Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Quantum diffusion of the random Schrödinger evolution in the scaling limit I. The non-recollision diagrams.

by

László Erdős, Manfred Salmhofer, and Horng-Tzer Yau

Preprint no.: 8 2006

Quantum diffusion of the random Schrödinger evolution in the scaling limit I. The non-recollision diagrams.

László Erdős[1]*, Manfred Salmhofer[2], Horng-Tzer Yau[3]†

[1] Mathematisches Institut, Universität München, Theresienstr. 39, D-80333 Munich, Germany

[2] Max–Planck Institute für Mathematik, Inselstr. 22, 04103 Leipzig, and

Theoretische Physik, Universität Leipzig, Postfach 100920, 04009 Leipzig, Germany

[3] Department of Mathematics, Harvard University, Cambridge MA-02138, USA

Abstract

We consider random Schrödinger equations on $\mathbb{R}^d$ for $d \geq 3$ with a homogeneous Anderson-Poisson type random potential. Denote by λ the coupling constant and ψ_t the solution with initial data ψ_0. The space and time variables scale as $x \sim \lambda^{-2-\kappa/2}, t \sim \lambda^{-2-\kappa}$ with $0 < \kappa < \kappa_0(d)$. We prove that, in the limit $\lambda \to 0$, the expectation of the Wigner distribution of ψ_t converges weakly to the solution of a heat equation in the space variable x for arbitrary L^2 initial data.

The proof is based on analyzing the phase cancellations of multiple scatterings on the random potential by expanding the propagator into a sum of Feynman graphs. In this paper we consider the non-recollision graphs and prove that the amplitude of the *non-ladder* diagrams is smaller than their "naive size" by an extra λ^c factor *per non-(anti)ladder vertex* for some $c > 0$. This is the first rigorous result showing that the improvement over the naive estimates on the Feynman graphs grows as a power of the small parameter with the exponent depending linearly on the number of vertices. This estimate allows us to prove the convergence of the perturbation series. The analysis of the recollision graphs is given in the companion paper [19].

AMS 2000 Subject Classification: 60J65, 81T18, 82C10, 82C44

*Partially supported by NSF grant DMS-0200235 and EU-IHP Network "Analysis and Quantum" HPRN-CT-2002-0027.

†Partially supported by NSF grant DMS-0307295 and MacArthur Fellowship.

1 Introduction

The fundamental equations governing the basic laws of physics, the Newton and the Schrödinger equations, are time reversible and have no dissipation. It is remarkable that dissipation is nevertheless ubiquitous in nature, so that almost all macroscopic phenomenological equations are dissipative. The oldest such example is perhaps the equation of heat conductance found by Fourier.

On a microscopic level, Brown observed almost two centuries ago that the motion of a pollen suspended in water was erratic [5]. This led to the kinetic explanation by Einstein in 1905 [14] that such a motion was created by the constant "kicks" on the relatively heavy pollen by the light water molecules. It should be noted that at that time even the atomic-molecular structure of matter was not universally accepted. Einstein's theory was strongly supported by Boltzmann's kinetic theory, which, however, was phenomenological and seriously debated at the time. Finally in 1908 Perrin [40] experimentally verified Einstein's theory and used it, among others, to give a precise estimate on the Avogadro number. These experiments gave the strongest evidence for atoms and molecules at that time.

In Einstein's kinetic theory both the heavy particle (the pollen) and the light particles (the water molecules) obey Newton's law. Therefore, Einstein's kinetic theory in fact postulated the emergence of the Brownian motion from a classical non-dissipative reversible dynamics. The key difficulty of a mathematically rigorous derivation of Brownian motion from reversible dynamics is similar to the justification of Boltzmann's molecular chaos assumption (Stoßzahlansatz); the dissipative character emerges only in a scaling limit, as the number of degrees of freedom goes to infinity.

The first mathematical definition of the Brownian motion was given in 1923 by Wiener, who constructed the Brownian motion as a scaling limit of random walks. This construction was built upon a stochastic microscopic dynamics which by itself is dissipative. The derivation of the Brownian motion from a time reversible Hamiltonian system, however, was not seriously considered until more than half a century later. Kesten-Papanicolaou [33] in 1978 proved that the velocity distribution of a particle moving in an environment consisting of random scatterers (i.e., Lorenz gas with random scatterers) converges to a Brownian motion in a weak coupling limit for $d \geq 3$. In this model the bath of light particles is replaced with random static impurities. The same result was obtained later in $d = 2$ dimensions by Dürr, Goldstein and Lebowitz [13]. In a recent work [36], Komorowski and Ryzhik have controlled the same evolution on a longer time scale and proved the convergence to Brownian motion of the position process as well. Bunimovich and Sinai [9] proved the convergence of the periodic Lorenz gas with a hard core interaction to a Brownian motion in 1980. In this model the only source of randomness is the distribution of the initial condition. Finally, Dürr, Goldstein and Lebowitz [12] proved that the velocity process of a heavy particle in a light ideal gas, which is a model with a dynamical environment, converges to the Ornstein-Uhlenbeck process.

Wiener's construction of Brownian motion is based on a random walk. The random walk could easily be replaced by the Markovian process generated by a linear Boltzmann equation. The linear Boltzmann equation was rigorously derived from the classical Lorenz gas by Gallavotti [29], Spohn [46] and Boldrighini, Bunimovich and Sinai [4]. (The nonlinear Boltzmann equation was derived by Lanford [37] for short time.) Although Brownian motion was discovered and theorized in the context of classical dynamics, we shall prove that it also describes the motion of a quantum particle in a random environment.

The random Schrödinger equation, or the quantum Lorentz model, is given by the evolution equation:

$$i\partial_t \psi_t(x) = H\psi_t(x), \qquad H = H_\omega = -\frac{1}{2}\Delta_x + \lambda V_\omega(x) \ , \tag{1.1}$$

where λ is the coupling constant and V_ω is the random potential. The first scale with a non-trivial limiting dynamics is the weak coupling limit, $\lambda \to 0$, where the space, time and the coupling constant are subject to the kinetic scaling:

$$t \to t\varepsilon^{-1}, \quad x \to x\varepsilon^{-1}, \quad \lambda = \sqrt{\varepsilon} \ . \tag{1.2}$$

Under this limit, the appropriately rescaled Wigner distribution (see (2.11)) of the solution to the Schrödinger evolution (1.1) converges weakly to a linear Boltzmann equation. This was first established by Spohn [45] for a random potential with Gaussian distribution and small macroscopic time. This method was extended to study higher order correlations in [28]. A different method (applicable to the lattice setting and general random potential, see remarks later on) was developed in [16] where the short time restriction was removed. This method was also extended to the phonon case [15] and Lukkarinen and Spohn [39] have employed a similar technique for studying the energy transport in a harmonic crystal with weakly perturbed random masses.

Since the long time limit of a Boltzmann equation is a heat equation, we shall take a time scale longer (see (2.17)) than in the kinetic scaling limit (1.2). Our aim is to prove that the limiting dynamics of the Schrödinger evolution in a random potential under this scaling is governed by a heat equation. This requires to control the Schrödinger dynamics up to a time scale $\lambda^{-2-\kappa}$, $\kappa > 0$. Quantum correlations that are small on the kinetic scale and are neglected in the first limit may contribute on the longer time scale. The derivation of the heat equation is thus much more difficult than first deriving the Boltzmann equation from Schrödinger dynamics on the kinetic scale and then showing that Boltzmann equation converges to a diffusive equation under a different limiting procedure. Notice that the limit in our approach is a long time scaling limit which involves *no semiclassical limit.*

The approach of this paper applies also to lattice models and yields a derivation of Brownian motion from the Anderson model [17, 18]. The dynamics of the Anderson model was postulated by Anderson to be localized for large coupling constant λ and extended for small coupling constant (away from the band edges and in dimension $d \geq 3$). The localization conjecture was first established rigorously by Goldsheid, Molchanov and Pastur [32] in one dimension, by Fröhlich-Spencer [25], and later by Aizenman-Molchanov [1] in several dimensions, and many other works have since contributed to this field. The progress for the extended state conjecture, however, has so far been very limited. It was proved by Klein [34] that all eigenfuctions are extended on the Bethe lattice (see also [2, 24]). In Euclidean space, Schlag, Shubin and Wolff [44] proved that the eigenfunctions cannot be localized in a region smaller than $\lambda^{-2+\delta}$ for some $\delta > 0$ in $d = 2$. Chen [10], partly based on the method [16], extended this result to all dimensions $d \geq 2$ and $\delta = 0$ (with logarithmic corrections).

A special class of random Schrödinger equation was proposed to understand the dynamics in the extended region. Instead of a spatially stationary random potential, one considers a random potential $V_\omega(x)$ with a power law decay, i.e.,

$$V_\omega(x) = h(x)\omega_x \ , \qquad h(x) \sim |x|^{-\eta}$$

where $\omega_x, x \in \mathbb{Z}^d$ are mean zero i.i.d. random variables and $\eta > 0$ is a fixed parameter.

If $\eta \geq 1$, i.e, the random potential decays at least as fast as the Coulomb potential, a standard scattering argument yields that for λ small enough H_ω has absolutely continuous spectrum. Using cancellation properties of the random potential, Rodnianski and Schlag [41] have improved the same result to $\eta > 3/4$ in $d \geq 2$ and recently, J. Bourgain [8] has extended it to $\eta > 1/2$ (see also [11]). Notice that the expected number of collisions for this model with $\eta > 1/2$ is of order one.

In summary, all known results [44, 41, 8, 10] for the Anderson model (or its modifications) in Euclidean space are in regions where the dynamics have either no effective collision or there are typically only finitely many of them. Under the diffusive scaling of this paper, see (2.17), the number of effective scatterings is a negative fractional power of the scaling parameter. In particular, it goes to infinity in the scaling limit, as it should be the case if we aim to obtain a Brownian motion. As in [10], one may derive from our dynamical result that the eigenfunctions cannot be localized in a region smaller than $\lambda^{-2-\kappa/2}$ and dimension $d \geq 3$.

Acknowledgement. The authors are grateful for the financial support and hospitality of the Erwin Schrödiger Institut, Vienna, Max Planck Institut, Leipzig, Stanford University and Harvard University, where part of this work has been done.

2 Statement of the main result

2.1 Notations

Let

$$H := -\frac{1}{2}\Delta + \lambda V \tag{2.1}$$

denote a random Schrödinger operator acting on $L^2(\mathbb{R}^d)$, $d \geq 3$, with a random potential $V = V_\omega(x)$ and a small positive coupling constant λ. The potential is given by

$$V_\omega(x) := \int_{\mathbb{R}^d} B(x-y)\mathrm{d}\mu_\omega(y)\ , \tag{2.2}$$

where B is a single site potential profile and μ_ω is a Poisson point process on $\mathbb{R}^d$ with homogeneous unit density and with independent, identically distributed random masses. More precisely, for almost all realizations ω consists of a countable, locally finite collection of points, $\{y_\gamma(\omega) \in \mathbb{R}^d \ : \ \gamma = 1, 2, \ldots\}$, and random weights $\{v_\gamma(\omega) \in \mathbb{R} \ : \ \gamma = 1, 2, \ldots\}$ such that the random measure is given by

$$\mu_\omega = \sum_{\gamma=1}^{\infty} v_\gamma(\omega)\delta_{y_\gamma(\omega)}\ , \tag{2.3}$$

where δ_y denotes the Dirac mass at $y \in \mathbb{R}^d$. The Poisson process $\{y_\gamma(\omega)\}$ is independent of the weights $\{v_\gamma(\omega)\}$. The weights are real i.i.d. random variables with distribution $\mathbf{P}_v$ and with moments $m_k := \mathbf{E}_v\, v_\gamma^k$ satisfying

$$m_2 = 1, \ \ m_{2d} < \infty, \quad m_1 = m_3 = m_5 = 0\,. \tag{2.4}$$

The expectation with respect to the random process $\{y_\gamma, v_\gamma\}$ is denoted by $\mathbf{E}$.

For the single-site potential, we assume that B is a spherically symmetric Schwarz function with 0 in the support of its Fourier transform, i.e.

$$0 \in \text{supp}\,(\widehat{B})\ . \tag{2.5}$$

More precisely, we introduce the norm

$$\|f\|_{m,n} := \sum_{|\alpha|\leq n} \|\langle x\rangle^m \partial^\alpha f(x)\|_\infty$$

with $\langle x\rangle := (2+x^2)^{1/2}$ (here α is a multiindex) and we assume

$$\|B\|_{k,k} < C_k \qquad \forall k \in \mathbf{N}\ . \tag{2.6}$$

Actually, it is sufficient to assume (2.6) for all $k \leq k_0(\kappa)$.

We note that the operator H_ω is not bounded from below due to the possible large concentration of Poisson points in some region. Nevertheless, H_ω is self-adjoint under very general conditions, see [35].

We introduce a few notational conventions. The letters x, y, z will always denote configuration space variables, while p, q, r, u, v, w will be reserved for d-dimensional momentum variables. The norm without indices, $\|\cdot\|$, will always denote the standard $L^2(\mathbb{R}^d)$ norm. The bracket $(\cdot\,,\cdot)$ denotes the standard scalar product on $L^2(\mathbb{R}^d)$ and $\langle\cdot\,,\cdot\rangle$ will denote the pairing between the Schwarz space and its dual on the phase space $\mathbb{R}^d \times \mathbb{R}^d$.

Integrals without explicit domains will always denote integration over $\mathbb{R}^d$ with respect to the Lebesgue measure. For any $f \in L^2(\mathbb{R}^d)$, the Fourier transform is given by

$$\widehat{f}(p) := \int e^{-2\pi i p\cdot x} f(x)\mathrm{d}x\ , \qquad p \in \mathbb{R}^d\ . \tag{2.7}$$

and the inverse Fourier transform is given by

$$g(x) = \int \widehat{g}(p) e^{2\pi i p\cdot x}\mathrm{d}p\ , \qquad x \in \mathbb{R}^d\ .$$

For functions defined on the phase space, $f(x,v)$, the Fourier transform will always be taken only in the space variable, i.e.

$$\widehat{f}(\xi, v) := \int e^{-2\pi i \xi\cdot x} f(x,v)\mathrm{d}x\ , \qquad \xi \in \mathbb{R}^d\ .$$

The Fourier transform of the kinetic energy operator is given by

$$\widehat{\big[-\frac{1}{2}\Delta f\big]}\ (p) = e(p)\widehat{f}(p)\ ,$$

where

$$e(p) := \frac{1}{2}\,p^2 \tag{2.8}$$

is the dispersion law. The velocity is given by $\frac{1}{2\pi}\nabla e(p) = \frac{1}{2\pi}\, p$.

Define the *Wigner transform* of a function $\psi \in L^2(\mathbb{R}^d)$

$$W_\psi(x,v) := \int e^{2\pi i v\cdot\eta}\overline{\psi(x+\frac{\eta}{2})}\psi(x-\frac{\eta}{2})\mathrm{d}x\ . \tag{2.9}$$

The Fourier transform of $W_\psi(x,v)$ in the x variable is therefore

$$\widehat{W}_\psi(\xi,v) = \overline{\widehat{\psi}\Big(v-\frac{\xi}{2}\Big)}\widehat{\psi}\Big(v+\frac{\xi}{2}\Big)\ . \tag{2.10}$$

Define the rescaled Wigner distribution as

$$W^\varepsilon_\psi(X,V) := \varepsilon^{-d}W_\psi\Big(\frac{X}{\varepsilon},V\Big)\ . \tag{2.11}$$

Its Fourier transform in X is given by

$$\widehat{W}^\varepsilon_\psi(\xi,V) = \overline{\widehat{\psi}\Big(V-\frac{\varepsilon\xi}{2}\Big)}\widehat{\psi}\Big(V+\frac{\varepsilon\xi}{2}\Big)\ .$$

For any function $h:\mathbb{R}^d\to\mathbf{C}$ and energy value $e\geq 0$ we introduce the notation

$$[h](e) := \int h(v)\delta(e-e(v))\mathrm{d}v := \int_{\Sigma_e} h(q)\,\frac{\mathrm{d}\nu(q)}{|\nabla e(q)|}\ , \tag{2.12}$$

where $\mathrm{d}\nu(q)$ is the restriction of the Lebesgue measure onto the energy surface $\Sigma_e := \{q\ :\ e(q) = e\}$ that is the ball of radius $\sqrt{2e}$. More explicitly,

$$[h](e) := (2e)^{\frac{d}{2}-1}\int_{S^{d-1}} h(\sqrt{2e}\phi)\mathrm{d}\phi\ .$$

Clearly

$$\int h(v)\mathrm{d}v = \int_0^\infty [h](e)\mathrm{d}e\ . \tag{2.13}$$

The normalization of the measure $[\cdot]_e$ is given by

$$[1](e) := c_{d-1}(2e)^{\frac{d}{2}-1}\ , \tag{2.14}$$

where c_{d-1} is the volume of the unit sphere S^{d-1}.

2.2 Main Theorem

The weak coupling limit is defined by the following scaling:

$$t = \mathcal{T}/\varepsilon,\quad x = \mathcal{X}/\varepsilon,\quad \varepsilon = \lambda^2\ . \tag{2.15}$$

The Wigner distribution $W^{\varepsilon}_{\psi_{\mathcal{T}/\varepsilon}}(\mathcal{X},\mathcal{V})$ converges weakly to a function $F_{\mathcal{T}}(\mathcal{X},V)$ that satisfies the Boltzmann equation

$$\begin{aligned}&\partial_{\mathcal{T}} F_{\mathcal{T}}(\mathcal{X},V) + V\cdot\nabla_{\mathcal{X}} F_{\mathcal{T}}(\mathcal{X},V)\\ &\qquad = 2\pi\int \mathrm{d}U|\widehat{B}(U-V)|^2\delta(e(U)-e(V))\Big[F_{\mathcal{T}}(\mathcal{X},U)-F_{\mathcal{T}}(\mathcal{X},V)\Big]\ .\end{aligned}\tag{2.16}$$

Note that the Boltzmann equation can be viewed as the generator of a Markovian semigroup on phase space. The proof of (2.16) for continuum Gaussian model was given in [16]; the $\mathbb{Z}^d$ lattice case with general i.i.d. random potential was considered in [10]. The derivation of the Boltzmann equation for potential (2.2) follows from these two proofs in a straightforward way.

In this paper we consider the long time scaling

$$t=\lambda^{-\kappa}\big(\lambda^{-2}T\big),\quad x=\lambda^{-\kappa/2}\big(\lambda^{-2}X\big)=X/\varepsilon,\quad \varepsilon=\lambda^{\kappa/2+2}\tag{2.17}$$

with some $\kappa>0$. This scaling corresponds to the long time limit of the Boltzmann equation with diffusive scaling.

For some energy $e>0$, let

$$L_e f(v) := \int \mathrm{d}u\ \sigma(u,v)[f(u)-f(v)],\qquad e(v)=e\ ,\tag{2.18}$$

be the generator of the momentum jump process $v(t)$ on Σ_e with collision kernel

$$\sigma(u,v) := 2\pi|\widehat{B}(u-v)|^2\delta(e(u)-e(v))\ .\tag{2.19}$$

A well-known argument shows that $B\not\equiv 0$ and the regularity of B guarantees the following properties. Some details will be given in [19].

Lemma 2.1 *For each $e>0$ the Markov process $\{v(t)\}_{t\geq 0}$ with generator L_e is uniformly exponentially mixing. The unique invariant measure is the uniform distribution, $[\,\cdot\,](e)/[1](e)$, on the energy surface Σ_e.*

Let

$$D_{ij}(e) := \frac{1}{(2\pi)^2}\int_0^\infty \mathcal{E}_e\big[v^{(i)}(t)v^{(j)}(0)\big]\mathrm{d}t\ ,\qquad v=(v^{(1)},\dots,v^{(d)}),\quad i,j=1,2,\dots d,$$

be the velocity autocorrelation matrix, where $\mathcal{E}_e$ denotes the expectation with respect to this Markov process in equilibrium. By the spherical symmetry of $\widehat{B}$ and $e(U)$, the autocorrelation matrix is constant times the identity:

$$D_{ij}(e)=D_e\ \delta_{ij},\qquad D_e := \frac{1}{(2\pi)^2 d}\int_0^\infty \mathcal{E}_e\big[v(t)\cdot v(0)\big]\mathrm{d}t\ .\tag{2.20}$$

The main result of the paper is the following theorem.

Theorem 2.2 *Let $d \geq 3$ and $\psi_0 \in L^2(\mathbb{R}^d)$ be a normalized initial wave function. Let $\psi(t) := \exp(-itH)\psi_0$ solve the Schrödinger equation (1.1). Let $\mathcal{O}(x,v)$ be a Schwarz function on $\mathbb{R}^d \times \mathbb{R}^d$. For any $e > 0$, let f be the solution to the heat equation*

$$\partial_T f(T,X,e) = \frac{D_e}{2}\Delta_X f(T,X,e) \tag{2.21}$$

with the initial condition

$$f(0,X,e) := \delta(X)\Big[|\widehat{\psi}_0(v)|^2\Big](e)\ .$$

Then there exist $0 < \kappa_0(d) \leq 2$ such that for $0 < \kappa < \kappa_0(d)$ and for ε and λ related by (2.17)*, the rescaled Wigner distribution satisfies*

$$\lim_{\lambda\to 0}\int \mathrm{d}X \int \mathrm{d}v\ \mathcal{O}(X,v)\mathbf{E}W^{\varepsilon}_{\psi(\lambda^{-\kappa-2}T)}(X,v) = \int \mathrm{d}X \int \mathrm{d}v\ \mathcal{O}(X,v) f(T,X,e(v))\ , \tag{2.22}$$

and the limit is uniform on $T \in [0,T_0]$ with any fixed T_0. In $d = 3$ one can choose $\kappa_0(3) = 1/500$.

Remark 1. The coefficient $\big[|\widehat{\psi}_0(v)|^2\big](e)$ in the initial condition $f(0,X,e)$ is finite for almost all e by using (2.13) for $h = |\widehat{\psi}_0|^2$.

Remark 2. The total cross section of the collision process (2.18),

$$\sigma_0(e) := \int \mathrm{d}u\ \sigma(u,v) \qquad e = e(v)\ , \tag{2.23}$$

is a function of $e = e(v)$ only. Assuming $\widehat{B}(0) \neq 0$, we see that $\sigma_0(e) \sim [1](e)$ for small e, and $\sigma_0(e) \sim e^{-1/2}$ for large e. It follows from Lemma 2.1 and from standard probability arguments that the diffusion constant (2.20) scales as $D_e \sim e/\sigma_0(e)$ for small $e \ll 1$ and $D_e \sim e^2/\sigma_0(e)$ for large $e \gg 1$. If $\widehat{B}$ vanishes at 0 (but (2.5) still holds), then the small energy behaviour of $\sigma_0(e)$ and D_e depends on the rate of vanishing of $\widehat{B}$ at 0 in a straighforward way.

