systems. In: Dynamical systems, theory and applications (Seattle, WA, 1974), Lecture Notes in Phys. 38, Springer, Berlin, 1–111 (1975) MR 0479206

[27] Lukkarinen, J., Spohn, H.: Weakly nonlinear Schrödinger equation with random initial data. Invent. Math. **183**, 79–188 (2011) Zbl 1235.76136 MR 2755061

[28] Nazarenko, S.: Wave turbulence. Lecture Notes in Phys. 825, Springer, Heidelberg (2011) Zbl 1220.76006 MR 3014432

[29] Newell, A. C., Nazarenko, S., Biven, L.: Wave turbulence and intermittency. Phys. D **152/153**, 520–550 (2001) Zbl 1049.76033 MR 1837928

[30] Peierls, R. E.: Zur kinetischen Theorie der Wärmeleitung in Kristallen. Ann. Physik **3**, 1055–1101 (1929) Zbl 55.0547.01

[31] Rosenzweig, M., Staffilani, G.: Uniqueness of solutions to the spectral hierarchy in kinetic wave turbulence theory. Phys. D **433**, article no. 133148, 16 pp. (2022) Zbl 07487492 MR 4383192

[32] Spohn, H.: Boltzmann hierarchy and Boltzmann equation. In: Kinetic theories and the Boltzmann equation (Montecatini, 1981), Lecture Notes in Math. 1048, Springer, Berlin, 207–220 (1984) Zbl 0559.76073 MR 0740726

[33] Staffilani, G., Tran, M.-B.: On the wave turbulence theory for a stochastic KdV type equation. arXiv:2106.09819v5 (2024)

[34] Zakharov, V. E., L'vov, V. S., Falkovich, C.: Kolmogorov spectra of turbulence I: Wave turbulence. Springer, Berlin (1992) Zbl 0786.76002