Remark 3. The condition (2.5) is not essential, but the theorem needs to be modified if $\widehat{B}$ vanishes on $D(0,\delta)$, a ball of radius $\delta > 0$ about the origin. Let $\delta > 0$ be the maximal radius so that $D(0,\delta) \cap \mathrm{supp}\,(\widehat{B}) = \emptyset$. In this case the total cross section $\sigma_0(e)$ is zero for all energy values $e \leq \delta^2/8$, because the diameter of the energy surface Σ_e is smaller than the minimal range of $\widehat{B}$. Therefore the evolution is ballistic for the part of the initial wave function that is supported on energy shells $e \leq \delta^2/8$. For the other part of the wave function the diffusion equation still holds.

Fig. 1 below shows the three different scales schematically. On the Schrödinger scale both time and space are of order 1 in atomic units. On the kinetic scale time and space are rescaled by λ^{-2}. The dynamics is given by the Boltzmann equation characterized by finitely many collisions. On the diffusive scale we rescaled the time and space by an additional factor $\lambda^{-\kappa}$ and $\lambda^{-\kappa/2}$ respectively. The typical number of collisions is of order $\lambda^2 t \sim \lambda^{-\kappa}$.

If we assume that the Boltzmann equation holds under all scalings, Theorem 2.2 can be easily understood. From the Boltzmann equation (2.16), the velocity distribution develops according to the Markovian generator L_e. Therefore, the Boltzmann equation (2.16) describes a process that a

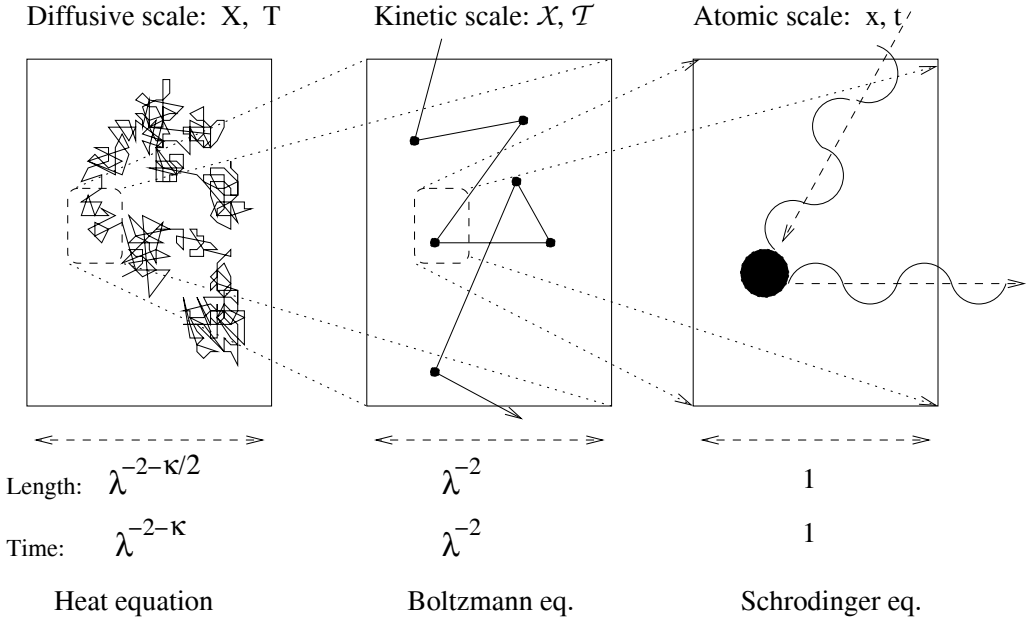

Figure 1: Evolution equations on three scales

particle travels with a fixed velocity v up to an exponentially distributed random time with average value $\sigma_0(e(v))^{-1}$, then it changes velocity from v to a new velocity u on the same energy surface Σ_e chosen by the probability distribution $P(u) = \sigma(u, v)/\sigma_0(e(v))$. The different energy sectors do not interact. Clearly, this process then converges to a Brownian motion in configuration space with a diffusion coefficient given by (2.20) and with momentum restricted to a fixed energy shell Σ_e.

Under the assumption that the Boltzmann equation is valid for all time, this argument applies in all dimensions. The random Schrödinger evolution, however, is expected to be localized for $d \leq 2$ even for small coupling constant. Therefore, even though the Boltzmann approximation was proved to be valid for $d \geq 2$ [16], [10] (it is not valid for $d = 1$) in the weak coupling limit, it will not be valid for all time in $d = 2$. It is expected that memory effects and quantum correlations eventually dominate the evolution and ruin the Markovian character of the Boltzmann picture. Heuristic ideas show this transition happens at an exponentially large time (see, e.g. [47]).

The effects of the quantum correlations and memory are not expected to change the Boltzmann picture drastically in $d \geq 3$, but one expects corrections to the diffusion equation and a transition between different energy shells for $\kappa \geq 2$ (see [38]).

2.3 Strategy of the proof

The above heuristic argument using the Boltzmann equation, besides being misleading for $d = 2$, also masks the difficulties in proving Theorem 2.2, namely that one has to follow the full *quantum mechanical* time evolution through infinitely many collisions. The main tool of our proof is to use the Duhamel expansion to decompose the wave function into elementary wave functions char-

acterized by their collision histories. We then apply two strategies to simplify the expansion: (i) renormalization of the propagator, i.e., resumming the two legged subdiagrams; (ii) stopping rules to control recollisions. Apart from these two steps, the bulk of our proof is devoted to giving sharp estimates for a large class of Feynman graphs.

To get an idea, imagine that we expand the solution to the Schrödinger equation by using the Duhamel formula repeatedly. This rewrites the solution into a sum of elementary wave functions, each of which is characterized by a sequence of collisions with random obstacles. When we take the expectation of $\|\psi_t\|^2$ with respect to a Gaussian randomness, we pair the random obstacles by Wick's theorem and obtain a sum of amplitudes of Feynman graphs. In case of a non-Gaussian randomness the higher order cumulants are also present (their contribution turns out to be negligible, but proving this is rather involved).

If we take only the Laplacian as the free part in the expansion, even the amplitudes of individual graphs diverge in the limit we consider. However, this can be remedied by a simple resummation of all two-legged insertions caused by the lowest order self-energy contribution (it turns out that higher order corrections to the self-energy do not play a role in the scaling limit we consider). The resummation is performed by choosing an appropriate reference Hamiltonian H_0 for the expansion. After this rearrangement, all graphs have a finite amplitude in our scaling limit, and the ladder graphs give the leading contribution.

However, we have to estimate not only individual graphs but the sum of all graph amplitudes, which requires beating down the factorial growth of the number of graphs. This problem has been addressed in constructive field theory. For field theories with bosons, the graphical expansion to infinite order diverges. Borel summability was proven by cluster expansion and renormalization group methods [30, 22, 7, 6]. In fermionic theories, the anticommutation relations entail cancellations which lead to analyticity in the presence of regulators [31, 23, 20, 42]. Our method to control the combinatorial growth is completely different: it is by very sharp bounds on the individual graphs. We give a classification of arbitrary large graphs, based on counting the number of vertices carrying extra oscillatory effects. The number of these vertices is called the *degree* of the graph and it measures the improvement over the standard power counting. For the ladder graphs, the degree is zero, for the anti-ladder (i.e., complete crossing) graph it is 2. For general graphs, the degree is roughly the number of vertices after removing all ladder and anti-ladder subgraphs. We thus obtain an extra λ^c factor (for some $c > 0$) *per non-(anti)ladder vertex*. This strong improvement is sufficient to beat the growth of the combinatorics in the time scale we consider. To our knowledge, nothing like this has been done in a graphical expansion before. Improved phase space estimates have been used to prove regularity in two-dimensional many–fermion systems, but the improvement exponent was fixed independently of the number of vertices [26, 27, 21].

For a comparison, the unperturbed Green functions in the perturbation expansion for the many–fermion systems and for the random Schrödinger equation are given by

$$\frac{1}{ip_0 + p^2 - \mu}, \qquad \frac{1}{p^2 - \alpha + i\eta}\ .$$

In the many–fermion case, $\mu > 0$, $p_0 \in M_F = \{\frac{\pi}{\beta}(2n+1) : n \in \mathbb{Z}\}$ where $\beta \sim T^{-1}$ is the inverse

temperature. In the random Schrödinger case, $\eta \sim t^{-1}$. Their L^2-properties are different:

$$\frac{1}{\beta}\sum_{p_0\in M_F}\int \mathrm{d}p\big|ip_0+p^2-\mu\big|^{-2}\sim|\log\beta|,\qquad \int \mathrm{d}p\big|p^2-\alpha+i\eta\big|^{-2}\sim\eta^{-1}\ .$$

Notice the divergence is more severe for the random Schrödinger equation case. In the many-fermion case, there is one p_0–summation per line of the graph; in the random Schrödinger case there are just two overall α–integrals for graphs with arbitrarily many lines.

Finally we note that the threshold $\kappa < \kappa_0(d)$ in our theorem may be improved up to $\kappa < 2/3$ simply with more detailed arguments. To go beyond $2/3$ but with $\kappa < 2$ requires substantially better estimates on the individual graphs than we have here. On the other hand, one cannot go beyond $\kappa = 2$ with only improvements on estimates of the individual graphs. The Duhamel formula must be expanded at least up to $k = \lambda^2 t = \lambda^{-\kappa}$, which is the typical number of collisions up to time t. Even if one proves for most graphs the best possible estimate, λ^{2k}, it cannot beat the $k!$ from the combinatorics when $k \gg \lambda^{-2}$, i.e., $\lambda^{2k}k! \gg 1$ for $k \gg \lambda^{-2}$. A different resummation procedure is needed beyond this threshold to exploit cancellations among these graphs.

This paper is organized as follows. In Section 3 we perform the self-energy renormalization, we smooth out the data and restrict the problem to a finite box. The Duhamel expansion is introduced in Section 4. In Section 5 we reduce the Main Theorem to Theorems 5.1, 5.2 and 5.3. The key result is Theorem 5.2 which we prove in the rest of this paper. The other two theorems are more technical and they are proven in the companion paper [19]. The Feynman graphs are introduced in Section 7. In Sections 8 and 9 we reduce all estimates to Theorem 8.4. This theorem is our main technical bound on Feynman graphs and it is proven in Section 10. Since the random potential in our model is given by general i.i.d. random variables, the rule for taking the expectation is different from the case of Gaussian random field used in [16]. In addition to the usual pairing from the Wick theorem, we have to introduce higher order partitions of the vertices, called *non-trivial lumps*. This produces significant technical difficulties and complicates our proof. We recommend the reader to ignore these lumps at first reading.

Universal constants and constants that depend only on the dimension d, on the final time T_0 and on the Schwarz norms $\|B\|_{k,k}$ from (2.6) will be denoted by C and their value may vary from line to line.

3 Preparations

3.1 Renormalization

The purpose of this procedure is to include immediate recollisions with the same obstacle into the propagator itself. This is also called the renormalization of "one-particle propagators" or two legged subdiagrams. Without renormalization, these graphs individually are exponentially large ("divergent"), but their sum is finite. Renormalization removes this instability and the analysis of the resulting Feynman graphs will become simpler.

The self-energy operator is given by the multiplication operator in momentum space

$$\theta(p) := \Theta(e(p)), \qquad \Theta(\alpha) := \lim_{\varepsilon\to 0+} \Theta_\varepsilon(\alpha)\ , \qquad \Theta_\varepsilon(\alpha) := \Theta_\varepsilon(\alpha, r) \tag{3.1}$$

for any r with $e(r) = \alpha$, where

$$\Theta_\varepsilon(\alpha, r) := \int \frac{|\widehat{B}(q-r)|^2 \mathrm{d}q}{\alpha - e(q) + i\varepsilon}\ . \tag{3.2}$$

Note that by spherical symmetry of B and $e(q)$, $\Theta_\varepsilon(\alpha, r)$ depends only on the length of r, therefore $\Theta_\varepsilon(\alpha)$ in (3.1) is well defined. Clearly $\theta(p)$ is spherically symmetric. The existence of the limit and related properties of Θ have been proven in [16] using that $\|\widehat{B}^2\|_{2d,2d} < \infty$. Here we summarize the results:

Lemma 3.1 *In $d \geq 3$ the following hold:*

$$|\Theta_\varepsilon(\alpha, r) - \Theta_e(\alpha, r')| \leq C\big|\ |r| - |r'|\ \big| \tag{3.3}$$

(Eq. (3.80) in [16]) and

$$|\Theta_\varepsilon(\alpha, r) - \Theta_{\varepsilon'}(\alpha', r)| \leq C(|\varepsilon - \varepsilon'| + |\alpha - \alpha'|)\varepsilon^{-1/2} \tag{3.4}$$

if $\varepsilon \geq \varepsilon' > 0$ (Eq. (3.68) in [16]). From this latter estimate the existence of the limit $\lim_{\varepsilon\to 0+0} \Theta_\varepsilon(\alpha, r)$ follows. Moreover, Θ is Hölder continuous

$$|\Theta(\alpha) - \Theta(\alpha')| \leq C|\alpha - \alpha'|^{1/2}\ . \tag{3.5}$$

Proof. We have only to prove the Hölder continuity. For any ε and any r, r' with $\alpha = e(r)$, $\alpha' = e(r')$ we have

$$\begin{aligned}|\Theta(\alpha) - \Theta(\alpha')| \leq & \lim_{\varepsilon'\to 0+0} |\Theta_{\varepsilon'}(\alpha, r) - \Theta_\varepsilon(\alpha, r)| + |\Theta_\varepsilon(\alpha, r) - \Theta_\varepsilon(\alpha, r')| \\ & + |\Theta_\varepsilon(\alpha, r') - \Theta_\varepsilon(\alpha', r')| + \lim_{\varepsilon'\to 0+0} |\Theta_\varepsilon(\alpha', r') - \Theta_{\varepsilon'}(\alpha', r')| \\ & \leq C\Big(\varepsilon^{1/2} + \big|\ |r| - |r'|\ \big| + |\alpha - \alpha'|\varepsilon^{-1/2}\Big)\ .\end{aligned}$$

By optimizing ε and using $e(r) = \alpha$, we obtain (3.5). □

We have the following estimate on $\theta(p)$ and in parallel on $\Theta(e)$:

Lemma 3.2 *For any $d \geq 3$ there exist universal positive constants c_1, c_2 such that*

$$|\theta(p)| \leq \frac{c_2 \log\langle p\rangle}{\langle p\rangle}\ , \qquad |\Theta(e)| \leq \frac{c_2 \log\langle e\rangle}{\langle e\rangle^{1/2}}\ , \tag{3.6}$$

$$\mathit{Im}\ \Theta(e) \leq -c_1 \min\{|e|^{\frac{d}{2}-1}, |e|^{-1/2}\}\ , \qquad \mathit{Im}\ \theta(p) \leq -c_1 \min\{|p|^{d-2}, |p|^{-1}\}\ . \tag{3.7}$$

Proof of Lemma 3.2. By performing the angular integration, we can write $\Theta_\varepsilon(\alpha, p)$ with $e(p) = \alpha$ as

$$\Theta_\varepsilon(\alpha, p) = \int_0^\infty \frac{(2e)^{\frac{d}{2}-1}\mathrm{d}e}{\alpha - e + i\varepsilon}\, S(e)\ , \qquad \text{with} \quad S(e) := \int_{S^{d-1}} |\widehat{B}(\sqrt{2e}(\phi_r - \phi))|^2 \mathrm{d}\phi\ , \tag{3.8}$$

where ϕ_r is a fixed vector on the unit sphere S^{d-1}. For small e values

$$|S(e)| = O(1), \qquad |\nabla S(e)| = O(e^{-1/2})\ .$$

For large e values, using the regularity of $\widehat{B}$,

$$|S(e)| = O(e^{-\frac{d-1}{2}}), \qquad |\nabla S(e)| = O(e^{-\frac{d}{2}})\ .$$

These estimates, in particular, stand behind the proof that $\lim_{\varepsilon\to 0+0}\Theta_\varepsilon(\alpha, p)$ is finite, since they guarantee the sufficient decay for large e and the sufficient smoothness around the singularity of the denominator in (3.8). The imaginary part therefore is

$$\mathrm{Im}\,\theta(p) = \mathrm{Im}\lim_{\varepsilon\to 0+0}\Theta_\varepsilon(\alpha, p) = -\pi(2\alpha)^{\frac{d}{2}-1}S(\alpha)$$

which behaves as $\sim -|p|^{d-2}$ for small p and as $\sim -|p|^{-1}$ for large p. The real part of $\Theta_\varepsilon(\alpha, r)$ is bounded for small α. For large α one splits the integration

$$\mathrm{Re}\,\Theta_\varepsilon(\alpha, r) = \left(\int_{\alpha-1}^{\alpha+1} + \int_{|\alpha-e|\geq 1}\right)\frac{(2e)^{\frac{d}{2}-1}\mathrm{d}e}{\alpha - e + i\varepsilon}\, S(e)\ .$$

After Taylor expanding $(2e)^{\frac{d}{2}-1}S(e)$ around α, the first term is bounded by

$$\int_{\alpha-1}^{\alpha+1}\Big|\frac{\mathrm{d}}{\mathrm{d}e}[(2e)^{\frac{d}{2}-1}S(e)]\Big|\mathrm{d}e = O(\alpha^{-1})$$

and the second term by

$$\int_{|\alpha-e|\geq 1}\frac{\mathrm{d}e}{|\alpha - e|e^{1/2}} \leq \frac{c\log\langle\alpha\rangle}{\langle\alpha\rangle^{1/2}}\ .$$

If we write $\Theta(e) = \mathcal{R}(e) - i\mathcal{I}(e)$, where $\mathcal{R}(e)$ and $\mathcal{I}(e)$ are real functions, and recall $Im(x + i0)^{-1} = -\pi\delta(x)$, we have

$$\mathcal{I}(e) = -\mathrm{Im}\,\Theta(e) = \pi\int\delta(e(q) - e)|\widehat{B}(q - r)|^2\mathrm{d}q \tag{3.9}$$

for any r satisfying $e = e(r)$. $\square$

We rewrite the Hamiltonian as

$$H = H_0 + \widetilde{V},$$

where

$$H_0 := \omega(p) := e(p) + \lambda^2\theta(p), \qquad \widetilde{V} := \lambda V - \lambda^2\theta(p)\ . \tag{3.10}$$

We note that our renormalization is only an approximation to the standard self-consistent renormalization given by the solution to the equation

$$\omega(p) = e(p) + \lambda^2 \lim_{\varepsilon\to 0+0} \int \frac{|\widehat{B}(p-q)|^2 \mathrm{d}q}{\omega(p) - \omega(q) + i\varepsilon}\ . \tag{3.11}$$

Due to our truncation procedure, the definition (3.1) is sufficient and is more convenient for us. Since $e(p)$ is spherically symmetric, so are $\theta(p)$ and $\omega(p)$.

The following lemma collects some estimates on the renormalized propagators we shall use to prove Theorem 2.2. The proof is fairly simple and will be given in [19].

Lemma 3.3 *Suppose that $\lambda^2 \geq \eta \geq \lambda^{2+4\kappa}$ with $\kappa \leq 1/12$. Then we have,*

$$\int \frac{|h(p-q)|\mathrm{d}p}{|\alpha - \omega(p) + i\eta|} \leq \frac{C\|h\|_{2d,0}\,|\log\lambda|\ \log\langle\alpha\rangle}{\langle\alpha\rangle^{1/2}\langle |q| - \sqrt{2|\alpha|}\rangle}\ , \tag{3.12}$$

and for $0 \leq a < 1$

$$\int \frac{|h(p-q)|\mathrm{d}p}{|\alpha - \omega(p) + i\eta|^{2-a}} \leq \frac{C_a\|h\|_{2d,0}\,\lambda^{-2(1-a)}}{\langle\alpha\rangle^{a/2}\langle |q| - \sqrt{2|\alpha|}\rangle}\ , \tag{3.13}$$

$$\int \frac{|h(p-q)|\mathrm{d}p}{|\alpha - e(p) + i\eta|^{2-a}} \leq \frac{C_a\|h\|_{2d,0}\,\eta^{-2(1-a)}}{\langle\alpha\rangle^{a/2}\langle |q| - \sqrt{2|\alpha|}\rangle}\ . \tag{3.14}$$

For $a = 0$ and with $h := \widehat{B}$, the following more precise estimate holds. There exists a constant C_0, depending only on finitely many constants C_k from (2.6) *such that*

$$\int \frac{\lambda^2|\widehat{B}(p-q)|^2\ \mathrm{d}p}{|\alpha - \overline{\omega}(p) - i\eta|^2} \leq 1 + C_0\lambda^{-12\kappa}\big[\lambda + |\alpha - \omega(q)|^{1/2}\big]\ . \tag{3.15}$$

3.2 Smoothing the initial data and the potential

In this section we show that it is sufficient to prove the Main Theorem under the assumptions that $\widehat{\psi}_0(p)$ is a bounded, smooth, compactly supported function and $\widehat{B}(p)$ is supported on $\{|p| \leq \lambda^{-\delta}\}$ for any fixed $\delta > 0$.

The approximation procedure relies on the following L^2-continuity property of the Wigner transform. If a random wave function is decomposed as $\psi = \psi_1 + \psi_2$, then

$$\Big|\mathbf{E}\langle\widehat{\mathcal{O}}, \widehat{W}^{\varepsilon}_{\psi}\rangle - \mathbf{E}\langle\widehat{\mathcal{O}}, \widehat{W}^{\varepsilon}_{\psi_1}\rangle\Big| \leq C\Big(\int \sup_v |\widehat{\mathcal{O}}(\xi, v)|\mathrm{d}\xi\Big)\sqrt{\mathbf{E}\big[\|\psi_1\|^2 + \|\psi_2\|^2\big]\cdot\mathbf{E}\|\psi_2\|^2} \tag{3.16}$$

by a simple Schwarz inequality. (Due to a misprint, the $\|\psi_2\|^2$ term was erroneously omitted in Section 2.1. of our earlier paper [16].)

Approximation of the initial data. Let $\widehat{\psi}_0 \in L^2$ and let $\widehat{\psi}_n$ be a sequence of smooth, compactly supported functions with $\|\widehat{\psi}_n - \widehat{\psi}_0\| \to 0$. We decompose $\widehat{\psi}_0 = \widehat{\psi}_n + (\widehat{\psi}_0 - \widehat{\psi}_n)$. Then

$$\widehat{\psi}(t) = e^{-itH}\widehat{\psi}_n + e^{-itH}(\widehat{\psi}_0 - \widehat{\psi}_n) \, .$$

Since

$$\|e^{-itH}(\widehat{\psi}_0 - \widehat{\psi}_n)\| = \|\widehat{\psi}_0 - \widehat{\psi}_n\| \to 0$$

as $n \to \infty$, uniformly in t, we see that

$$\lim_{n\to\infty} \Big| \mathbf{E}\langle \widehat{\mathcal{O}}, \widehat{W}^{\varepsilon}_{\psi(t)}\rangle - \mathbf{E}\langle \widehat{\mathcal{O}}, \widehat{W}^{\varepsilon}_{\psi_n(t)}\rangle \Big| = 0$$

uniformly in t (and thus in ε), where $\psi_n(t) := e^{-itH}\widehat{\psi}_n$ is the time evolution of the approximated initial data. This means that the approximation procedure is continuous on the left hand side of (2.22).

Similarly, on the right hand side of (2.22), we can define $f_n(T, X, e)$ to be the solution to (2.21) with initial data $f_n(0, X, e) := \delta(X)\big[|\widehat{\psi}_n|^2\big](e)$. Clearly $\big[|\widehat{\psi}_n|^2\big](e)$ converges to $\big[|\widehat{\psi}_0|^2\big](e)$ in $L^1(\mathrm{d}e)$. Therefore

$$f_n(T, X, e) \to f(T, X, e) \tag{3.17}$$

in $L^1(\mathrm{d}X\,\mathrm{d}e)$, uniformly in T. The right hand side of (2.22) is therefore also continuous as $n \to \infty$.

We remark, that if $\widehat{\psi}_0$ is smooth, e.g. $|\nabla_p \widehat{\psi}_0(p)| \leq C\langle p\rangle^{-4d}$, then a bounded, smooth and compactly supported approximant, $\widehat{\psi}_n$, can be chosen so that $\big[|\widehat{\psi}_n|^2\big](e) \to \big[|\widehat{\psi}_0|^2\big](e)$ for every $e > 0$ and then the convergence in (3.17) also holds in $L^1(\mathrm{d}X)$ for any e. The smoothness of $\widehat{\psi}_0$ is used only at the point when we explicitly compute the main term of the perturbation expansion and identify it with the Boltzmann equation, see [19].

Propagation estimate. To verify that a truncation is allowed for $\widehat{B}$, we first need a crude propagation estimate. Define the following event for any $Z > 0$

$$\Omega_Z := \Big\{ \omega \ : \ \int_{|y-k|\leq 1} \mathrm{d}|\mu_\omega|(y) \leq Z\langle k\rangle \ , \forall k \in \mathbb{Z}^d \Big\} \, ,$$

where $|\mu_\omega|$ denotes the total variation of the (random) measure μ_ω. A standard large deviation estimate on the Poisson point process and on the random weights easily gives

$$\lim_{Z\to\infty} \mathbf{P}(\Omega_Z) = 1 \, . \tag{3.18}$$

We decompose

$$\mathbf{E}\langle \widehat{\mathcal{O}}, \widehat{W^{\varepsilon}_{\psi_t}}\rangle = \mathbf{E}\left[\mathbf{1}(\Omega_Z)\langle \widehat{\mathcal{O}}, \widehat{W^{\varepsilon}_{\psi_t}}\rangle\right] + \mathbf{E}\left[\mathbf{1}(\Omega_Z^c)\langle \widehat{\mathcal{O}}, \widehat{W^{\varepsilon}_{\psi_t}}\rangle\right],$$

where $\mathbf{1}(\cdot)$ is the characteristic function. On the set Ω_Z^c we use

$$\Big|\mathbf{E}\left[\mathbf{1}(\Omega_Z^c)\langle \widehat{\mathcal{O}}, \widehat{W^{\varepsilon}_{\psi_t}}\rangle\right]\Big| \leq \Big(\int \sup_v |\widehat{\mathcal{O}}(\xi, v)|\mathrm{d}\xi\Big)\|\psi_t\|^2 \mathbf{P}(\Omega_Z^c) \to 0 \tag{3.19}$$

as $Z \to \infty$, uniformly in t (hence in λ).

For $\omega \in \Omega_Z$ we have $|V_\omega(x)| \leq CZ\langle x\rangle$ using the decay properties of B. Computing the time derivative of the mean square displacement, we obtain $\partial_t(\psi_t, x^2\psi_t) = i(\psi_t, [H, x^2]\psi_t)$. Using $[H, x^2] = -(\nabla \cdot x + x \cdot \nabla)$ and a Schwarz estimate we have

$$\left|\partial_t(\psi_t, x^2\psi_t)\right| \leq C(\psi_t, x^2\psi_t)^{1/2}\big[E + \lambda(\psi_t, |V_\omega|, \psi_t)\big]^{1/2} \tag{3.20}$$

with $E := (\psi_t, H\psi_t) = (\psi_0, H\psi_0)$ by energy conservation. We estimate $(\psi, |V_\omega|\psi) \leq CZ + CZ(\psi, x^2\psi)^{1/2}$, in particular the energy E is bounded (depending on ψ_0 and Z). From (3.20) we thus have

$$(\psi_t, x^2\psi_t) \leq c_1(Z, \psi_0)t^4 + c_2(Z, \psi_0) \tag{3.21}$$

on Ω_Z with some constants $c_{1,2}(Z, \psi_0)$.

Approximation of the potential. We define the truncation of B in Fourier space as $\widehat{B}^\delta(p) := \varphi(\lambda^\delta\langle p\rangle)\widehat{B}(p)$, where $\varphi : \mathbb{R}_+ \to [0, 1]$ is a fixed smooth cutoff function with $\varphi(a) \equiv 1$ for $a \leq 1/2$ and $\varphi(a) \equiv 0$ for $a \geq 1$. In position space, we have for any $M \in \mathbf{N}$,

$$|B(x) - B^\delta(x)| \leq \langle x\rangle^{-2d} \int \left|\langle\nabla_p\rangle^{2d}\big[\widehat{B}(p)[1 - \varphi(\lambda^\delta\langle p\rangle)\big]\right| \mathrm{d}p \leq C_{\delta,M}\lambda^M\langle x\rangle^{-2d} \tag{3.22}$$

by using that B is in Schwarz space (2.6).

Let

$$H^\delta := -\frac{1}{2}\Delta_x + \lambda\int_{\mathbb{R}^d} B^\delta(x - y)\mathrm{d}\mu_\omega(y)$$

be the Hamiltonian with the truncated potential. Let $\psi_t^\delta := e^{-itH^\delta}\psi_0$ be the evolution of the wave function under the modified Hamiltonian H^δ. On the set Ω_Z and for $t \ll \lambda^{-4}$

$$\partial_t\|\psi_t - \psi_t^\delta\|^2 = -2\,\mathrm{Im}\,(\psi_t^\delta, (H - H^\delta)\psi_t) \leq C_\delta Z\lambda^{13}(\psi_t, \langle x^2\rangle\psi_t)^{1/2} \leq C(Z, \delta, \psi_0)\lambda^5$$

by using (3.22) with $M = 12$. In particular, ψ_t and ψ_t^δ remain close up to time scale $t \sim \lambda^{-2-\kappa}$, $\kappa < 2$. This bound, together with the L^2-continuity of the Wigner transform (3.16) guarantees that the truncation of B does not influence the left hand side of (2.22).

As for the right hand side of (2.22), notice that the collision kernel, $\sigma(U, V)$, of the momentum jump process (2.18) is restricted to the energy surface $e(V) = e(U) = e$. Therefore U, V are bounded, depending on e, so $\widehat{B}(U - V) = \widehat{B}^\delta(U - V)$ for these momenta, if λ is sufficiently small. Thus the truncation of B does not influence the right hand side of (2.22).

Armed with these results, we assume for the rest of the paper that $\widehat{\psi}_0(p)$ is smooth, compactly supported, bounded and $\widehat{B}(p)$ is supported on $\{|p| \leq \lambda^{-\delta}\}$ for any fixed $\delta > 0$. We thus extend the convention from the end of Section 2 that general constants denoted by C may depend on the truncated version of $\widehat{B}$ and $\widehat{\psi}_0$. The same applies to the hidden constants in the $O(\cdot)$ and $o(\cdot)$ notations.

3.3 Restriction to a finite box

We will reduce the problem to a finite box of size L, $L \gg 1$, with periodic boundary conditions. In this way, for technical convenience, we avoid the infinite summation in (2.3). Let $\Lambda_L := [-L/2, L/2]^d \subset \mathbb{R}^d$ be a finite torus and let $\Lambda_L^* := (\mathbb{Z}/L)^d$ be the dual lattice. We introduce the notation

$$\int_{\Lambda_L^*} f(x)\mathrm{d}p := \frac{1}{|\Lambda_L^*|} \sum_{p\in\Lambda_L^*} f(p) \; . \tag{3.23}$$

The integrals $\int_{\Lambda_L}$ and $\int_{\Lambda_L^*}$ converge to their infinite volume counterparts as $L \to \infty$. Let $(\cdot,\cdot)_L$ and $\|\cdot\|_L$ denote the scalar product and the norm on $L^2(\Lambda_L)$.

For any $L, M \gg 1$ we consider the random Schrödinger operator

$$H' = H'_{L,M} := -\frac{1}{2}\Delta + \lambda V'_\omega \qquad V'_\omega(x) := \sum_{\gamma=1}^{M} v'_\gamma \, B(x - y'_\gamma) = \int_{\Lambda_L} B(x-y)\mathrm{d}\mu'_\omega \; ,$$

with periodic boundary conditions on Λ_L and $\mu'_\omega := \sum_{\gamma=1}^M v'_\gamma \delta_{y'_\gamma}$. Here $\{y'_\gamma \; : \; \gamma = 1, \dots, M\}$ are i.i.d. random variables uniformly distributed on Λ_L and $\{v'_\gamma \; : \; \gamma = 1, \dots, M\}$ are i.i.d. variables distributed according to $\mathbf{P}_v$ and they are independent of the y'_γ. M itself will be random; it is chosen to be an independent Poisson variable with expectation $|\Lambda_L|$. The expectation with respect to the joint measure of $\{M, y'_\gamma, v'_\gamma\}$ is denoted by $\mathbf{E}'$. Sometimes we will use the decomposition

$$\mathbf{E}' = \mathbf{E}_M \mathbf{E}_y^{\otimes M} \mathbf{E}_v^{\otimes M} \tag{3.24}$$

referring to the expectation of M, $\{y_\gamma\}$ and $\{v_\gamma\}$ separately. The parameter L is implicit in these notations. In particular, $\mathbf{E}_y^{\otimes M}$ stands for the normalized integral

$$\frac{1}{|\Lambda_L|^M} \int_{\Lambda_L} \prod_{\gamma=1}^{M} \mathrm{d}y_\gamma \; . \tag{3.25}$$

It is well known that the restriction of the random measure μ_ω (see (2.3)) to the box Λ_L has the same distribution as μ'_ω. In particular, given a realization ω of the infinite volume random measure μ_ω, we can associate to it the number of points in Λ_L ($M = M(\omega)$) and the operator $H'_\omega = H_{L,M(\omega)}$ with random measure μ'_ω. We can thus realize the random operator H'_ω on the same probability space as H_ω. Due to the periodic boundary and the nontrivial support of B, the potential of H_ω and H'_ω will not be the same on Λ_L, but the difference will be negligible far away from the boundary.

Let χ_L be a smooth cutoff function, supported on Λ_L, with $\chi_L \equiv 1$ on $\Lambda_{L/2}$ and $|\nabla\chi_L| \le CL^{-1}$. Let $\psi_L(t) := \chi_L e^{-itH}\psi'_0$ and let $\psi'(t) := e^{-itH'}\psi'_0$ be the two dynamics applied to the cutoff initial data $\psi'_0 := \chi_L\psi_0$ supported on Λ_L. We also define the cutoff observable $\mathcal{O}_L := \chi_L\mathcal{O}$. Clearly

$$\lim_{L\to\infty} \mathbf{E}\langle \widehat{\mathcal{O}}_L, \widehat{W}^\varepsilon_{\psi_L(t)}\rangle_L = \mathbf{E}\langle \widehat{\mathcal{O}}, \widehat{W}^\varepsilon_{\psi(t)}\rangle \tag{3.26}$$

for any t. We estimate

$$\partial_t \|\psi_L(t) - \psi'_L(t)\|_L^2 \le C\|(H-H')\psi_L(t)\|_L^2 + C\|[H,\chi_L]\psi_L(t)\|_L^2 \; . \tag{3.27}$$

The second term is bounded by $CL^{-1}\|\nabla\psi_L(t)\|$ and on Ω_Z it can be estimated by the total energy as in (3.20). With a propagation estimate similar to (3.21) but applied to the evolution $e^{-itH}\psi_0'$, we easily obtain that the right hand side of (3.27) vanishes as $L\to\infty$ for any t. On the complement set, $\omega\in\Omega_Z^c$, we use the uniform bound (3.19) and finally let $Z\to\infty$. In summary, we have shown the following

Lemma 3.4 *Let $\psi'(t) := e^{-itH'_{L,M}}\psi_0'$, where M is a Poisson random variable with mean $|\Lambda_L|$, then*

$$\limsup_{L\to\infty}\Big|\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi(t)}\rangle-\mathbf{E}'\langle\widehat{\mathcal{O}}_L,\widehat{W}^{\varepsilon}_{\psi'(t)}\rangle_L\Big|=0\qquad\square$$

whenever $\int\sup_v|\widehat{\mathcal{O}}(\xi,v)|\mathrm{d}\xi<\infty$. $\square$

4 The Duhamel expansion

We expand the unitary kernel of $H=H_0+\widetilde{V}$ (see (3.10)) by the Duhamel formula. Due to the restriction to Λ_L, we really work with $H'_{L,M}=H_0'+\widetilde{V}'$, where the renormalized free evolution, H_0', is given by $\omega(p)$ in Fourier space and $\widetilde{V}'=\lambda V'-\lambda^2\theta(p)$, $p\in\Lambda_L^*$. The prime indicates the restriction to Λ_L and the dependence on L and M. In this section we work on Λ_L but we will mostly omit the primes in the notation.

For any fixed integer $N\geq 1$

$$\psi_t := e^{-itH}\psi_0=\sum_{n=0}^{N-1}\psi_n(t)+\Psi_N(t)\;, \tag{4.1}$$

with

$$\psi_n(t):=(-i)^n\int_0^t[\mathrm{d}s_j]_1^{n+1}\;\; e^{-is_{n+1}H_0}\widetilde{V}e^{-is_nH_0}\widetilde{V}\dots\widetilde{V}e^{-is_1H_0}\psi_0 \tag{4.2}$$

being the fully expanded terms and

$$\Psi_N(t):=(-i)\int_0^t\mathrm{d}s\,e^{-i(t-s)H}\widetilde{V}\psi_{N-1}(s) \tag{4.3}$$

is the non-fully expanded or error term. We used the shorthand notation

$$\int_0^t[\mathrm{d}s_j]_1^n:=\int_0^t\dots\int_0^t\Big(\prod_{j=1}^n\mathrm{d}s_j\Big)\delta\Big(t-\sum_{j=1}^n s_j\Big)\;.$$

Since each potential $\widetilde{V}$ in (4.2), (4.3) is a summation itself, $\widetilde{V}=-\lambda^2\theta(p)+\sum_{\gamma=1}^M V_\gamma$, $V_\gamma(x):=v_\gamma B(x-y_\gamma)$, both of these terms in (4.2) and (4.3) are actually big summations over so-called elementary wave functions, which are characterized by their collision history, i.e. by a sequence of obstacles and, occasionally, by $\theta(p)$. Denote by $\tilde{\Gamma}_n$, $n\leq\infty$, the set of sequences

$$\tilde{\gamma}=(\tilde{\gamma}_1,\tilde{\gamma}_2,\dots,\tilde{\gamma}_n),\qquad\tilde{\gamma}_j\in\{1,2,\dots,M\}\cup\{\vartheta\} \tag{4.4}$$

and by $W_{\tilde\gamma}$ the associated potential

$$W_{\tilde\gamma} := \begin{cases} \lambda V_{\tilde\gamma} & \text{if} \quad \tilde\gamma \in \{1,\dots,M\} \\ -\lambda^2\theta(p) & \text{if} \quad \tilde\gamma = \vartheta\ . \end{cases}$$

The tilde refers to the fact that the additional $\{\vartheta\}$ symbol is also allowed. An element $\tilde\gamma \in \{1,\dots,M\}\cup\{\vartheta\}$ is identified with the potential $W_{\tilde\gamma}$ and it is called *potential label* if $\tilde\gamma \in \{1,\dots,M\}$, otherwise it is a *$\vartheta$-label.* A potential label carries a factor λ, a ϑ-label carries λ^2.

For any $\tilde\gamma \in \tilde\Gamma_n$ we define the following fully expanded wave function with truncation

$$\psi_{*t,\tilde\gamma} := (-i)^{n-1}\int_0^t [\mathrm{d}s_j]_1^n\ \ W_{\tilde\gamma_n}e^{-is_nH_0}W_{\tilde\gamma_{n-1}}\dots e^{-is_2H_0}W_{\tilde\gamma_1}e^{-is_1H_0}\psi_0 \tag{4.5}$$

and without truncation

$$\psi_{t,\tilde\gamma} := (-i)^{n}\int_0^t [\mathrm{d}s_j]_1^{n+1}\ \ e^{-is_{n+1}H_0}W_{\tilde\gamma_n}e^{-is_nH_0}W_{\tilde\gamma_{n-1}}\dots e^{-is_2H_0}W_{\tilde\gamma_1}e^{-is_1H_0}\psi_0\ . \tag{4.6}$$

In the notation the star $(*)$ will always refer to truncated functions. Note that

$$\psi_{t,\tilde\gamma} = (-i)\int_0^t \mathrm{d}s\ e^{-i(t-s)H_0}\psi_{*s,\tilde\gamma}\ .$$

Each term (4.6) along the expansion procedure is characterized by its order n and by a sequence $\tilde\gamma \in \tilde\Gamma_n$. We now identify the main terms.

Denote by $\Gamma_k^{nr} \subset \tilde\Gamma_k$ the set of *non-repetitive* sequences that contain only potential labels, i.e.

$$\Gamma_k^{nr} := \Big\{\gamma = (\gamma_1,\dots,\gamma_k)\ :\ \gamma_j \in \{1,\dots,M\},\ \gamma_i \neq \gamma_j \text{ if } i \neq j\Big\}\ .$$

Let

$$\psi_{t,k}^{nr} := \sum_{\gamma\in\Gamma_k^{nr}} \psi_{t,\gamma}$$

denote the corresponding elementary wave functions.

The typical number of collisions up to time t is of order $\lambda^2 t$. To allow us for some room, we set

$$K := [\lambda^{-\delta}(\lambda^2 t)]\ , \tag{4.7}$$

($[\ \cdot\ ]$ denotes integer part), where $\delta = \delta(\kappa) > 0$ is a small positive number to be fixed later on. K will serve as an upper threshold for the number of collisions in the expansion.

5 Proof of the Main Theorem

The proof is divided into three theorems. The first one states that all terms other than $\psi_{t,k}^{nr}$, $0 \le k < K$, are negligible. For the precise statement we use the previous notations, in particular we recall that the prime indicates the dependence on L, M

Theorem 5.1 (L^2-estimate of the error terms) *Let $t = O(\lambda^{-2-\kappa})$ and K given by* (4.7)*. If $\kappa < \kappa_0(d)$ and δ is sufficiently small (depending only on κ), then*

$$\lim_{\lambda\to 0}\lim_{L\to\infty} \mathbf{E}'\Big\|\psi'_t - \sum_{k=0}^{K-1} \psi'^{\,nr}_{t,k}\Big\|_L^2 = 0\ .$$

In $d = 3$ dimensions, one can choose $\kappa_0(3) = \frac{1}{500}$.

The second key theorem gives an explicit formula for the main terms, $\psi'^{\,nr}_{t,k}$. It really identifies the so-called ladder diagram as the only contributing term. We introduce the notation

$$R_\eta(\alpha, v) := \frac{1}{\alpha - \omega(v) + i\eta}\ ,$$

for the renormalized propagator.

Theorem 5.2 (Only the ladder diagram contributes) *Let $\kappa < \frac{2}{34d+39}$, $\varepsilon = \lambda^{2+\kappa/2}$, $t = O(\lambda^{-2-\kappa})$, and K given by* (4.7)*. For a sufficiently small positive δ, for $\eta \geq \lambda^{2+2\kappa}$ and for any $1 \leq k < K$ we have*

$$\lim_{L\to\infty} \mathbf{E}'\|\psi'^{\,nr}_{t,k}\|_L^2 = V_\lambda(t,k) + O\Big(\lambda^{\frac{1}{3}-(\frac{17}{3}d+\frac{13}{2})\kappa-O(\delta)}\Big) \tag{5.1}$$

$$\lim_{L\to\infty} \langle \widehat{\mathcal{O}}_L, \mathbf{E}'\widehat{W}^{\varepsilon}_{\psi'^{\,nr}_{t,k}}\rangle_L = W_\lambda(t,k,\mathcal{O}) + O\Big(\lambda^{\frac{1}{3}-(\frac{17}{3}d+\frac{13}{2})\kappa-O(\delta)}\Big) \tag{5.2}$$

as $\lambda \ll 1$. Here

$$\begin{aligned} V_\lambda(t,k) :=& \frac{\lambda^{2k}e^{2t\eta}}{(2\pi)^2}\iint_{-\infty}^{\infty} \mathrm{d}\alpha\mathrm{d}\beta\ e^{i(\alpha-\beta)t}\int \Big(\prod_{j=1}^{k+1}\mathrm{d}p_j\Big)\ |\widehat{\psi}_0(p_1)|^2 \\ &\times \prod_{j=1}^{k+1}\overline{R_\eta(\alpha,p_j)}R_\eta(\beta,p_j)\prod_{j=1}^{k}|\widehat{B}(p_{j+1}-p_j)|^2 \end{aligned} \tag{5.3}$$

$$\begin{aligned} W_\lambda(t,k,\mathcal{O}) :=& \frac{\lambda^{2k}e^{2t\eta}}{(2\pi)^2}\iint_{-\infty}^{\infty} \mathrm{d}\alpha\mathrm{d}\beta\ e^{i(\alpha-\beta)t}\int \mathrm{d}\xi\int \Big(\prod_{j=1}^{k+1}\mathrm{d}v_j\ \Big)\widehat{\mathcal{O}}(\xi,v_{k+1})\overline{\widehat{W}^{\varepsilon}_{\psi_0}}(\xi,v_1) \\ &\times \prod_{j=1}^{k+1}\overline{R_\eta\Big(\alpha, v_j+\frac{\varepsilon\xi}{2}\Big)}R_\eta\Big(\beta, v_j-\frac{\varepsilon\xi}{2}\Big)\prod_{j=1}^{k}|\widehat{B}(v_j-v_{j+1})|^2\ . \end{aligned} \tag{5.4}$$

We adopt the notation $O(\delta)$ in the exponent of λ. This always means (const.)δ with universal, explicitly computable positive constants that depend on κ and that can be easily computed from the proof.

The formula (5.3) is the value of the so-called *ladder Feynman graph* in the diagrammatic expansion of $\mathbf{E}'\|\psi'^{\,nr}_{t,k}\|^2$. We will see in Proposition 7.2 that this expansion generates $k!B_k$ terms, where B_k is the number of partitions of a set with k elements (note that B_k is almost of order $k!$). Theorem

5.2 states that only one diagram is relevant; the contribution of all the other Feynman graphs is negligible even after summation. The extension of (5.1) to the Wigner transform (5.2) is straightforward. Theorem 5.2 is the most important step in the proof of the Main Theorem.

The third theorem identifies the limit of $\sum_k W_\lambda(t,k,\mathcal{O})$ as $\lambda \to 0$ with the solution to the heat equation. We note that the definition (5.4) does not apply literally to the free evolution term $k=0$; this term is defined separately:

$$W_\lambda(t,k=0,\mathcal{O}) := \int \mathrm{d}\xi \mathrm{d}v \; e^{it\varepsilon v\cdot\xi} \; e^{2t\lambda^2 \mathrm{Im}\,\theta(v)} \; \widehat{\mathcal{O}}(\xi,v)\overline{\widehat{W}_0}(\varepsilon\xi,v) \; . \tag{5.5}$$

Theorem 5.3 (The ladder diagram converges to the heat equation) *Under the conditions of Theorem 5.2 and setting $t=\lambda^{-2-\kappa}T$, we have*

$$\lim_{\lambda\to 0} \sum_{k=0}^{K-1} W_\lambda(t,k,\mathcal{O}) = \int \mathrm{d}X \int \mathrm{d}v \; \mathcal{O}(X,v) f(T,X,e(v)) \; , \tag{5.6}$$

where f is the solution to the heat equation (2.21).

Proof of the Main Theorem 2.2 using Theorems 5.1, 5.2 and 5.3. We compute the expectation of the rescaled Wigner transform, $\mathbf{E}W_t^\varepsilon = \mathbf{E}W_{\psi_t}^\varepsilon$, tested against a Schwarz function

$$\int \mathrm{d}X \int \mathrm{d}v \; \mathcal{O}(X,v)\mathbf{E}W_t^\varepsilon(X,v) = \int \mathrm{d}\xi \int \mathrm{d}v \; \widehat{\mathcal{O}}(\xi,v)\mathbf{E}\overline{\widehat{W_t^\varepsilon}}(\xi,v) = \langle \, \mathcal{O}, \mathbf{E}W_t^\varepsilon \rangle \; .$$

Combining Lemma 3.4, Theorem 5.1 and the finite box version of the L^2-continuity of the Wigner transform (3.16), it is sufficient to compute the Wigner transform of $\psi'(t,K) := \sum_{k=0}^{K-1} \psi'^{\,nr}_{t,k}$. The Wigner transform $W_{\psi'(t,K)}$ is quadratic in ψ', so it contains a double sum over k and k'

$$W_{\psi'(t,K)} = \sum_{k,k'=0}^{K-1} \overline{\psi'^{\,nr}_{t,k}}(\cdots)\psi'^{\,nr}_{t,k'}(\cdots) \; .$$

The potential labels are not repeated within $\overline{\psi}$ and ψ. Moreover, the expectation of a single potential in (4.6) is zero. Thus the potential labels in the ψ and $\overline{\psi}$ must pair, in particular taking expectation reduces this double sum to a single sum over k

$$\mathbf{E}' \, W_{\psi'(t,K)} = \sum_{k=0}^{K-1} \mathbf{E}' \, W_{\psi'^{\,nr}_{t,k}} \; .$$

By using (5.2) and (5.6) together with $K = O(\lambda^{-\kappa-\delta})$, we obtain Theorem 2.2. $\square$

The main result of the present paper is the proof of Theorem 5.2. The proofs of Theorem 5.1 and Theorem 5.3 will be given in the companion paper [19]. For the reader's convenience, we summarize below the key ideas of the proof of Theorem 5.1 from [19].

The Duhamel expansion allows for the flexibility that at every new term of the expansion we perform the separation into elementary waves, $\psi_{*s,\widetilde{\gamma}}$, and we can decide whether we want to stop (keeping the full propagator as in (4.3)) or we continue to expand that term further. This decision will depend on the collision history, $\widetilde{\gamma}$. In particular, not every error term will be expanded up to the same order N, in some cases we may decide to stop the expansion earlier.

To estimate a non-fully expanded term, we will use the unitarity of the full evolution,

$$\left\|(-i)\int_0^t e^{-i(t-s)H}\psi_{*s,\widetilde{\gamma}}\mathrm{d}s\right\|^2 \leq t\int_0^t \|\psi_{*s,\widetilde{\gamma}}\|^2\mathrm{d}s\ . \tag{5.7}$$

Typically we lose a factor of t by using this estimate since the oscillatory character of the time integration is lost. We can use this crude estimate only if the fully expanded term, $\|\psi_{*s,\widetilde{\gamma}}\|^2$, is small, i.e. if $\widetilde{\gamma}$ represents an atypical collision sequence. Once $\widetilde{\gamma}$ is "sufficiently" atypical, we stop the expansion for that elementary wave function to reduce the number and the complexity of the expanded terms.

There are basically two patterns how a collision history can become atypical; either the total number of collisions exceeds the typical number of collisions, $O(\lambda^2 t)$, or there is a recollision. This explains why only the non-repetition terms $\psi_{t,k}^{nr}$ with $k \leq K$ contribute to the main term.

A recollision is typically penalized by a factor λ^2 in the weak coupling environment. This is, however, not the case for the immediate repetition of a potential label, $\widetilde{\gamma}_j = \widetilde{\gamma}_{j+1} \in \{1,\ldots,M\}$. The renormalization (3.10) compensates for these terms. Up to the highest order, the contribution of a sequence with an immediate repetition cancels that of the same sequence where the repetition is replaced by a θ-label. Technically, all these estimates have to be combined with the key method of the present paper (proof of Theorem 5.2) to show that the sum of all $k!B_k$ repetition diagrams is sufficiently small to compensate for the unitarity estimate (5.7).

6 Pairing potential labels

The wave function

$$\psi_{t,k}'^{\,nr} = (-i)^k \sum_{\gamma\in\Gamma_k^{nr}} \int_0^t [\mathrm{d}s_j]_1^{k+1}\ \ e^{-is_{k+1}H_0'}V_{\gamma_k}' e^{-is_kH_0'}V_{\gamma_{k-1}}'\ldots e^{-is_2H_0'}V_{\gamma_1}' e^{-is_1H_0'}\psi_0'$$

contains k potential terms with different potential labels. Every term in

$$\mathbf{E}'\|\psi_{t,k}'^{\,nr}\|_L^2 = \sum_{\gamma,\gamma'} \mathbf{E}'\ \overline{\psi_{t,\gamma}}\psi_{t,\gamma'}$$

has $2k$ potential terms, and their expectation is

$$\mathbf{E}'\ \overline{V_{\gamma_1}'V_{\gamma_2}'\ldots V_{\gamma_k}'}V_{\gamma_1'}'V_{\gamma_2'}'\ldots V_{\gamma_k'}'\ . \tag{6.1}$$

Since there is no repetition within γ and γ', and $\mathbf{E}'V_\gamma' = 0$, the expectation in (6.1) is nonzero only if there is a complete pairing between γ and γ'. Such pairings correspond to permutations on $I_k = \{1,2,\ldots,k\}$. We denote by $\mathfrak{S}_k$ the set of all permutations on k elements.

We recall the K-identity from Lemma 3.1 of [16] (with a corrected $(2\pi)^{-1}$ factor)

$$\int_0^t [\mathrm{d}s_j]_1^{k+1} \prod_{j=1}^{k+1} e^{-is_j\omega(p_j)} = \frac{ie^{\eta t}}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \; e^{-i\alpha t} \prod_{j=1}^{k+1} \frac{1}{\alpha - \omega(p_j) + i\eta} \tag{6.2}$$

for any $\eta > 0$. Therefore, we have

$$\begin{aligned} \mathbf{E}' \|\psi_{t,k}'^{nr}\|_L^2 = \frac{\lambda^{2k} e^{2t\eta}}{(2\pi)^2} \sum_{\sigma \in \mathfrak{S}_k} \iint_{\Lambda_L^*} \mathrm{d}\mathbf{p} \mathrm{d}\tilde{\mathbf{p}} \; \delta(p_{k+1} - \tilde{p}_{k+1}) \\ \times \mathbf{E}' \sum_{\substack{\gamma_1,\dots,\gamma_k = 1 \\ \gamma_i \neq \gamma_j}}^{M} \prod_{j=1}^{k} \overline{\widehat{V}_{\gamma_j}(p_{j+1} - p_j)} \widehat{V}_{\gamma_j}(\tilde{p}_{\sigma(j)+1} - \tilde{p}_{\sigma(j)}) \overline{\widehat{\psi}_0'(p_1)} \widehat{\psi}_0'(\tilde{p}_1) \\ \times \iint_{\mathbb{R}} \mathrm{d}\alpha \mathrm{d}\beta \; e^{i(\alpha-\beta)t} \left(\prod_{j=1}^{k+1} \frac{1}{\alpha - \overline{\omega}(p_j) - i\eta} \; \frac{1}{\beta - \omega(\tilde{p}_j) + i\eta} \right), \end{aligned} \tag{6.3}$$

where the summation runs over all ordered k-tuples $(\gamma_1, \dots, \gamma_k)$ of $\{1, 2, \dots, M\}$ with disjoint elements. We compute the expectation, using $m_2 = 1$ from (2.4) and the factorization of $\mathbf{E}'$ from (3.24)

$$\mathbf{E} \prod_{j=1}^{k} \overline{\widehat{V}_{\gamma_j}(p_{j+1} - p_j)} \widehat{V}_{\gamma_j}(\tilde{p}_{\sigma(j)+1} - \tilde{p}_{\sigma(j)}) = P(\sigma, \mathbf{p}, \tilde{\mathbf{p}}) \overline{\mathcal{B}(\mathbf{p})} \mathcal{B}(\tilde{\mathbf{p}}) \tag{6.4}$$

with

$$\mathcal{B}(\mathbf{p}) := \prod_{j=1}^{k} \widehat{B}(p_{j+1} - p_j) \tag{6.5}$$

and

$$P(\sigma, \mathbf{p}, \tilde{\mathbf{p}}) := \mathbf{E}_M \mathbf{E}_y^{\otimes M} \sum_{\substack{\gamma_1,\dots,\gamma_k = 1 \\ \gamma_i \neq \gamma_j}}^{M} \prod_{j=1}^{k} \exp\left[2\pi i y_{\gamma_j} (p_{j+1} - p_j - (\tilde{p}_{\sigma(j)+1} - \tilde{p}_{\sigma(j)})) \right]. \tag{6.6}$$

We obtain from (6.3) that

$$\begin{aligned} \mathbf{E}' \|\psi_{t,k}'^{nr}\|_L^2 = \lambda^{2k} \sum_{\sigma \in \mathfrak{S}_k} \sum_{\substack{\gamma_1,\dots,\gamma_k \\ \gamma_i \neq \gamma_j}} \iint_{\Lambda_L^*} \mathrm{d}\mathbf{p} \mathrm{d}\tilde{\mathbf{p}} \; \delta(p_{k+1} - \tilde{p}_{k+1}) \\ \times P(\sigma, \mathbf{p}, \tilde{\mathbf{p}}) M^\circ(k, \mathbf{p}, \tilde{\mathbf{p}}) \overline{\widehat{\psi}_0(p_1)} \widehat{\psi}_0(\tilde{p}_1) \end{aligned} \tag{6.7}$$

with

$$M^\circ(k, \mathbf{p}, \tilde{\mathbf{p}}) := \frac{e^{2t\eta}}{(2\pi)^2} \iint_{\mathbb{R}} \mathrm{d}\alpha \mathrm{d}\beta \; e^{i(\alpha-\beta)t} \left(\prod_{j=1}^{k+1} \frac{\overline{\widehat{B}}(p_{j+1} - p_j)}{\alpha - \overline{\omega}(p_j) - i\eta} \; \frac{\widehat{B}(\tilde{p}_{j+1} - \tilde{p}_j)}{\beta - \omega(\tilde{p}_j) + i\eta} \right) \tag{6.8}$$

(with the convention that for $j = k+1$ we set the superfluous term $\widehat{B}(p_{j+1} - p_j) := 1$).

Due to the constraint $\gamma_i \neq \gamma_j$, the formula (6.6) is not a simple product of delta functions and we have to use a connected graph expansion that is well-known from field theory.

Let $\mathcal{A}_n$ be the set of partitions of $I_k := \{1, 2, \dots, k\}$, i.e. $\mathbf{A} = \{A_\mu \ : \ \mu \in I(\mathbf{A})\} \in \mathcal{A}_k$ if $\cup_{\mu \in I(\mathbf{A})} A_\mu = I_k$ and the elements of $\mathbf{A}$ are disjoint. The sets in the partition are labelled by the index set $I(\mathbf{A})$ and let $m(\mathbf{A}) = |I(\mathbf{A})|$ denote the number of elements in $\mathbf{A}$. The elements of the partition $\mathbf{A}$ will be called *lumps*. A lump is *trivial* if it has only one element. The trivial partition, where every lump is trivial, is denoted by $\mathbf{A}_0$.

Lemma 6.1 *For any fixed L, k and M, $k \leq M$, and any fixed momenta $q_j \in \Lambda_L^*$,*

$$\mathbf{E}_y^{\otimes M} \sum_{\substack{\gamma_1, \dots, \gamma_k = 1 \\ \gamma_i \neq \gamma_j}}^{M} \prod_{j=1}^{k} \exp\Big[2\pi i q_j y_{\gamma_j}\Big] = \Big(\frac{M}{|\Lambda_L|}\Big)^k \sum_{\mathbf{A} \in \mathcal{A}_k} \prod_{\nu \in I(\mathbf{A})} c(|A_\nu|) \delta\Big(\sum_{\ell \in A_\nu} q_\ell\Big) \tag{6.9}$$

with

$$c(n) := \sum_{\substack{\Gamma \subset K_n \\ \Gamma \ connected}} (-1)^{|\Gamma|}$$

where K_n denotes the complete graph on n vertices and $|\Gamma|$ denotes the number of edges in the subgraph Γ. The following estimate holds for $n \geq 2$

$$|c(n)| \leq n^{n-2} \ . \tag{6.10}$$

Remark. An analogous formula holds if the natural index set $I_k = \{1, 2, \dots, k\}$ is replaced by an arbitrary finite set S. In this case, the summation on the right hand side is over all partitions of S. The set of these partitions is denoted by $\mathcal{A}(S)$.

Proof. By the connected graph expansion

$$\prod_{i \neq j = 1}^{k} (1 - \delta_{\gamma_i, \gamma_j}) = \sum_{\mathbf{A} \in \mathcal{A}_k} \prod_{\nu \in I(\mathbf{A})} \delta_c(A_\nu) \ ,$$

where

$$\delta_c(A) = c(|A|) \prod_{\ell, \ell' \in A} \delta_{\gamma_\ell, \gamma_{\ell'}}$$

is the Ursell coefficients of the hard-core lattice gas (see: e.g. [43]). Therefore

$$\begin{aligned} \text{L.h.s of (6.9)} \ &= \ \mathbf{E}_y^{\otimes M} \sum_{\gamma_1, \dots, \gamma_k = 1}^{M} \sum_{\mathbf{A} \in \mathcal{A}_k} \prod_{\nu \in I(\mathbf{A})} \left(e^{2\pi i \sum_{\ell \in A_\nu} q_\ell y_{\gamma_\ell}} c(|A_\nu|) \prod_{\ell, \ell' \in A_\nu} \delta_{\gamma_\ell, \gamma_{\ell'}} \right) \\ &= \ \Big(\frac{M}{|\Lambda_L|}\Big)^k \sum_{\mathbf{A} \in \mathcal{A}_k} \prod_{\nu \in I(\mathbf{A})} \Big[c(|A_\nu|) \delta(\sum_{\ell \in A_\nu} q_\ell) \Big] \ . \qquad \square \end{aligned}$$

We will use the identity (6.9) to express P in (6.6) as a linear combination of products of delta functions of the momenta and insert it into (6.7). After the limit $L \to \infty$, each term in the summation $\sum_\sigma \sum_{\mathbf{A}}$ will be expressed by a Feynman graph. The precise definitions will be given in the next Section.

7 Graphical representation

Traditionally, the Feynman graphs consist of interaction vertices and particle lines among them. In case of Gaussian random potentials, the interaction vertices are paired according to the Wick theorem [16]. For non-Gaussian randomness, the non-vanishing higher order cumulants correspond to joining several vertices [10]. In our case, the appearance of the non-trivial subsets are due to selecting the non-repetition sequences. This requires us to define Feynman graphs in a more general setup than usual. In this section we introduce the necessary graphical representation in full generality and we will define the value of a Feynman graph, $V^\circ(\mathbf{A}, \sigma)$, with permutation σ and partition $\mathbf{A}$ in (7.21). The final result of this section is given in Proposition 7.2 at the end.

7.1 Circle graphs and their values

We start with an oriented circle graph with two distinguished vertices, denoted by $0, 0^*$. The number of vertices is N. The vertex set is $\mathcal{V}$, the set of oriented edges is $\mathcal{L}(\mathcal{V})$. For $v \in \mathcal{V}$ we use the notation $v-1$ and $v+1$ for the vertex right before and after v in the circular ordering. We also denote $e_{v-} = (v-1, v)$ and $e_{v+} = (v, v+1)$ the edge right before and after the vertex v, respectively. In particular $e_{(v+1)-} = e_{v+}$. For each $e \in \mathcal{L}(\mathcal{V})$, we introduce a momentum w_e and a real number α_e associated to this edge. The collection of all momenta is denoted by $\mathbf{w} = \{w_e \,:\, e \in \mathcal{L}(\mathcal{V})\}$ and $\mathrm{d}\mathbf{w} = \otimes_e \mathrm{d}w_e$ is the Lebesgue measure. We sometimes use the notation $v \sim e$ to indicate that an edge e is adjacent to a vertex v.

Let $\mathbf{P} = \{P_\mu \,:\, \mu \in I\}$ be a partition of the set $\mathcal{V} \setminus \{0, 0^*\}$

$$\mathcal{V} \setminus \{0, 0^*\} = \bigcup_{\mu \in I} P_\mu \,,$$

(all P_μ nonempty and pairwise disjoint) where $I = I(\mathbf{P})$ is the index set to label the sets in the partition. Let $m(\mathbf{P}) := |I(\mathbf{P})|$. The sets P_μ are called $\mathbf{P}$-*lumps* or just *lumps*. If two elements $v, v' \in \mathcal{V} \setminus \{0, 0^*\}$ belong to the same lump within a partition $\mathbf{P}$, we denote it by $v \equiv v' \ (mod\ \mathbf{P})$. We assign a variable, $u_\mu \in \mathbb{R}^d$, $\mu \in I(\mathbf{P})$, to each lump. We call them *auxiliary momenta*; they will be needed for a technical reason. We always assume that the auxiliary momenta add up to 0

$$\sum_{\mu \in I(\mathbf{P})} u_\mu = 0 \,. \tag{7.1}$$

The vector of auxiliary momenta is denoted by $\mathbf{u} := \{u_\mu \,:\, \mu \in I(\mathbf{P})\}$.

The set of all partitions of the vertex set $\mathcal{V} \setminus \{0, 0^*\}$ is denoted by $\mathcal{P}_\mathcal{V}$. For any $P \subset \mathcal{V}$, we let

$$L_+(P) := \{(v, v+1) \in \mathcal{L}(\mathcal{V}) \,:\, v+1 \notin P,\ v \in P\}$$

denote the set of edges that go out of P, with respect to the orientation of the circle graph, and similarly $L_-(P)$ denote the set of edges that go into P. We set $L(P) := L_+(P) \cup L_-(P)$.

For any $\xi \in \mathbb{R}^d$ we define the following product of delta functions

$$\Delta(\mathbf{P}, \mathbf{w}, \mathbf{u}) := \delta\Big(\xi + \sum_{e \in L_\pm(\{0^*\})} \pm w_e\Big) \prod_{\mu \in I(\mathbf{P})} \delta\Big(\sum_{e \in L_\pm(P_\mu)} \pm w_e - u_\mu\Big) \,, \tag{7.2}$$

where $\mathbf{u} := \{u_\mu \ : \ \mu \in I(\mathbf{P})\} \in \mathbb{R}^d$ is a set of auxiliary momenta. The sign $\pm$ indicates that momenta w_e is added or subtracted depending whether the edge e is outgoing or incoming, respectively. The function $\Delta(\cdots) = \Delta_\xi(\cdots)$ depends on ξ, but we will mostly omit this fact from the notation. All estimates will be uniform in ξ.

Summing up all arguments of these delta functions and using (7.1) we see that these delta functions force the two momenta corresponding to the two edges adjacent to 0 to differ by ξ: $w_e - w_{e'} = \xi$ for $e \in L_+(\{0\})$, $e' \in L_-(\{0\})$.

As a motivation for these definitions, we mention that the lumps naturally arise from the connected graph formula (Lemma 6.1). According to this formula, the Kirchhoff Law must be satisfied for all lumps, i.e. the incoming and outgoing momenta must sum up to zero. This fact would be described by the delta functions (7.2) with all $u_\mu = 0$. In certain recollision terms, however, the non-repetition condition leading to Lemma 6.1 is not fully satisfied and the Kirchhoff Law breaks down for a few lumps. The nontrivial auxiliary momenta will bookkeep this deviation from the Kirchhoff Law (see [19] for more details). Finally, the shift by ξ at the vertex 0^* in (7.2) will be used when computing the Wigner transform in Fourier representation (2.10).

For each subset $\mathcal{G} \subset \mathcal{V} \setminus \{0, 0^*\}$, we define

$$\mathcal{N}_\mathcal{G}(\mathbf{w}) := \prod_{e \sim 0} |\widehat{\psi}_0(w_e)| \prod_{v \in \mathcal{V} \setminus \{0,0^*\} \setminus \mathcal{G}} |\widehat{B}(w_{e_{v-}} - w_{e_{v+}})| \prod_{v \in \mathcal{G}} \langle w_{e_{v-}} - w_{e_{v+}} \rangle^{-2d} \ . \tag{7.3}$$

In our application, the subset $\mathcal{G}$ collects those vertices, where the original potential decay $|\widehat{B}(w_{e_{v-}} - w_{e_{v+}})|$ could not be explicitly kept along the estimates and this will happen only at a few places; the size of $\mathcal{G}$ will be at most 8. For the purpose of this paper, i.e. for the proof of Theorem 5.2, we will need only $\mathcal{G} = \emptyset$, but for the analysis of the repetition terms in [19] we need the more general definition.

Due to the support properties of $\widehat{B}$ and $\widehat{\psi}_0$, we will see that all intermediate momenta w_e satisfy $|w_e| \leq N\lambda^{-\delta}$. The maximal number of vertices in our graphs will be $N \leq 2K + 2 = O(\lambda^{-\kappa-\delta})$, therefore all intermediate momenta will be smaller than $\zeta := \lambda^{-\kappa-3\delta}$. This justifies to define the restricted Lebesgue measures

$$\mathrm{d}\mu(w) := \mathbf{1}(|w| \leq \zeta)\mathrm{d}w \ , \quad \zeta := \lambda^{-\kappa-3\delta}, \qquad \mathrm{d}\mu(\mathbf{w}) := \otimes_e \mathrm{d}\mu(w_e) \ . \tag{7.4}$$

Moreover, each auxiliary momenta, u_μ, will always be a sum (or difference) of different w_e momenta (see (7.2)), therefore each of them always satisfies $|u_\mu| \leq O(\lambda^{-2\kappa-4\delta})$. We will often take the supremum of all possible auxiliary momenta and $\sup_\mathbf{u}$ is always considered subject to this bound.

With these notations, we define, for any $\mathbf{P} \in \mathcal{P}_\mathcal{V}$ and $g = 0, 1, 2, \ldots$, the **E-value of the partition**

$$E_g(\mathbf{P}, \mathbf{u}, \boldsymbol{\alpha}) := \lambda^{N-2} \sup_{\mathcal{G} \,:\, |\mathcal{G}| \leq g} \int \mathrm{d}\mu(\mathbf{w}) \prod_{e \in \mathcal{L}(\mathcal{V})} \frac{1}{|\alpha_e - \omega(w_e) + i\eta|} \, \Delta(\mathbf{P}, \mathbf{w}, \mathbf{u}) \mathcal{N}_\mathcal{G}(\mathbf{w}) \ . \tag{7.5}$$

The prefactor λ^{N-2} is due to the fact that in the applications all but the two distinguished vertices, $\{0, 0^*\}$, will carry a factor λ. The E-value depends also on the parameters λ, η, but we will not

specify them in the notation. In the applications, the regularization η will be mostly chosen as $\eta = \lambda^{2+\kappa}$.

We will also need a slight modification of these definitions, indicated by a lower star in the notation:

$$E_{*g}(\mathbf{P}, \mathbf{u}, \boldsymbol{\alpha}) := \lambda^{N-2} \sup_{\mathcal{G}\,:\,|\mathcal{G}|\le g} \int \mathrm{d}\mu(\mathbf{w}) \prod_{\substack{e\in\mathcal{L}(\mathcal{V})\\ e\not\in L(\{0^*\})}} \frac{1}{|\alpha_e - \omega(w_e) + i\eta|}\, \Delta(\mathbf{P}, \mathbf{w}, \mathbf{u}) \mathcal{N}_{\mathcal{G}}(\mathbf{w})\ . \quad (7.6)$$

The only difference is that the denominators carrying the momenta associated to edges that are adjacent to 0^* are not present in E_{*g}. We call E_{*g} the **truncation** of E_g. We will see that Feynman diagrams arising from the perturbation expansion can naturally be estimated by quantities of the form (7.5) or (7.6).

7.2 Feynman graphs

We apply this general setup to the following situation that we will call **Feynman graph**. Every quantity in our perturbation expansion will be expressed by values of Feynman graphs that are defined below.

For experts we mention that our Feynman graphs differ from those that one typically obtains after averaging over a Gaussian disorder. In the latter, potential lines never appear as external lines but only as pairing lines, and one can identify vertices connected by pairing lines so that the graph becomes four–valent. In our case, the graph is still trivalent and has external potential lines, with a corresponding dependence on momentum variables $\mathbf{u}$. Also, averaging the disorder will not simply pair up lines but can also join more than two potential lines. which correspond to the higher moments.

Consider the cyclically ordered set $\mathcal{V}_{n,n'} := \{0, 1, 2, \ldots, n, 0^*, \tilde{n}', \widetilde{n'-1}, \ldots, \tilde{1}\}$ and view this as the vertex set of an oriented circle graph on $N = n + n' + 2$ vertices. We set $I_n := \{1, 2, \ldots n\}$ and $\widetilde{I}_{n'} := \{\widetilde{1}, \widetilde{2}, \ldots, \widetilde{n}'\}$ and the vertex set can be identified with $\mathcal{V}_{n,n'} = I_n \cup \widetilde{I}_{n'} \cup \{0, 0^*\}$.

The set of edges $\mathcal{L}(\mathcal{V}_{n,n'})$ is partitioned into $\mathcal{L}(\mathcal{V}_{n,n'}) = \mathcal{L} \cup \widetilde{\mathcal{L}}$ such that $\mathcal{L}$ contains the edges between $I_n \cup \{0, 0^*\}$ and $\widetilde{\mathcal{L}}$ contains the edges between $\widetilde{I}_{n'} \cup \{0, 0^*\}$.

Let $\mathcal{P}_{n,n'}$ be the set of all partitions $\mathbf{P}$ on the set $I_n \cup \widetilde{I}_{n'}$. The lumps of a partition containing only one vertex will be called *single lumps*. The vertices 0 and 0^* are not part of the partitions hence they will not be considered single lumps. Let $G = G(\mathbf{P})$ be the set of edges that go into a single lump and let $g(\mathbf{P}) := |G(\mathbf{P})|$ be its cardinality. In case of $n = n'$, we will use the shorter notation $\mathcal{V}_n = \mathcal{V}_{n,n}$, $\mathcal{P}_n = \mathcal{P}_{n,n}$ etc. The Feynman graphs arising from the non-repetition terms will always have $n = n'$ and no single lumps, $g(\mathbf{P}) = 0$, but the more general definition will be needed for the repetition terms in [19]. We remark that even in [19] we will always have

$$|n - n'| \le 2, \qquad |n - n'| \le g(\mathbf{P}) \le 4, \qquad n, n' \le K\ . \quad (7.7)$$

We also introduce a function Q that will represent the momentum dependence of the observable. In our estimates, we will always bound Q in supremum norm; no decay or smoothness will be

necessary. We will need extra conditions on the observable only to evaluate the ladder in the proof of Theorem 5.3 (see [19] for details). Since Q will always appear linearly in our formulae, we can assume, for convenience, that $\|Q\|_\infty \leq 1$. General Q can be accommodated by a multiplicative factor $\|Q\|_\infty$ in the final estimate but it will not be carried along the proofs.

We define the following function to collect all potential terms:

$$\mathcal{M}(\mathbf{w}) := \prod_{e\in\mathcal{L}\cap G}[-\overline{\theta(w_e)}]\prod_{e\in\widetilde{\mathcal{L}}\cap G}[-\theta(w_e)]\prod_{\substack{e\in\mathcal{L}\setminus G\\ e\not\sim 0^*}}\overline{\widehat{B}(w_e-w_{e+1})}\prod_{\substack{e\in\widetilde{\mathcal{L}}\setminus G\\ e\not\sim 0}}\widehat{B}(w_e-w_{e+1}) \tag{7.8}$$
$$\times\,\overline{\widehat{\psi}(w_{e_{0+}})}\widehat{\psi}(w_{e_{0-}})Q\Big[\frac{1}{2}(w_{e_{0^*-}}+w_{e_{0^*+}})\Big]$$

with $\mathbf{w} := \{w_e \,:\, e\in\mathcal{L}\cup\widetilde{\mathcal{L}}\}$ and recalling that for any $e\in\mathcal{L}(\mathcal{V})$, the edge $e+1$ denotes the edge succeeding e in the circular ordering.

The delta function $\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)$ ensures that the two momenta adjacent to each single lump coincide. This holds even for $\xi\neq 0$, recall that $0,0^*$ are not considered lumps. Therefore the distribution $\mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)$ is supported on the regime with $|w_e|\leq\zeta$ for all momenta w_e, thanks to the support properties of $\widehat{\psi}_0$, $\widehat{B}$ and to the control on the number of terms, $n,n'\leq K$ (see Section 7.1). In particular

$$\mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathrm{d}\mathbf{w} = \mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathrm{d}\mu(\mathbf{w})\;. \tag{7.9}$$

Using the boundedness of θ and $|\widehat{\psi}_0(w)|\leq C\langle w\rangle^{-10d}$, we easily obtain

$$|\mathcal{M}(\mathbf{w})|\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathrm{d}\mathbf{w}\leq C^{g(\mathbf{P})}\mathcal{N}_{\mathcal{G}}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathrm{d}\mu(\mathbf{w})\;, \tag{7.10}$$

where $\mathcal{G}$ is the set of single lumps and $g(\mathbf{P})=|\mathcal{G}|$, since the delta function also guarantees that there is no additional decay at the vertices $v\in\mathcal{G}$ in $\mathcal{N}_{\mathcal{G}}(\mathbf{w})$ (the last product in (7.3) is a constant).

Let $\alpha,\beta\in\mathbb{R}$, $\mathbf{P}\in\mathcal{P}_{n,n'}$ and

$$V(\mathbf{P},\alpha,\beta) := \lambda^{n+n'+g(\mathbf{P})}\int\mathrm{d}\mathbf{w}\prod_{e\in\mathcal{L}}\frac{1}{\alpha-\overline{\omega}(w_e)-i\eta}\prod_{e\in\widetilde{\mathcal{L}}}\frac{1}{\beta-\omega(w_e)+i\eta} \tag{7.11}$$
$$\times\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathcal{M}(\mathbf{w})\;.$$

Thanks to (7.9), the integration measure could be changed to $\mathrm{d}\mu(\mathbf{w})$. The truncated version, $V_*(\mathbf{P},\alpha,\beta)$, is defined analogously but the α and β denominators that correspond to $e\in L(\{0^*\})$ are removed.

We set $Y:=\lambda^{-100}$ and define

$$V_{(*)}(\mathbf{P}) := \frac{e^{2t\eta}}{(2\pi)^2}\iint_{-Y}^{Y}\mathrm{d}\alpha\mathrm{d}\beta\; e^{it(\alpha-\beta)}V_{(*)}(\mathbf{P},\alpha,\beta) \tag{7.12}$$

and

$$E_{(*)g}(\mathbf{P},\mathbf{u}) := \frac{e^{2t\eta}}{(2\pi)^2}\iint_{-Y}^{Y}\mathrm{d}\alpha\mathrm{d}\beta\; E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha})\;, \tag{7.13}$$

where $\boldsymbol{\alpha}$ in $E_{(*)g}(\mathbf{P}, \mathbf{u}, \boldsymbol{\alpha})$ is defined as $\alpha_e = \alpha$ for $e \in \mathcal{L}$ and $\alpha_e := \beta$ for $e \in \widetilde{\mathcal{L}}$. We will call these numbers the *V-value* and *E-value of the partition* $\mathbf{P}$, or sometimes, of the corresponding Feynman graph. Strictly speaking, the V- and E-values depend on ξ through $\Delta = \Delta_\xi$. When this dependence is important, we will make it explicit in the notation, e.g. $V = V_\xi$. Both values depend on the choice of Q as well. When necessary, the notations $V_\xi(\mathbf{P}; Q)$, $E_{\xi,g}(\mathbf{P}, \mathbf{u}; Q)$ etc. will indicate this fact.

Clearly, by using (7.10),

$$\left|V_{(*)}(\mathbf{P})\right| \leq (C\lambda)^{g(\mathbf{P})}\ E_{(*)g}(\mathbf{P}, \mathbf{u} \equiv 0) \tag{7.14}$$

with $g = g(\mathbf{P})$. We will use the notation $E_{(*)g}(\mathbf{P}) := E_{(*)g}(\mathbf{P}, \mathbf{u} \equiv 0)$.

As we will see in (6.2), for the graphical representation of the Duhamel expansion we will really need

$$V^\circ_{(*)}(\mathbf{P}) := \frac{e^{2t\eta}}{(2\pi)^2} \int\!\!\int_{\mathbb{R}} \mathrm{d}\alpha \mathrm{d}\beta\ V_{(*)}(\mathbf{P}, \alpha, \beta)\ , \tag{7.15}$$

i.e. a version of $V_{(*)}(\mathbf{P})$ with unrestricted $\mathrm{d}\alpha\,\mathrm{d}\beta$ integrations. (The circle superscript in V° will refer to the unrestricted version of V). However, the difference between the restricted and unrestricted V-values are negligible even when we sum them up for all partitions:

Lemma 7.1 *Assuming that $\eta \geq \lambda^{2+4\kappa}$ and* (7.7)*, we have*

$$\sum_{\mathbf{P} \in \mathcal{P}_{n,n'}} \left|V_{(*)}(\mathbf{P}) - V^\circ_{(*)}(\mathbf{P})\right| = O(\lambda^{5(n+n')})\ . \tag{7.16}$$

The same result holds if $V_{(*)}(\mathbf{P})$ were defined by restricting the α, β-integral to any domain that contains $[-Y, Y] \times [-Y, Y]$.

Proof. First we consider the case $n, n' \geq 1$. To estimate the difference, we consider the integration domain where either $|\alpha| \geq Y$ or $|\beta| \geq Y$. We assume, for definiteness, that $|\alpha| \geq Y$, and we estimate all α denominators trivially,

$$\frac{1}{|\alpha - \overline{\omega}(w_e) - i\eta|} \leq \frac{C}{\langle\alpha\rangle}\ ,$$

by using that $|\omega(w_e)| \leq \frac{1}{2}w_e^2 + O(\lambda^2) \leq \frac{1}{2}Y + O(\lambda^2)$ on the support of $\mathrm{d}\mu(w_e)$. Then we estimate all but the last β-denominators in (7.11) trivially by $\eta^{-1}\langle\beta/Y\rangle^{-1}$. Thus all w_e integrations are trivial except the last one where we use (3.12). Thanks to the bounds $|\widehat{\psi}(w)|, |\widehat{B}(w)| \leq C\langle w\rangle^{-10d}$, one easily obtains that

$$|V(\mathbf{P}, \alpha, \beta)| \leq \frac{(C\lambda)^{n+n'+g(\mathbf{P})} |\log\lambda|\ \ \log\langle\beta\rangle}{\langle\alpha\rangle^{n+1}\eta^{n'}\langle\beta/Y\rangle^{n'}\langle\beta\rangle^{1/2}}\ .$$

Therefore, we have

$$\int_{\mathbb{R}} \mathrm{d}\beta \int_{\{|\alpha|\geq Y\}} \mathrm{d}\alpha\ \ |V(\mathbf{P}, \alpha, \beta)| \leq \frac{(C\lambda)^{n+n'+g(\mathbf{P})} |\log\lambda|}{\eta^{n'} Y^{n-\frac{1}{2}-2\delta}} = O(\lambda^{6(n+n')})$$

by using (7.7). Similar bound holds for the truncated values, V_*. Thus

$$\Big|V_{(*)}(\mathbf{P}) - V^{\circ}_{(*)}(\mathbf{P})\Big| = O(\lambda^{6(n+n')}) \; . \tag{7.17}$$

Since the total number of partitions, $|\mathcal{P}_{n,n'}|$, is smaller than $(n+n')^{n+n'}$ and in our applications $n, n' \le K \ll \lambda^{-\kappa-2\delta}$, we see that the restriction of the α, β-integral to any domain that contains $[-Y, Y] \times [-Y, Y]$ has a negligible effect of order $O(\lambda^{5(n+n')})$ even after summing up all V-values.

If the condition $n, n' \ge 1$ is not satisfied, say the number of α-denominators is one ($n = 0$), then we will not introduce the auxiliary variable α as in (6.2), because the $\int \mathrm{d}\alpha$ integral would be logarithmically divergent after taking the absolute value. In this case, we use the definition

$$V^{\circ}(\mathbf{P}) := \frac{e^{t\eta}}{2\pi} \int_{\mathbb{R}} \mathrm{d}\beta \; e^{-it\beta} V(\mathbf{P}, \beta)$$

with

$$V^{\circ}(\mathbf{P}, \beta) := \lambda^{n+n'+g(\mathbf{P})} \int \Big(\prod_{e \in \mathcal{L} \cup \widetilde{\mathcal{L}}} \mathrm{d}w_e \Big) \prod_{e \in \widetilde{\mathcal{L}}} \frac{1}{\beta - \omega(w_e) + i\eta}$$
$$\times e^{it\overline{\omega}(w)} \Delta(\mathbf{P}, \mathbf{w}, \mathbf{u} \equiv 0) \mathcal{M}(\mathbf{w}) \; ,$$

directly, instead of V° given in (7.15). Similar modifications hold for the other cases ($n' = 0, n \ge 1$ and $n' = n = 0$) as well. In particular, in our expansion (including all the cases in [19]) only a few such graphs may appear due to $|n - n'| \le 2$. The estimates leading to (7.16) in these cases are similar but much easier than in the general case and they are left to the reader (the same estimates were covered in [16] as well, without the renormalization of the dispersion relation). $\square$

Sometimes we will use a numerical labelling of the edges (see Fig. 2). In this case, we label the edge between $(j-1, j)$ by e_j, the edge between $(\widetilde{j}, \widetilde{j-1})$ by $e_{\widetilde{j}}$. At the special vertices $0, 0^*$ we denote the edges as follows: $e_{n+1} := (n, 0^*)$, $e_{\widetilde{n'+1}} := (0^*, \tilde{n}')$, $e_1 = (0, 1)$ and $e_{\widetilde{1}} := (\widetilde{1}, 0)$. Therefore the edge set $\mathcal{L} = \mathcal{L}(\mathcal{V}_{n,n'})$ is identified with the index set $I_{n+1} \cup \widetilde{I}_{n'+1}$ and we set

$$p_j := w_{e_j}, \qquad \tilde{p}_j := w_{e_{\widetilde{j}}}. \tag{7.18}$$

These two notations will sometimes be used in parallel: the $\mathbf{p}$, $\tilde{\mathbf{p}}$ notation is preferred when distinction is needed between momenta on $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ edges and the $\mathbf{w}$ notation is used when no such distinction is necessary. Note that we always have

$$p_1 - \tilde{p}_1 = \xi \; . \tag{7.19}$$

7.3 Non-repetition Feynman graphs

A partition $\mathbf{P} \in \mathcal{P}_n$ of $I_n \cup \widetilde{I}_n$ is called **even** if for any $P_\mu \in \mathbf{P}$ we have $|P_\mu \cap I_n| = |P_\mu \cap \widetilde{I}_n|$. In particular, in an even partition there are no single lumps, $G(\mathbf{P}) = \emptyset$.

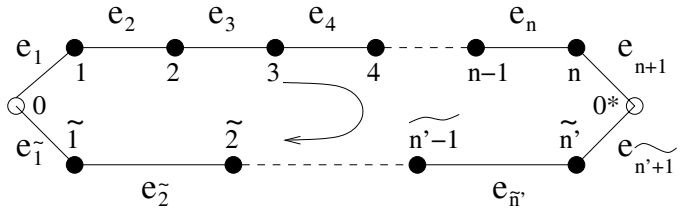

Figure 2: Vertex and edge labels

Let $\mathfrak{S}_n$ be the set of permutations on I_n and let id be the identity permutation. Note that $\mathbf{A} \in \mathcal{A}_n$ and $\sigma \in \mathfrak{S}_n$, uniquely determine an even partition in $\mathbf{P}(\mathbf{A}, \sigma) \in \mathcal{P}_n$, by $I(\mathbf{P}) := I(\mathbf{A})$ and $P_\mu := A_\mu \cup \sigma(A_\mu)$.

Conversely, given an even partition $\mathbf{P} \in \mathcal{P}_n$, we can define its projection onto I_n, $\mathbf{A} := \pi(\mathbf{P}) \in \mathcal{A}_n$, by $I(\mathbf{A}) := I(\mathbf{P})$ and $A_\mu := P_\mu \cap I_n$. We let

$$\mathfrak{S}_n(\mathbf{P}) := \{\sigma \in \mathfrak{S}_n \ : \ \mathbf{P}(\pi(\mathbf{P}), \sigma) = \mathbf{P}\}$$

be the set of permutations that are **compatible** with a given even partition $\mathbf{P}$. In other words, $\sigma \in \mathfrak{S}_n(\mathbf{P})$ if for each $i \in I_n$ the pair $(i, \sigma(i))$ belongs to the same $\mathbf{P}$-lump. Clearly

$$|\mathfrak{S}_n(\mathbf{P})| = \prod_{\mu \in I(\mathbf{P})} \Big(\frac{|P_\mu|}{2}\Big)! = \prod_{\mu \in I(\pi(\mathbf{A}))} |A_\mu|\,! \ . \tag{7.20}$$

We will use the notation

$$V_{(*)}(\mathbf{A}, \sigma, Q) := V_{(*)}(\mathbf{P}(\mathbf{A}, \sigma), Q) \tag{7.21}$$

and similarly for $E_{(*)g}$ and $V^\circ_{(*)}$. In the proofs, Q will be omitted. We also introduce

$$c(\mathbf{A}) := \prod_{\nu \in I(\mathbf{A})} c(|A_\nu|) \ . \tag{7.22}$$

With these notations we can state the representation of the non-repetition terms as a summation over Feynman diagrams:

Proposition 7.2 *With $Q \equiv 1$ and $\xi = 0$ we have*

$$\lim_{L\to\infty} \mathbf{E}' \|\psi'^{\,nr}_{t,k}\|^2_L = \sum_{\sigma \in \mathfrak{S}_k} \sum_{\mathbf{A} \in \mathcal{A}_k} c(\mathbf{A}) V^\circ_{\xi=0}(\mathbf{A}, \sigma, Q \equiv 1) \tag{7.23}$$

and with $Q_\xi(v) := \widehat{\mathcal{O}}(\xi, v)$ we have

$$\lim_{L\to\infty} \mathbf{E}' \langle \widehat{\mathcal{O}}_L, \widehat{W}^\varepsilon_{\psi'^{\,nr}_{t,k}} \rangle_L = \sum_{\sigma \in \mathfrak{S}_k} \sum_{\mathbf{A} \in \mathcal{A}_k} c(\mathbf{A}) \int \mathrm{d}\xi \ V^\circ_{\varepsilon\xi}(\mathbf{A}, \sigma, Q_\xi) \ . \tag{7.24}$$

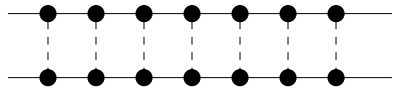

Figure 3: Ladder graph

Proof of Proposition 7.2. We insert (6.6) and (6.9) into (7.7) and we take the limit $L \to \infty$. We use that $\mathbf{E}_M\big[M/|\Lambda_L|\big]^k \to 1$ for any fixed k. We also replace every Riemann sum (3.23) with integrals and we use $\widehat{\psi}_0' \to \widehat{\psi}_0$. By recalling (7.19) and by choosing $Q \equiv 1$ in the definition (7.8), we obtain (7.23). The proof of (7.24) is identical. □.

Proof of Theorem 5.2. We will prove only (5.1); the proof of (5.2) is analogous. Starting with (7.23), we notice that the graph with the trivial partition $\mathbf{A}_0$ and with the identity permutation on I_k gives the main term in Theorem 5.2 since

$$V(k) = V^{\circ}_{\xi=0}(\mathbf{A}_0, id) \ .$$

This graph is called the **ladder graph** (Fig. 3). All other graphs will be negligible.

We first replace $V^{\circ}(\cdots)$ with $V(\cdots)$; the error is negligible by Lemma 7.1. In Section 8 we then estimate $V(\mathbf{A}, \sigma)$ for the trivial partition $\mathbf{A} = \mathbf{A}_0$, where every lump has one element. The result is Proposition 8.6. In this case we set $V(\sigma) := V(\mathbf{A}_0, \sigma)$. In Section 9 we treat the general case $\mathbf{A} \neq \mathbf{A}_0$. The final result of this section is Proposition 9.2. The proof of both propositions rely on Theorem 8.4 that is the core of our method. Its proof is given separately in Section 10. Finally, the proof of Theorem 5.2 follows from Proposition 9.2, together with (7.23), (7.16) and (6.10). □

We remark that the E- and V-values of the partitions depend on the parameters $\lambda, t, \xi, \zeta, k$ and g; a fact that is not explicitly included in the notation. In Sections 8, 9 and 10 we will always assume the following relations

$$\eta = \lambda^{2+\kappa}, \quad t = \lambda^{-2-\kappa}T, \quad T \in [0, T_0], \quad K = [\lambda^{-\delta}(\lambda^2 t)], \quad k < K, \quad \zeta = \lambda^{-\kappa-3\delta}, \quad g \leq 8 \tag{7.25}$$

with a sufficiently small positive $\delta > 0$ that is independent of λ but depends on κ. All estimates will be uniform in ξ and in $T \in [0, T_0]$. We mention that for the proof of Theorem 5.2 we need only $g = 0$, but the more general case is used in [19].

8 Estimates on Feynman graphs without nontrivial lumps

We use the letters p_j, $\tilde{p}_j$, $j \in I_{k+1}$ for the momenta variables (see the convention at the end of Section 7.2) and $I(\mathbf{A}_0) = I_k$ for the index set of the trivial partition. In the following sections we always assume $Q \equiv 1$.

We introduce the restricted version of M° (see (6.8)) as

$$M(k, \mathbf{p}, \tilde{\mathbf{p}}) := \frac{e^{2t\eta}}{(2\pi)^2} \int\!\!\int_{-Y}^{Y} \mathrm{d}\alpha \mathrm{d}\beta \, e^{i(\alpha-\beta)t} \left(\prod_{j=1}^{k+1} \frac{\overline{\widehat{B}}(p_{j+1} - p_j)}{\alpha - \overline{\omega}(p_j) - i\eta} \, \frac{\widehat{B}(\tilde{p}_{j+1} - \tilde{p}_j)}{\beta - \omega(\tilde{p}_j) + i\eta} \right) \tag{8.1}$$

and we also define the trivial estimate of M as

$$N(k,\mathbf{p},\tilde{\mathbf{p}}) := \frac{e^{2t\eta}}{(2\pi)^2}\int\!\!\int_{-Y}^{Y} \mathrm{d}\alpha\mathrm{d}\beta \left(\prod_{j=1}^{k+1} \frac{|\widehat{B}(p_{j+1}-p_j)|}{|\alpha-\overline{\omega}(p_j)-i\eta|}\ \frac{|\widehat{B}(\tilde{p}_{j+1}-\tilde{p}_j)|}{|\beta-\omega(\tilde{p}_j)+i\eta|}\right) . \tag{8.2}$$

The truncated versions of these quantities, denoted by $M_*(k,\mathbf{p},\tilde{\mathbf{p}})$ and $N_*(k,\mathbf{p},\tilde{\mathbf{p}})$, are defined by removing the $(k+1)$-th α and β denominators from the definitions (8.1), (8.2) but keeping all numerators and all other denominators.

From the definition (7.11), we obtain

$$V_{(*)}(\sigma) = \lambda^{2k}\int\!\!\int \mathrm{d}\mathbf{p}\mathrm{d}\tilde{\mathbf{p}}\ M_{(*)}(k,\mathbf{p},\tilde{\mathbf{p}})\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u}\equiv 0)\overline{\widehat{\psi}_0}(p_1)\widehat{\psi}_0(\tilde{p}_1)\ , \tag{8.3}$$

$$E_{(*)}(\sigma,\mathbf{u}) = \lambda^{2k}\int\!\!\int \mathrm{d}\mathbf{p}\mathrm{d}\tilde{\mathbf{p}}\ N_{(*)}(k,\mathbf{p},\tilde{\mathbf{p}})\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u})\overline{\widehat{\psi}_0}(p_1)\widehat{\psi}_0(\tilde{p}_1) \tag{8.4}$$

with

$$\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u}) := \delta(\tilde{p}_{k+1}-p_{k+1}+\xi)\prod_{\ell=1}^{k}\delta\Big(\ p_{\ell+1}-p_\ell-(\tilde{p}_{\sigma(\ell)+1}-\tilde{p}_{\sigma(\ell)})-u_\ell\ \Big)\ .$$

Clearly $|V_{(*)}(\sigma)| \le E_{(*)g=0}(\sigma,\mathbf{u}\equiv 0)$ for any ξ (see (7.14)).

We introduce a convenient notation. For any $(k+1)\times(k+1)$ matrix M and for any vector of momenta $\mathbf{p}=(p_1,\dots p_{k+1})$, we let $M\mathbf{p}$ denote the following $(k+1)$-vector of momenta

$$M\mathbf{p} := \Big(\sum_{j=1}^{k+1} M_{1j}p_j,\ \sum_{j=1}^{k+1} M_{2j}p_j,\dots\Big)\ . \tag{8.5}$$

Furthermore, we introduce the vector $\mathbf{v}=(v_1,\dots,v_{k+1})$ as

$$v_\ell := \xi+u_1+u_2+\dots+u_{\ell-1},\quad \text{for all}\ \ \ell=1,2,\dots,k+1\ . \tag{8.6}$$

Note that $v_{k+1}=\xi$ by (7.1).

Given a permutation $\sigma\in\mathfrak{S}_k$, we define a $(k+1)\times(k+1)$ matrix $M=M(\sigma)$ as follows

$$M_{ij}(\sigma) := \begin{cases} 1 & \text{if} \quad \tilde{\sigma}(j-1)<i\le\tilde{\sigma}(j) \\ -1 & \text{if} \quad \tilde{\sigma}(j)<i\le\tilde{\sigma}(j-1) \\ 0 & \text{otherwise,}\end{cases} \tag{8.7}$$

where, by definition, $\tilde{\sigma}$ is the **extension** of σ to a permutation of $\{0,1,\dots,k+1\}$ by $\tilde{\sigma}(0):=0$ and $\tilde{\sigma}(k+1):=k+1$. In particular $[M\mathbf{p}]_1=p_1$, $[M\mathbf{p}]_{k+1}=p_{k+1}$.

It is easy to check that

$$\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u}) = \prod_{j=1}^{k+1}\delta\Big(\ \tilde{p}_j-[M\mathbf{p}]_j-[M\mathbf{v}]_j\Big)\ , \tag{8.8}$$

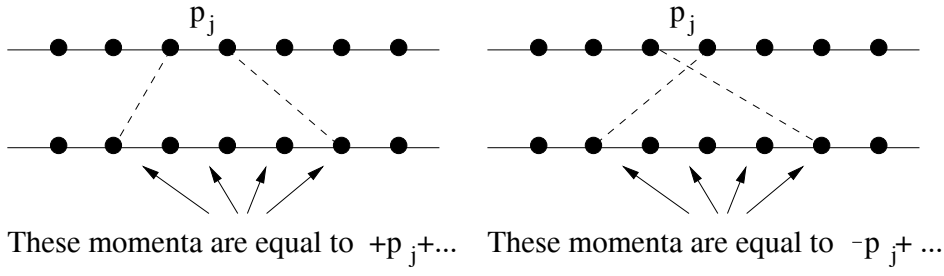

Figure 4: Domain of dependencies of the momenta

in other words, the matrix M encodes the dependence of the $\tilde{p}$-momenta on the p-momenta and the v-momenta. This rule is transparent in the graphical representation of the Feynman graph: the momentum p_j appears in those $\tilde{p}_i$'s which fall into its "domain of dependence", i.e. the section between the image of the two endpoints of p_j, and the sign depends on the ordering of these images (Fig. 4).

Definition 8.1 *A matrix M with entries $0, +1$ or -1 is called* **tower matrix** *if in each column the non-zero entries are consecutive and identical. The collection of these consecutive 1 or -1 entries are called the* **tower** *of that column.*

By construction, the matrix $M(\sigma)$ is a tower matrix.

Proposition 8.2 *For any permutation $\sigma \in \mathfrak{S}_k$ the matrix $M(\sigma)$ is*
(i) invertible;
(ii) totally unimodular, i.e. each subdeterminant is 0 or ± 1.

Proof. The invertibility follows from the fact that $\mathbf{p}$ and $\tilde{\mathbf{p}}$ play a symmetric role in (8.8) if $\mathbf{v} \equiv 0$, $\xi = 0$, in particular $M(\sigma)^{-1} = M(\sigma^{-1})$. It is easy to prove by induction on the size of the matrix that any tower matrix is totally unimodular. □.

The following definition is crucial. It establishes the necessary concepts to measure the complexity of a permutation.

Definition 8.3 (Valley, peak and slope) *Given a permutation $\sigma \in \mathfrak{S}_k$ let $\tilde{\sigma}$ be its extension. A point $(j, \sigma(j))$, $j \in I_k$, on the graph of σ is called* **peak** *if $\tilde{\sigma}(j-1) > \sigma(j) < \tilde{\sigma}(j+1)$, it is called* **valley** *if $\tilde{\sigma}(j-1) < \sigma(j) > \tilde{\sigma}(j+1)$, otherwise it is called* **slope**. *Additionally, the point $(k+1, k+1)$ is also called valley, but $(0,0)$ is not a peak.*

Let $I = \{1, 2, \dots, k+1\}$ denote the set of row indices of M. This set is partitioned into three disjoint subsets, $I = I_v \cup I_p \cup I_s$, such that $i \in I_v, I_p$ or I_s depending on whether $(\tilde{\sigma}^{-1}(i), i)$ is a valley, peak or slope, respectively. For $i \in I_p$ the $\tilde{\sigma}^{-1}(i)$ and $\tilde{\sigma}^{-1}(i)+1$ columns of M are called the **walls** *of the valley $(\tilde{\sigma}^{-1}(i), i)$. Finally, an index $i \in I_v \cup I_s$ is called* **ladder index** *if $|\tilde{\sigma}^{-1}(i) - \tilde{\sigma}^{-1}(i-1)| = 1$. The set of ladder indices is denoted by $I_\ell \subset I$ and their cardinality is denoted by $\ell = \ell(\sigma) := |I_\ell|$. The number of non-ladder indices,*

$$\deg(\sigma) := k + 1 - \ell(\sigma) \,, \tag{8.9}$$

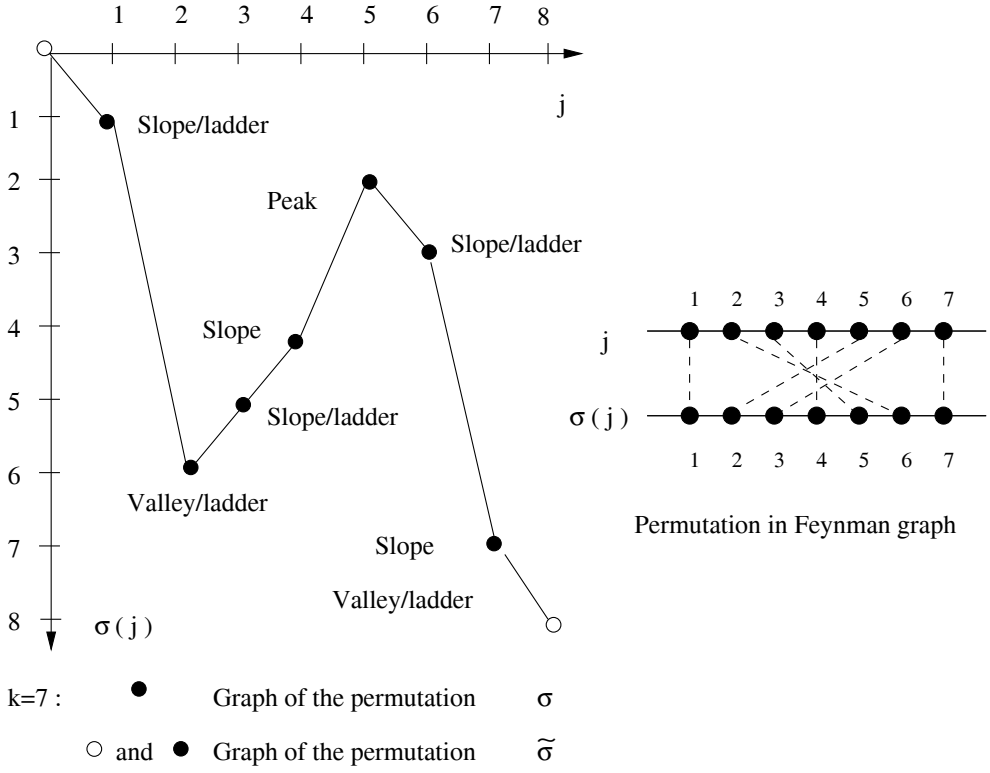

Figure 5: Graph of a permutation

is called the **degree** *of the permutation* σ.

A maximal collection of consecutive ladder indices, $i+1, \ldots, i+b \in I_\ell$*, is called a* **ladder** *of length* b*. The index* i *and* $i+b$ *are called the* **top** *and the* **bottom index of the ladder***, respectively. The set of bottom and top indices are denoted by* $I_b, I_t \subset \{0, 1, \ldots, k+1\}$.

Remarks: (i) The terminology of peak, valley, slope, ladder comes from the graph of the permutation $\tilde{\sigma}$ drawn in a coordinate system where the axis of the dependent variable, $\sigma(j)$, is oriented downward (see Fig. 5). The nonzero entries in the matrix $M(\sigma)$ follow the same geometric pattern as the graph: each downward segment of the graph corresponds to a column with a few consecutive 1's, upward segments correspond to columns with (-1)'s.

(ii) By the definition of the ladder, the index i of the top of a ladder is not a ladder index. Moreover, the permutation $\tilde{\sigma}^{-1}$ restricted to $(i, i+1, \ldots, i+b)$ is clearly a monotonic function.

(iii) All these index sets depend on the permutation σ but we usually omit this dependence from the notation.

(iv) We note that for the special case of the identity permutation $\sigma = id$ we have $I_p = \emptyset$, $I_s = \{1, 2, \ldots, k\}$, $I_v = \{k+1\}$ and $I_\ell = \{1, 2, \ldots, k+1\}$. In particular, $\deg(id) = 0$ and $\deg(\sigma) \geq 2$ for any other permutation $\sigma \neq id$.

An example is shown on Fig. 5. The matrix corresponding to the permutation on this figure is

the following (zero entries are left empty)

$$M(\sigma) := \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & 1 & & & -1 & 1 & & \\ & 1 & & & -1 & & 1 & \\ & 1 & & -1 & & & 1 & \\ & 1 & -1 & & & & 1 & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} 1 & s/\ell \\ 2 & p \\ 3 & s/\ell \\ 4 & s \\ 5 & s/\ell \\ 6 & v/\ell \\ 7 & s \\ 8 & v/\ell \end{matrix} \tag{8.10}$$

The numbers on the right indicate the column indices and the letters show whether it is peak/slope/valley or ladder. In this case $I_p = \{2\}$, $I_v = \{6, 8\}$, $I_s = \{1, 3, 4, 5, 7\}$. $I_\ell = \{1, 3, 5, 6, 8\}$, $I_t = \{0, 2, 4, 7\}$, $I_b = \{1, 6, 3, 8\}$ and $\deg(\sigma) = 3$. There are four ladders, three of them is of length one, one is of length two.

Now we are ready to estimate $|V(\sigma)| \le E_{g=0}(\sigma, \mathbf{u} \equiv 0)$. The following theorem shows that the degree of the permutation, $\deg(\sigma)$, measures the size of $V(\sigma)$. The proof is the key step in our method and it is given in Section 10.

Theorem 8.4 *Assume* (7.25) *with* $\kappa < \frac{2}{10d+9}$ *and let* $\sigma \in \mathfrak{S}_k$. *Then the* E*-value of the graph of the trivial partition with permutation* σ *is estimated by*

$$\sup_{\mathbf{u}} E_{(*)g}(\sigma, \mathbf{u}) \le C\Big(\lambda^{\frac{1}{3}-(\frac{5}{3}d+\frac{3}{2})\kappa-O(\delta)}\Big)^{\deg(\sigma)} |\log\lambda|^2 \tag{8.11}$$

if $\lambda \ll 1$.

This theorem is complemented by the following combinatorial lemma.

Lemma 8.5 *Let* $k \le O(\lambda^{-\kappa-\delta})$, $D \ge 0$ *integer, and let* $\gamma > \kappa + \delta$ *be fixed. Then*

$$\sum_{\substack{\sigma\in\mathfrak{S}_k \\ \deg(\sigma)\ge D}} \lambda^{\gamma\deg(\sigma)} \le O\Big(\lambda^{D(\gamma-\kappa-\delta)}\Big) \tag{8.12}$$

for $\lambda \ll 1$.

Since $\deg(\sigma) \ge 2$ if $\sigma \ne id$, from Theorem 8.4, Lemma 8.5, $g(\mathbf{P}) = 0$ and the estimate $|V(\sigma)| \le E_{g=0}(\sigma, \mathbf{u} \equiv 0)$ we immediately obtain:

Proposition 8.6 *Assuming* (7.25) *with* $\kappa < \frac{2}{10d+15}$ *we have*

$$\sum_{\substack{\sigma\in\mathfrak{S}_k \\ \sigma\ne id}} \sup_{\mathbf{u}} |V(\sigma)| \le O\Big(\lambda^{\frac{2}{3}-(\frac{10}{3}d+5)\kappa-O(\delta)}\Big) \tag{8.13}$$

for $\lambda \ll 1$. $\square$

Proof of Lemma 8.5. Notice that $\ell(\sigma) = k+1$ only if $\sigma = id$, for all other permutations $\ell(\sigma) \le k-1$. We shall prove that, for any ℓ,

$$\#\{\sigma \in \mathfrak{S}_k \ : \ \ell(\sigma) = \ell\} \le (Ck)^{k-\ell+1} \ . \tag{8.14}$$

Then (8.12) follows by recalling $k - \ell(\sigma) + 1 = \deg(\sigma)$ and $k \le K = O(\lambda^{-\kappa-\delta})$ and summing up the geometric series

$$\sum_{\substack{\sigma\in\mathfrak{S}_k \\ \deg(\sigma)\ge D}} \lambda^{\gamma\deg(\sigma)} \le \sum_{m=D}^{k+1} \left[C\lambda^{(\gamma-\kappa-\delta)}\right]^m .$$

To prove (8.14), let σ be a permutation with m ladders of size $b_1, \dots, b_m$ such that $\sum_j b_j = \ell$ and $b_j \ge 1$. If we remove these ladder indices, we have $k-\ell+1$ indices in $I \setminus I_\ell = \{i_1, i_2, \dots, i_{k-\ell+1}\}$. The permutation σ induces a unique permutation $\sigma^* \in \mathfrak{S}_{k-\ell+1}$ on the indices of this set by $\sigma^*(j) < \sigma^*(j')$ iff $\sigma(i_j) < \sigma(i_{j'})$.

Suppose that $k+1 \notin I_\ell$, then the number of allowed permutations σ^* is $(k-\ell)!$ since $\sigma(k+1) = k+1$ determines σ^* on the last element.

Let $I^* := (I \setminus I_\ell) \cup \{0\}$, then $g := |I^*| = k - \ell + 2$. Clearly the top of any ladder belongs to I^* and each element of I^* can be the top of at most one ladder, except $k+1$. We assign the length b_j of the ladder to its top and for simplicity, we assign the value zero to any other element of $I^* \setminus \{k+1\}$. Thus we obtain numbers $b_1, \dots, b_{g-1}$ with $\sum_{j=1}^{g-1} b_j = \ell$ and $b_j \ge 0$.

If the permutation σ^* and the numbers $b_1, \dots, b_{g-1}$ are given, then we have 2^{g-1} ways to reconstruct the original permutation σ. To see this, first notice that the tops are identified by the condition $b_j > 0$ and the corresponding ladder in σ can grow either "to the right" or "to the left" from its bottom point on the graph of σ. Therefore, the number of permutations $\sigma \in \mathfrak{S}_k$ with ℓ ladder indices is bounded by

$$\begin{aligned} &2^{g-1}(k-\ell)! \times \#\Big\{(b_1, b_2, \dots, b_{g-1}) \ : \ \sum_{j=1}^{g-1} b_j = \ell, \ b_j \ge 0\Big\} \\ &\le 2^{k-\ell+1}(k-\ell)! \binom{g-2+\ell}{g-2} \le (Ck)^{k-\ell+1} \ . \end{aligned}$$

The estimate for the case $k+1 \in I_\ell$ is similar. □

9 Estimates on Feynman graphs with nontrivial lumps

In this section we estimate $V(\mathbf{A}, \sigma)$ for a general partition $\mathbf{A}$. We start with a definition.

Definition 9.1 *(i) Let $\mathbf{A} \in \mathcal{A}_k$. Set $a_\nu := |A_\nu|$, $\nu \in I(\mathbf{A})$, to be the size of the ν-th lump. Let*

$$S(\mathbf{A}) := \bigcup_{\substack{\nu\in I(\mathbf{A}) \\ a_\nu \ge 2}} A_\nu$$

be the union of nontrivial lumps. The cardinality of this set, $s(\mathbf{A}) := |S(\mathbf{A})|$, is called the **degree of the partition A**.

(ii) Let $\mathbf{A} \in \mathcal{A}_k$ and $\sigma \in \mathfrak{S}_k$. The number

$$q(\mathbf{A}, \sigma) := \max\Big\{\deg(\sigma), \frac{1}{2}s(\mathbf{A})\Big\} \tag{9.1}$$

is called the **joint degree** *of the pair $(\sigma, \mathbf{A})$ of the permutation σ and partition $\mathbf{A}$.*

The goal is the following generalization of Proposition 8.6 that includes summations over nontrivial lumps as well.

Proposition 9.2 *We assume* (7.25)*. Let D and s be given integers, let $q := \max\{D, \frac{1}{2}s\}$. For any $\kappa < \frac{2}{34d+39}$ we have*

$$\Xi(k, D, s) := \sum_{\substack{\sigma\in\mathfrak{S}_k \\ \deg(\sigma)\ge D}} \sum_{\substack{\mathbf{A}\in\mathcal{A}_k \\ s(\mathbf{A})\ge s}} \sup_{\mathbf{u}, g\le 8} E_{(*)g}(\mathbf{A}, \sigma, \mathbf{u})|c(\mathbf{A})| \le C\Big(\lambda^{\frac{1}{3}-(\frac{17}{3}d+\frac{13}{2})\kappa-O(\delta)}\Big)^q |\log\lambda|^2 \,. \tag{9.2}$$

Proof of Proposition 9.2. The following lemma shows that any even partition $\mathbf{P} \in \mathcal{P}_k$ can be generated by a permutation with high degree, depending on the size of nontrivial lumps. The proof will be given later.

Lemma 9.3 *For any even partition $\mathbf{P} \in \mathcal{P}_k$ there exists a compatible permutation $\widehat{\sigma} = \widehat{\sigma}(\mathbf{P}) \in \mathfrak{S}_k(\mathbf{P})$ such that*

$$\deg(\widehat{\sigma}) \ge \frac{1}{2}s(\pi(\mathbf{P})) \,. \tag{9.3}$$

Corollary 9.4 *Given $\sigma \in \mathfrak{S}_k$ and $\mathbf{A} \in \mathcal{A}_k$, we have, for $\kappa \le \frac{2}{34d+9}$*

$$\sup_{\mathbf{u}} E_{(*)g}(\mathbf{A}, \sigma, \mathbf{u}) \le C|\log\lambda|^2\Big(\lambda^{\frac{1}{3}-(\frac{17}{3}d+\frac{3}{2})\kappa-O(\delta)}\Big)^{q(\mathbf{A},\sigma)} \,. \tag{9.4}$$

Proof of Corollary 9.4. We define a permutation $\sigma^* := \sigma^*(\mathbf{A}, \sigma)$ as $\sigma^* := \sigma$ if $\deg(\sigma) \ge \frac{1}{2}s(\mathbf{A})$, and $\sigma^* := \widehat{\sigma}(\mathbf{P}(\mathbf{A}, \sigma))$ otherwise. By Lemma 9.3 we have $\deg(\sigma^*) = q(\mathbf{A}, \sigma)$. Clearly $\mathbf{P}(\mathbf{A}, \sigma) = \mathbf{P}(\mathbf{A}, \sigma^*)$, in particular, $E_{(*)g}(\mathbf{A}, \sigma, \mathbf{u}) = E_{(*)g}(\mathbf{A}, \sigma^*, \mathbf{u})$.

We wish to estimate the value of an arbitrary partition $\mathbf{A}$ by that of the trivial partition $\mathbf{A}_0$. We can artificially break up lumps and using the auxiliary momenta to account for the additional Kirchhoff rules. We describe this procedure in full generality for any circle graph. We will call it Operation I because further similar operations will be introduced in the companion paper [19].

Operation I: Breaking up lumps

Consider a circle graph on N vertices (Section 7.1). Given a partition of the set $\mathcal{V} \setminus \{0, 0^*\}$, $\mathbf{P} = \{P_\mu \,:\, \mu \in I(\mathbf{P})\} \in \mathcal{P}_\mathcal{V}$, we define a new partition $\mathbf{P}^*$ by breaking up one of the lumps into two smaller nonempty lumps. Let $P_\nu = P_{\nu'} \cup P_{\nu''}$ with $P_{\nu'} \cap P_{\nu''} = \emptyset$ and $\mathbf{P}^* = \{P_{\nu'}, P_{\nu''}, P_\mu \,:\, \mu \in I(\mathbf{P}) \setminus \{\nu\}\}$. In particular $I(\mathbf{P}^*) = I(\mathbf{P}) \cup \{\nu', \nu''\} \setminus \{\nu\}$ and $m(\mathbf{P}^*) = m(\mathbf{P}) + 1$.

Lemma 9.5 *With the notation above, we have*

$$E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha}) \le \int_{|r|\le N\zeta} \mathrm{d}r\; E_{(*)g}(\mathbf{P}^*,\mathbf{u}^*(r,\nu),\boldsymbol{\alpha})\;,$$

where the new set of momenta $\mathbf{u}^* = \mathbf{u}^*(r,\nu)$ *is given by* $u^*_\mu := u_\mu$, $\mu \in I(\mathbf{P})\setminus\{\nu\}$ *and* $u^*_{\nu'} = u_\nu - r$, $u^*_{\nu''} = r$. *In our estimates we will always have* $N \le 2K$ *and then*

$$\sup_{\mathbf{u}} E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha}) \le \Lambda E_{(*)g}\sup_{\mathbf{u}}(\mathbf{P}^*,\mathbf{u},\boldsymbol{\alpha})$$

with $\Lambda := [CK\zeta]^d = O(\lambda^{-2d\kappa - O(\delta)})$ *(see* (4.7) *and* (7.4)*).*

Proof of Lemma 9.5. The break-up of the lump P_ν corresponds to

$$\delta\Big(\sum_{e\in L_\pm(P_\nu)} \pm w_e - u_\nu\Big) = \int \mathrm{d}r\;\delta\Big(\sum_{e\in L_\pm(P_{\nu'})} \pm w_e - u_\nu + r\Big)\delta\Big(\sum_{e\in L_\pm(P_{\nu''})} \pm w_e - r\Big)\;. \tag{9.5}$$

Note that $L(P_\nu)\subset L(P_{\nu'})\cup L(P_{\nu''})$ and for any edge $e\in (L(P_{\nu'})\cup L(P_{\nu''}))\setminus L(P_\nu)$, we inserted an extra $w_e - w_e$ in the left hand side of (9.5). Note that the property (7.1) on the sum of auxiliary momenta is preserved. The integration in (9.5) can be restricted to $|r|\le N\zeta$ since $|w_e|\le\zeta$ for all e. □

We return to the proof of Corollary 9.4. By applying the break-up operation not more than $s(\mathbf{A})$ times, and by using Lemma 9.5, we obtain

$$\sup_{\mathbf{u}} E_{(*)g}(\mathbf{A},\sigma^*,\mathbf{u}) \le \Lambda^{s(\mathbf{A})}\sup_{\mathbf{u}} E_{(*)g}(\mathbf{A}_0,\sigma^*,\mathbf{u})$$

with $\Lambda := [CK\zeta]^d$. Then (9.4) immediately follows from $s(\mathbf{A})\le 2q(\mathbf{A},\sigma)$ and from Theorem 8.4. □

Proof of Lemma 9.3. For any $\sigma\in\mathfrak{S}_k$ we define the set of **internal ladder indices** as

$$I^*_\ell = I^*_\ell(\sigma) := \{i\in I_k\;:\;|\tilde\sigma^{-1}(i-1)-\tilde\sigma^{-1}(i)| = |\tilde\sigma^{-1}(i+1)-\tilde\sigma^{-1}(i)| = 1\}\subset I_\ell$$

where $\tilde\sigma$ is the extension of σ as before. The indices $i-1, i+1$ are called the **protectors** of the internal ladder index $i\in I^*_\ell$. They ensure that the index i is neither the bottom nor the top index of the ladder.

We first claim that for any $\sigma\in\mathfrak{S}_k$ we have

$$k - |I^*_\ell(\sigma)| \le 2\mathrm{deg}(\sigma) \tag{9.6}$$

To see this inequality, we first recall the definition of a ladder and its bottom and top indices from Definition 8.3. Since every ladder has a unique bottom and top index, that are not internal ladder indices, we see that the sets I^*_ℓ, I_b and I_t are disjoint subsets of $\{0,1,\dots,k+1\}$, hence $|I^*_\ell|+|I_t|+|I_b|\le k+2$. Moreover $|I_t| = |I_b|$. Finally, notice that $I_\ell = I^*_\ell\cup I_b$ is a disjoint union, therefore $\ell = |I^*_\ell| + |I_b|$. From these three relations and $\mathrm{deg}(\sigma) = k-\ell+1$, (9.6) follows.

Next we show that there exists a compatible permutation $\widehat{\sigma} \in \mathfrak{S}_k(\mathbf{P})$,

$$I^*_\ell(\widehat{\sigma}) \cap \widehat{\sigma}(S(\mathbf{A})) = \emptyset \;, \tag{9.7}$$

where we set $\mathbf{A} := \pi(\mathbf{P})$ for simplicity. Combining this fact with I^*_ℓ and $\widehat{\sigma}(S(\mathbf{A})) \subset \{1, 2, \dots, k\}$ and with (9.6), we will obtain (9.3).

To construct $\widehat{\sigma}$ satisfying (9.7), we apply a greedy algorithm. Since $\mathbf{P}$ is even, $\mathfrak{S}_k(\mathbf{P})$ is nonempty and we pick a $\sigma_0 \in \mathfrak{S}_k(\mathbf{P})$. If (9.7) is not satisfied for σ_0, then some internal ladder index i is in the image of a nontrivial lump $A \in \mathbf{A}$; $i = \sigma_0(i')$, $i' \in A$. Let $j' \in A$ be another element of this lump. Flip the permutation σ_0 on these two elements, i.e. define a new permutation σ_1 such that $\sigma_1(i') := \sigma_0(j') = j$, $\sigma_1(j') := \sigma_0(i') = i$, otherwise $\sigma_1 := \sigma_0$. Clearly $\sigma_1 \in \mathfrak{S}_k(\mathbf{P})$, $\sigma_1(s(\mathbf{A})) = \sigma_0(s(\mathbf{A}))$. We claim that

$$|I^*_\ell(\sigma_1) \cap \sigma_1(S(\mathbf{A}))| < |I^*_\ell(\sigma_0) \cap \sigma_0(S(\mathbf{A}))| \;, \tag{9.8}$$

i.e. the total number of internal ladder indices in nontrivial lumps decreased. Continuing this flipping process for σ_1 etc., we obtain a permutation $\widehat{\sigma}$ satisfying (9.7).

To see (9.8) we note that after the flip i is not an internal ladder index any more. This is clear if $j \neq i-1, i+1$; in that case the points $(\tilde{\sigma}^{-1}(i-1), i-1)$ and $(\tilde{\sigma}^{-1}(i+1), i+1)$ have not changed and they would uniquely fix the location of an internal ladder index in between. The preimage of the index i has moved out from this position, $\sigma_1^{-1}(i) \neq \sigma_0^{-1}(i)$. The index j however would not become internal ladder since $\sigma_1^{-1}(j) = i'$ is between $\sigma_1^{-1}(i-1)$ and $\sigma_1^{-1}(i+1)$, but j is not between $i-1$ and $i+1$. It is easy to see that the fixed points $(\tilde{\sigma}^{-1}(i-1), i-1)$ and $(\tilde{\sigma}^{-1}(i+1), i+1)$ also prevent any other indices from becoming an internal ladder index after the flip. This could only be possible if due to the new point $(\tilde{\sigma}_1^{-1}(j), j) = (\tilde{\sigma}_0^{-1}(i), j)$, one of the neighbors of j, say $j+1$, would become an internal ladder index. It is easy to see that then $j+1$ must be equal to $i-1$ and the other protector of the new internal ladder index $j+1$ must be i. In this case $i-1$ was already an internal ladder index before the flip as well, so no new internal ladder was created.

A similar but simpler argument shows that if $j = i-1$ or $j = i+1$, the number of internal ladder indices also decreases. This completes the proof of Lemma 9.3 $\square$

We continue the proof of Proposition 9.2. Given $\sigma \in \mathfrak{S}_k$ and $\mathbf{A} \in \mathcal{A}_k$, we recall the permutation $\sigma^* = \sigma^*(\mathbf{A}, \sigma)$ defined in the proof of Corollary 9.4. We also note that $s(\mathbf{A}) \le 2\deg(\sigma^*)$. Hence

$$\Xi(k, D, s) = \sum_{\substack{\mathbf{A} \in \mathcal{A}_k \\ s(\mathbf{A}) \ge s}} |c(\mathbf{A})| \sum_{\substack{\sigma^* \in \mathfrak{S}_k \\ \deg(\sigma^*) \ge q}} \sum_{\substack{\sigma \in \mathfrak{S}_k \\ \sigma^*(\mathbf{A}, \sigma) = \sigma^*}} \sup_{\mathbf{u}} E_{(*)g}(\mathbf{A}, \sigma, \mathbf{u}) \;.$$

Note that $\sigma \in \mathfrak{S}_k(\mathbf{P}(\mathbf{A}, \sigma^*))$, so by (7.20) the summation over σ contributes by a factor of at most $\prod_\nu a_\nu!$ and we obtain

$$\Xi(k, D, s) \le \sum_{\substack{\sigma^* \in \mathfrak{S}_k \\ \deg(\sigma^*) \ge q}} \sum_{\substack{\mathbf{A} \in \mathcal{A}_k \\ s(\mathbf{A}) \le 2\deg(\sigma^*)}} \Big(\prod_{\nu \in I(\mathbf{A})} a_\nu^{a_\nu - 2} a_\nu! \Big) \sup_{\mathbf{u}} E_{(*)g}(\mathbf{A}, \sigma^*, \mathbf{u}) \;.$$

We also used the estimate (6.10). By using (9.4) and $\deg(\sigma^*) = q(\mathbf{A}, \sigma)$, we obtain

$$\Xi(k, D, s) \le \sum_{\substack{\sigma^* \in \mathfrak{S}_k \\ \deg(\sigma^*) \ge q}} \Big(C\lambda^{\frac{1}{3} - (\frac{17}{3}d + \frac{3}{2})\kappa - O(\delta)} \Big)^{\deg(\sigma^*)} \sum_{\substack{\mathbf{A} \in \mathcal{A}_k \\ s \le s(\mathbf{A}) \le 2\deg(\sigma^*)}} \Big(\prod_{\nu \in I(\mathbf{A})} a_\nu^{a_\nu - 2} a_\nu! \Big) .$$

We introduce the notation

$$\sum\nolimits^* f(a_\nu) := \sum_{\substack{\nu \in I(\mathbf{A}) \\ a_\nu \ge 2}} f(a_\nu) , \qquad \prod\nolimits^* f(a_\nu) := \prod_{\substack{\nu \in I(\mathbf{A}) \\ a_\nu \ge 2}} f(a_\nu) .$$

First we fix the sizes of the nontrivial lumps $a_\nu \ge 2$. Given these sizes, the number of $\mathbf{A}$ partitions is bounded by

$$\binom{k}{a_1}\binom{k-a_1}{a_2}\binom{k-a_1-a_2}{a_3}\cdots \le \frac{k!}{(k - \sum^* a_\nu)! \prod^* a_\nu!} \le \frac{k^{\sum^* a_\nu}}{\prod^* a_\nu!} .$$

Recalling $s(\mathbf{A}) = \sum^* a_\nu$, and that $s(\mathbf{A}) \le 2\deg(\sigma^*)$, we have

$$\Xi(k, D, s) \le \sum_{\substack{\sigma^* \in \mathfrak{S}_k \\ \deg(\sigma^*) \ge q}} (Ck^2\lambda^\tau)^{\deg(\sigma^*)} \sum_{a_\nu \,:\, \sum^* a_\nu \le 2\deg(\sigma^*)} \Big(\prod\nolimits^* a_\nu^{a_\nu - 2} \Big) \tag{9.9}$$

with $\tau < \frac{1}{3} - (\frac{17}{3}d + \frac{3}{2})\kappa$. We use the bound $a^{a-2} \le C^{a-1}(a-1)!$. To estimate the summation over a_ν's we use the following inequality. For any fixed m, H we have

$$\sum\nolimits^{\#} \Big(\prod\nolimits^* (a_\nu - 1)! \Big) \le (H-1)! \tag{9.10}$$

where the summation $\#$ is over all sequences $(a_1, a_2, \dots, a_m)$ of positive integers at least 2, whose sum is H. The proof of (9.10) is easily obtained by induction on m from

$$\sum_{a=2}^{H-2} (a-1)!(H-a-1)! \le \sum_{a=2}^{H-2} (H-2)! < (H-1)! .$$

Summing (9.10) over all $H \le 2\deg(\sigma^*)$ and $m \le H/2$, we obtain the bound

$$\sum_{a_\nu \,:\, \sum^* a_\nu \le 2\deg(\sigma^*)} \Big(\prod\nolimits^* a_\nu^{a_\nu - 2} \Big) \le 2 \big[2\deg(\sigma^*) \big]! \le (Ck)^{2\deg(\sigma^*)}$$

for the a_ν summation in (9.9) since $\deg(\sigma^*) \le k+1$ by definition.

In summary, we obtain from (9.9)

$$\Xi(k, D, s) \le \sum_{\substack{\sigma \in \mathfrak{S}_k \\ \deg(\sigma) \ge q}} (Ck^4\lambda^\tau)^{\deg(\sigma)}$$

Recalling that $k = O(\lambda^{-\kappa-\delta})$, we can apply Lemma 8.5 with $\gamma = \tau - 4(\kappa+\delta)$ as long as $\gamma > \kappa + \delta$. For sufficiently small positive δ this gives the condition $\kappa < \frac{2}{34d+39}$ in Proposition 9.2 and the estimate (9.2). $\square$

10 Proof of Theorem 8.4

We set

$$E(M) := \sup_{\tilde{\mathbf{u}},\xi} \lambda^{2k} \int\!\!\int_{-Y}^{Y} \mathrm{d}\alpha \mathrm{d}\beta \int \mathrm{d}\mu(\mathbf{p})\, |\mathcal{B}(\mathbf{p})|\, |\mathcal{B}(M\mathbf{p}+\tilde{\mathbf{u}})| |\widehat{\psi}(p_1)| |\widehat{\psi}(p_1+\tilde{u}_1)| \times \prod_{j=1}^{k+1} \frac{1}{|\alpha - \overline{\omega}(p_j) - i\eta|} \frac{1}{|\beta - \omega([M\mathbf{p}+\tilde{\mathbf{u}}]_j) + i\eta|} \ . \tag{10.1}$$

The key step in the proof of Theorem 8.4 is the following Lemma:

Lemma 10.1 *For any $\sigma \in \mathfrak{S}_k$ we have*

$$E(M(\sigma)) \le \left(\lambda^{\frac{1}{3}-(\frac{5}{3}d+\frac{3}{2})\kappa - O(\delta)}\right)^{\deg(\sigma)} |\log \eta|^2 \ , \tag{10.2}$$

where the matrix $M(\sigma)$ was defined in (8.7) *and the degree of the permutation, $\deg(\sigma)$, was defined in* (8.9).

From (7.13), (8.4), (8.8), clearly

$$\sup_{\mathbf{u}} E_{g=0}(\sigma, \mathbf{u}) \le \frac{e^{2t\eta}}{(2\pi)^2} E(M(\sigma)) \tag{10.3}$$

after integrating out all $\tilde{p}_j$ variables in (8.4) and by using $\tilde{p}_1 = p_1 - \xi$. The estimate (10.2) will then complete the proof of Theorem 8.4 for $g = 0$.

The proof of Theorem 8.4 for other (but finitely many) g values follows exactly in the same way. This requires to slightly redefine $\mathcal{B}(\mathbf{p})$ (see (6.5)) in the definition of $E(M)$ by allowing the factor $\langle p_{j+1} - p_j\rangle^{-2d}$ instead of $\widehat{B}(p_{j+1} - p_j)$ at a few places exactly as in the definition of $\mathcal{N}_{\mathcal{G}}$ (see (7.3)). As we will see along the proof of (10.2), this change will require using the less precise bound (3.13) with $a = 0$ and $h(p-q) = \langle p-q\rangle^{-2d}$ instead of the more accurate estimate (3.15) at most g times. Each time we lose a constant factor compared with the proof for $g = 0$. Since $g \le 8$, this results only in a constant factor. Finally, the proof for the truncated E-values requires to define a truncated version of $E(M)$, where the last product in (10.1) runs only up to $j = k$, i.e. the last α and β denominators are not present. It will be clear from the proof of Lemma 10.1 that the same bound holds for the truncated version of $E(M)$ as well. □

10.1 Pedagogical detour

The size of the multiple integral in (10.1) heavily depends on the structure of $M = M(\sigma)$. Before we go into the algorithm to evaluate this multiple integral, we present two calculations, that introduce the techniques that we are going to use in the actual proof. The second calculation also provides the bound (10.2), hence (8.11), for the case of the trivial permutation $\deg(\sigma) = 0$.

10.1.1 Method I. Pointwise bound

The most straightforward bound on (10.1) estimates all but one of the β-denominators by L^∞ norm

$$\sup_{\beta,\mathbf{p},\tilde{\mathbf{u}},j} \frac{1}{|\beta - \omega([M\mathbf{p}+\tilde{\mathbf{u}}]_j) + i\eta|} \le \eta^{-1} . \tag{10.4}$$

It would be possible to estimate this denominator by $C\lambda^{-2}$ apart from a neighborhood of zero using (3.7) and treat the $|p| \sim 0$ regime separately. Here we choose the simplest argument and we do not optimize for the best possible exponent κ.

We use (10.4) k times to obtain

$$E(M) \le \Big(\frac{C\lambda^2}{\eta}\Big)^k \sup_{\tilde{\mathbf{u}}} \int \frac{\mathrm{d}\mu(p_{k+1})}{\langle p_{k+1}+\tilde{u}_{k+1}\rangle^{2d}} \iint_{-Y}^{Y} \frac{\mathrm{d}\alpha\mathrm{d}\beta}{|\alpha - \overline{\omega}(p_{k+1}) - i\eta||\beta - \omega(p_{k+1}+\tilde{u}_{k+1}) + i\eta|}$$
$$\times \int \dots \int \mathrm{d}p_1 \dots \mathrm{d}p_k \prod_{j=1}^{k} \frac{|\widehat{B}(p_{j+1}-p_j)|}{|\alpha - \overline{\omega}(p_j) - i\eta|} .$$

Here we used the following bound for any $\mathbf{q} = (q_1, \dots, q_{k+1})$:

$$|\mathcal{B}(\mathbf{q})||\widehat{\psi}_0(q_1)| \le C^k \langle q_{k+1}\rangle^{-2d} \tag{10.5}$$

to obtain the decay in $\tilde{p}_{k+1} = p_{k+1} + \tilde{u}_{k+1}$. We integrate out $p_1, p_2, \dots, p_k$ using (3.12), then we perform the $\mathrm{d}\alpha, \mathrm{d}\beta$ integrals and finally $\mathrm{d}p_{k+1}$ to obtain

$$E(M) \le (C|\log\lambda|)^{k+2}(\lambda^2\eta^{-1})^k . \tag{10.6}$$

The estimate (10.6) is off by a factor $(\lambda^2\eta^{-1})^k = (\lambda^{-\kappa})^k$ because we did not use the stronger estimate mentioned after (10.4). We also collected many logarithmic factors and the constant is not optimal. We note that in the typical term $k \sim \lambda^2 t \sim \lambda^{-\kappa} \gg 1$, so even an error C^k may not be affordable. To improve this estimate, for a typical matrix M, we will not use the pointwise bound (10.4) for all β-denominators. We will carefully select those β-denominators whose singularities cannot overlap with other singularities, hence they can be integrated out at a $|\log\eta|$ expense instead of η^{-1}.

Before we explain this algorithm, we show another method to estimate $E(M)$. It practically estimates $E(M)$ by $E(I)$, i.e. by the ladder graph, that can be computed more precisely. The same calculation will be important when evaluating embedded ladder graphs.

10.1.2 Method II. Successive integration scheme

We separate all but one α and β denominator by a Schwarz' inequality. We obtain

$$E(M) \le \lambda^{2k} \sup_{\tilde{\mathbf{u}}} \int \mathrm{d}\mu(\mathbf{p}) \iint_{-Y}^{Y} \frac{\mathrm{d}\alpha\mathrm{d}\beta}{|\alpha - \overline{\omega}(p_1) - i\eta||\beta - \omega(p_1+\tilde{u}_1) + i\eta|}$$
$$\times \left[|\widehat{\psi}_0(p_1)|^2 \prod_{j=2}^{k+1} \frac{|\widehat{B}(p_j - p_{j-1})|^2}{|\alpha - \overline{\omega}(p_j) - i\eta|^2} + |\widehat{\psi}_0(q_1)|^2 \prod_{j=2}^{k+1} \frac{|\widehat{B}(q_j - q_{j-1})|^2}{|\beta - \omega(q_j) + i\eta|^2} \right]$$

with the shorthand notation $\mathbf{q} := M\mathbf{p} + \tilde{\mathbf{u}}$. By using that M is an invertible matrix with determinant ± 1 (Proposition 8.2), the contributions of the two terms in the square bracket are identical up to exchange of α and β. To estimate the first term, we use iteratively (3.15) and (3.13) (with $a = 1/2$) to integrate out $p_{k+1}, p_k, \ldots, p_2$ (in this order):

$$\lambda^2 \int \frac{|\widehat{B}(p_{k+1} - p_k)|^2}{|\alpha - \overline{\omega}(p_{k+1}) - i\eta|^2} \mathrm{d}p_{k+1} \leq (1 + C\lambda^{1-12\kappa}) \Big[1 + \lambda^{-1} |\alpha - \omega(p_k)|^{1/2} \Big] \tag{10.7}$$

$$\lambda^2 \int \frac{|\widehat{B}(p_k - p_{k-1})|^2}{|\alpha - \overline{\omega}(p_k) - i\eta|^2} \Big[1 + \lambda^{-1} |\alpha - \omega(p_k)|^{1/2} \Big] \mathrm{d}p_k \leq (1 + C\lambda^{1-12\kappa}) \Big[1 + \lambda^{-1} |\alpha - \omega(p_{k-1})|^{1/2} \Big] \tag{10.8}$$

etc. In the last step we use only (3.13) once for $a = 0$ and once for $a = 1/2$:

$$\lambda^2 \int \frac{|\widehat{B}(p_2 - p_1)|^2}{|\alpha - \overline{\omega}(p_2) - i\eta|^2} \Big[1 + \lambda^{-1} |\alpha - \omega(p_2)|^{1/2} \Big] \mathrm{d}p_2 \leq C \, . \tag{10.9}$$

Then we integrate $\mathrm{d}\alpha \, \mathrm{d}\beta$ and finally $\mathrm{d}p_1$ to obtain

$$E(M) \leq C(1 + C\lambda^{1-12\kappa})^k |\log \lambda|^2 \leq C |\log \lambda|^2 \tag{10.10}$$

by using $k \leq K \ll \lambda^{-1+12\kappa}$ as $\kappa < 1/13$.

We note that this method also gives a robust bound for the truncated E-value, since the truncation means that Lemma 3.3 is used only $k - 1$ times. Summarizing, we have proved

Lemma 10.2 *We assume* (7.25) *and* $\kappa < 1/13$. *Then*

$$\sup_{\sigma \in \mathfrak{S}_k} \sup_{\mathbf{u}} E(\sigma, \mathbf{u}) \leq C |\log \lambda|^2 \tag{10.11}$$

$$\sup_{\sigma \in \mathfrak{S}_k} \sup_{\mathbf{u}} E_*(\sigma, \mathbf{u}) \leq C\lambda^2 |\log \lambda|^2 \, . \qquad \square \tag{10.12}$$

10.2 Choice of the integration variables

Before we start the actual proof of Lemma 10.1, we explain the main idea. We use the combination of Methods I and II. We will assume in the sequel that $\sigma \neq id$. The lemma for the trivial $\sigma = id$ case has been proven in (10.10).

Note that each factor in the integrand in (10.1) is almost singular on a set of codimension one of the form $\{\alpha = \text{Re } \omega(p_j)\}$ or $\{\beta = \text{Re } \omega([M\mathbf{p} + \tilde{\mathbf{u}}]_j)\}$ in the high dimensional space of integration, $(\mathbb{R}^d)^{k+1} \times \mathbb{R}^2$. The singularities are regularized by η and Im ω and the two regularizations always have the same sign. The matrix M may enhance the strength of these singularities by forcing these "almost singularity" sets to overlap. For example, in the ladder diagram with $\tilde{\mathbf{u}} \equiv 0$, we have $\tilde{\mathbf{p}} \equiv \mathbf{p}$, hence the singularity sets are pairwise identical if $\alpha = \beta$. Therefore singularities of quadratic type necessarily occur. It is expected that this is the only mechanism that creates relevant overlaps of singularities. Hence, ideally, one would integrate out the ladder momenta with the precise bound

(3.15). This would remove all denominators with indices $j \in I_\ell$ and the remaining integral should be of $O(\lambda^{2(k-\ell)})$ if $\sigma \neq id$ (and for $\sigma = id$ it is $O(1)$) with possible logarithmic corrections.

However, the singularity sets of the remaining denominators may still intersect on higher codimensional manifolds. These enhancements of singularities are expected to have contributions of lower order, but their estimate is not easy. Note that every integration variable p_j may appear in many denominators. This interdependence renders the effective L^1-estimate of each integral practically impossible. A Schwarz inequality (Method II) can remove all correlations between denominators, but the resulting L^2-estimate is of the same order as the main (ladder) term and we would not gain anything from the higher degree of the permutation σ.

The idea is to estimate many, but not all β-denominators in (10.1) in the trivial way (10.4). These denominators are chosen in such a way that the remaining ones can be successively integrated out without ever computing an integrand with more than two denominators. This choice is given by an algorithmic procedure that we describe now.

We recall Definition 8.3 on partitioning the set of rows of $M = M(\sigma)$. Let $p := |I_p|$ be the number of peaks. We will derive new matrices from M by removing certain rows and entries. However, for convenience, we shall keep the original labelling of the rows (even after removing some of them in between).

Step 1. Construct a $(k+1-p) \times (k+1)$ matrix M_1 by removing all rows $i \in I_p$ from M. Note that M_1 has no empty column and it is a tower matrix. Let $I_1 := I \setminus I_p$ denote the set of row indices of M_1.

Step 2. For each row $h \in I_1$ in M_1 we define a column index $c(h)$ such that the bottom nonzero entry of this column in M_1 is in the row h. If $(\sigma^{-1}(h), h)$ was a slope then we have a unique choice, if it was a valley then both $\sigma^{-1}(h)$ and $\sigma^{-1}(h)+1$ are possible; we choose one of them arbitrarily if $h \notin I_\ell$. If $h \in I_v \cap I_\ell$, i.e., $|\tilde\sigma^{-1}(h) - \tilde\sigma^{-1}(h-1)| = 1$, then choose $c(h) := \sigma^{-1}(h)$ if $\tilde\sigma^{-1}(h) - \tilde\sigma^{-1}(h-1) = 1$, and $c(h) := \sigma^{-1}(h)+1$ otherwise. The column indices $\{c(h) \ : \ h \in I_v \cup I_s\}$ are called the pivot columns; the corresponding p_j momenta will be used for integration.

The matrix M_2 is constructed from M_1 by zeroing the bottom nonzero entry of each pivot column. The set of row-indices of M_2 is I_1, the same as for M_1. Let $r = rank(M_2)$. Consider those columns that became empty in M_2. In these columns there was only a single nonzero entry in M_1, so these are called *single* columns, their number is denoted by s. Note that these are not only the columns $c(h)$ of the ladder rows $h \in I_\ell$ in M, because the removal of the I_p rows can result in additional single columns.

Step 3. Construct a full row-rank matrix M_3 of dimension $r \times (k+1)$ from M_2 by selecting r linearly independent rows. Let $I_3 \subset I_1$ be the set of indices of the rows in M_3. These rows will determine those β-denominators that are *not* estimated trivially by (10.4).

It trivially follows from the definition that M_1, M_2, M_3 are tower matrices. Below we show the matrices M_1, M_2 and M_3 derived from the matrix $M(\sigma)$ in (8.10). The boxed entries are the pivot elements $(h, c(h))$, $h \in I_1$ that are removed in M_2. In this example, $k = 7$, $p = 1$, $\ell = 5$, $s = 5$,

$r = 2$ and the final index set $I_3 = \{3, 4\}$.

$$M_1 := \begin{pmatrix} \boxed{1} & & & & & & & \\ 1 & & & -1 & \boxed{1} & & & \\ 1 & & & \boxed{\text{-1}} & & 1 & & \\ 1 & & \boxed{\text{-1}} & & & 1 & & \\ 1 & \boxed{\text{-1}} & & & & 1 & & \\ & & & & & \boxed{1} & & \\ & & & & & & \boxed{1} & \end{pmatrix} \begin{matrix} 1 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix}$$

$$M_2 := \begin{pmatrix} 0 & & & & & & \\ 1 & & -1 & 0 & & \\ 1 & & 0 & & 1 & \\ 1 & 0 & & & 1 & \\ 1 & 0 & & & 1 & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} \begin{matrix} 1 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} \qquad M_3 := \begin{pmatrix} 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{matrix} 3 \\ 4 \end{matrix}$$

A few inequalities follow from this construction.

Lemma 10.3 *We have*

$$r \geq (k + 1 - p) - s \,, \tag{10.13}$$

$$s \leq \ell + p \,, \tag{10.14}$$

$$r \geq v - 1 = p \,. \tag{10.15}$$

From the inequalities (10.13)–(10.15), it follows that

$$r \geq \frac{1}{3}(k + 1 - \ell) \,. \tag{10.16}$$

Proof. Since the bottoms of the towers in the pivot columns all are located in different rows in M_1, M_1 has full rank and thus $rank(M_1) = k + 1 - p$. Furthermore, the towers in the nonempty columns of M_2 all become shorter exactly by one element, so their bottoms are still located in different rows, thus these columns are independent. This proves (10.13).

For the proof of (10.14), we claim that, apart from the ladders, (i.e., $h \in I_\ell$ whose pivot column $c(h)$ is single in both M and M_1), each peak can create at most one extra single. Granting this claim, we have proved (10.14). The last claim is obvious; a proof is given in the following.

Creating a single in M_1 that was non-ladder in M requires removing the second to the bottom nonzero element of that column, i.e. it requires that there be a peak $(\sigma^{-1}(p), p)$, $p \in I_p$ right above of the single that sits in row with index $p + 1$. Because M is obtained from a permutation, the removal of the row right above a peak removes the second to the bottom nonzero entry of two, one or zero columns, depending whether $(\sigma^{-1}(p+1), p+1)$ is a valley, slope or peak, respectively. However, for each valley only one wall can become single since only one of them was chosen as a pivot column. Therefore every peak can increase the number of singles only by at most one.

To prove (10.15), notice that the non-pivot columns of the valleys (except the last valley $(k + 1, k+1)$) all have different height in M_2. The relation $v-1 = p$ follows from the fact that $(k+1, k+1)$ was defined to be a valley, but $(0, 0)$ is not a peak. This concludes the proof. □

10.3 Integration procedure

Our goal is to estimate (10.1) for $M = M(\sigma)$ when $\sigma \neq id$. We eliminate the β-denominators one by one either by a trivial L^∞-estimate or by integrating out a well-chosen p_j variable. Unfortunately, new denominators with point-like singularities arise along this way. This makes a generalization of (10.1) necessary.

We start with defining

$$\|\!|q|\!\| := \eta + \min\{|q|, 1\}, \qquad q \in \mathbb{R}^d \,. \tag{10.17}$$

This is not a norm, but it satisfies the triangle inequality, $\|\!|p+q|\!\| \leq \|\!|p|\!\| + \|\!|q|\!\|$. For any index set $I' \subset I = \{1, 2, \dots, k+1\}$ with $1 \notin I'$, we define the function $\mathcal{U}_{I'}$ as the product of those potential terms, $|\widehat{B}|$, in (10.1) that depend only on momenta $\{p_j \,:\, j \in I'\}$. More precisely,

$$\mathcal{U}_{I'}(\mathbf{p}, \tilde{\mathbf{u}}) := \left| \prod_j^* \widehat{B}(p_{j+1} - p_j) \prod_j^* \widehat{B}\Big([M\mathbf{p} + \tilde{\mathbf{u}}]_{j+1} - [M\mathbf{p} + \tilde{\mathbf{u}}]_j\Big) \right| ,$$

where the star indicates a product on a restricted index set. The first product is taken over all indices j for which $j, j+1 \in I'$. The second product is taken over those j's, for which $M_{j+1,b} = M_{j,b}$ for all $b \notin I'$.

For any $|I'| \times (k+1)$ matrix M, any ν integer and any $\nu \times (k+1)$ matrix $\mathcal{E}$ we define

$$\begin{aligned} E(I', M, \mathcal{E}) := \lambda^{2k} \sup_{\tilde{\mathbf{u}}, \mathbf{v}} \int\!\!\int_{-Y}^{Y} \mathrm{d}\alpha \mathrm{d}\beta \int \frac{|\widehat{\psi}_0(p_1)||\widehat{\psi}_0(p_1 + \tilde{u}_1)| \mathrm{d}\mu(p_1)}{|\alpha - \overline{\omega}(p_1) - i\eta| \; |\beta - \omega(p_1 + u_1) + i\eta|} \\ \times \sup_{p_j \,:\, j \notin I'} \sup_{|\alpha|, |\beta| \leq Y} \int \prod_{j \in I'} \mathrm{d}\mu(p_j) \mathcal{U}_{I'}(\mathbf{p}, \tilde{\mathbf{u}}) \prod_{\mu=1}^{\nu} \frac{1}{\|\!|[\mathcal{E}\mathbf{p} + \mathbf{v}]_\mu|\!\|} \\ \times \left(\prod_{j \in I'} \frac{1}{|\alpha - \overline{\omega}(p_j) - i\eta|} \frac{1}{|\beta - \omega\big([M\mathbf{p} + \tilde{\mathbf{u}}]_j\big) + i\eta|} \right) . \end{aligned} \tag{10.18}$$

Here $\{v_\mu \,:\, 1 \leq \mu \leq \nu\}$ is an additional set of dummy momenta and $\mathcal{E}\mathbf{p}$ is defined analogously to (8.5). The $\mathrm{d}\alpha\, \mathrm{d}\beta$ integrations can be directly performed in (10.18), yielding a factor $C|\log\lambda|^2$ and eliminating the propagators carrying momenta p_1 and $\tilde{p}_1 = p_1 + \tilde{u}_1$. The index 1 therefore plays a special role. From the definition of $M(\sigma)$ it follows that either $1 \in I_p \cup I_\ell$, in particular $1 \notin I_3$.

For each $h \in I$ we will define an index set $I^{(h)}$ and matrices $M^{(h)}, \mathcal{E}^{(h)}$ such that $E(I^{(h)}, M^{(h)}, \mathcal{E}^{(h)})$ is the intermediate value of the integral (10.18) after integrating out h momentum variables p_i. We will bookkeep this integration by removing the corresponding h rows from M and we will say in short that these rows have been integrated out. On the matrix $\mathcal{E}^{(h)}$ we perform a Gaussian row operation.

Definition 10.4 *Let B be a matrix and let b_{ij} be a non-zero entry. We say that the j-th column is* **Gaussian-eliminated** *by b_{ij} if appropriate multiples of the i-th row is added to each rows to zero out all but the b_{ij} entry in the j-th column.*

Definition 10.5 *Let B be a matrix. We say that a matrix C is* **derived** *from B if it is obtained from B by applying the following operations arbitrary many times:*

(a) zeroing out some columns completely;

(b) removing some rows;

(c) Gaussian-eliminating a column by one of its nonzero element;

(d) swapping two rows.

The following Lemma is obvious:

Lemma 10.6 *Any matrix derived from a totally unimodular matrix is also totally unimodular.* □

Now we define $M^{(h)}$, $I^{(h)}$ and $\mathcal{E}^{(h)}$. For any $h \geq 1$, define $M^{(h)}$ to be the matrix $M = M(\sigma)$ with the first h rows, $i = 1, 2, \ldots, h$, removed. Let $I^{(1)} := I = \{1, 2, \ldots, k+1\}$ and for $h \geq 2$ let

$$I^{(h)} := I \setminus \{c(h') \ : \ h' \leq h, h' \in I_3 \cup I_\ell\}$$

be the index set of columns that have not been integrated out up to the h-th step.

For the definition of $\mathcal{E}^{(h)}$, we will need a sequence of auxiliary matrices, $\mathcal{F}^{(h)}$, of dimensions $r \times (k+1)$ that will be defined inductively below. The rows of $\mathcal{F}^{(h)}$ will be naturally indexed by I_3. By definition, throughout this section $\mathcal{E}^{(h)}$ will always denote the matrix consisting of the rows of $\mathcal{F}^{(h)}$ with index $i \in I_3$, $i \leq h$.

We define the initial matrix $\mathcal{F}^{(1)} := M_3$. The initial $\mathcal{E}^{(1)}$ matrix is then the trivial matrix with no rows at all recalling that $1 \notin I_3$. At every level h, there may be several $\mathcal{F}^{(h)}$ matrices (and therefore also several $\mathcal{E}^{(h)}$ matrices) according to a ramification in the integration procedure. But each matrix $\mathcal{F}^{(h)}$ at level h will have linearly independent rows and will be derived from M_3, in particular $\mathcal{F}^{(h)}$ will be totally unimodular. The same statements then obviously hold for each $\mathcal{E}^{(h)}$.

Suppose that a collection of matrices $\mathcal{F}^{(h-1)}$ at level $(h-1)$ has been defined with the properties above and we now define the matrices $\mathcal{F}^{(h)}$ at level $h \geq 2$.

Case 1: $h \in I \setminus (I_3 \cup I_\ell)$ (recall that I_ℓ denotes the set of ladder indices from Definition 8.3). In this case $I^{(h)} = I^{(h-1)}$ and we set $\mathcal{F}^{(h)} := \mathcal{F}^{(h-1)}$. We estimate the denominator $|\beta - \omega([M\mathbf{p} + \tilde{\mathbf{u}}]_h) + i\eta|$ trivially as in (10.4) and we obtain

$$E(I^{(h-1)}, M^{(h-1)}, \mathcal{E}^{(h-1)}) \leq \eta^{-1} E(I^{(h)}, M^{(h)}, \mathcal{E}^{(h)}) \ . \tag{10.19}$$

Case 2: $h \in I_\ell \setminus I_3$. Let $h, h+1, \ldots, h+\tau-1 \in I_\ell \setminus I_3$ be a maximal sequence of consecutive ladder indices, i.e. $h + \tau \notin I_\ell \setminus I_3$ with some $\tau \geq 1$. By the definition of the pivot indices, $c(h), c(h+1), \ldots, c(h+\tau-1)$ are consecutive numbers, which, for definiteness, are assumed to be increasing (the other case is similar). Let $c = c(h)$. From the fact that $h, \ldots, h+\tau-1$ are all consecutive ladder indices, it follows that $[M\mathbf{p}]_i = \mathbf{p}_i + w$, for all $c \leq i \leq c+\tau-1$ with a vector $w = w(\mathbf{p})$ that does not depend on $\{p_j \ : \ c \leq j \leq c+\tau-1\}$. Set $\mathcal{F}^{(k)} := \mathcal{F}^{(h-1)}$ for all $k = h, h+1, \ldots, h+\tau-1$. We claim

$$E(I^{(h-1)}, M^{(h-1)}, \mathcal{E}^{(h-1)}) \leq C\lambda^{-2\tau}\zeta^d \, E(I^{(h+\tau-1)}, M^{(h+\tau-1)}, \mathcal{E}^{(h+\tau-1)}) \ . \tag{10.20}$$

This inequality entails integrating out the variables $p_c, p_{c+1}, \dots, p_{c+\tau-1}$ (in this order) to remove the rows $h, h+1, \dots h+\tau-1$. This requires estimating the following integral

$$\mathcal{I} := \sup_{\alpha,\beta,w} \sup_{q,r} \int \prod_{j=c}^{c+\tau-1} \frac{\mathrm{d}\mu(p_j)}{|\alpha - \overline{\omega}(p_j) - i\eta||\beta - \omega(p_j + w + \tilde{u}_j) + i\eta|} \times \prod_{j=c}^{c+\tau-2} |\widehat{B}(p_{j+1} - p_j)||\widehat{B}(p_{j+1} + \tilde{u}_{j+1} - p_j - \tilde{u}_j)| \tag{10.21}$$

Due to the ladder structure it is easy to see that $\mathcal{U}_{I^{(h-1)}}$ still contains all the $\widehat{B}$ factors shown in (10.21).

After a Schwarz inequality we use the iterative integration scheme similar to (10.7)–(10.9) to estimate

$$\int \prod_{j=c}^{c+\tau-1} \frac{|\widehat{B}(p_{j+1} - p_j)|^2}{|\alpha - \overline{\omega}(p_j) - i\eta|^2} \, \mathrm{d}\mu(p_j) \leq \big[\lambda^{-2}(1 + C\lambda^{1-12\kappa})\big]^{\tau-1} \int \frac{1 + \lambda^{-1}|\alpha - \omega(p_c)|^{1/2}}{|\alpha - \overline{\omega}(p_c) - i\eta|^2} \mathrm{d}\mu(p_c) \, .$$

In the last integral we use (3.13) with $a = 0$ and $a = 1/2$ after inserting a partition of unity subordinated to unit cubes on the support of $\mathrm{d}\mu(p_c)$ (see (7.4)). Since $\tau \leq K \ll \lambda^{1-12\kappa}$, this gives

$$\mathcal{I} \leq C\lambda^{-2\tau}\zeta^d \, . \tag{10.22}$$

It would be possible to eliminate the additional ζ^d factor by keeping one more $\widehat{B}$ factor that connects the ladder to a non-ladder index in the definition of $\mathcal{U}$, but we will not need it since we do not aim at the optimal κ.

Case 3: $h \in I_3$. We define $A_h := A(h, \mathcal{F}^{(h-1)}) := \{a \ : \ \mathcal{F}^{(h-1)}_{a,c(h)} \neq 0\}$ to be the set of row indices with nonzero entry in the $c(h)$ column of $\mathcal{F}^{(h-1)}$. For any $a \in A_h$, the matrix $\mathcal{F}^{(h),a}$ is constructed from $\mathcal{F}^{(h-1)}$ as follows:

(i) we Gauss-eliminate the $c(h)$-th column of $\mathcal{F}^{(h-1)}$ by the entry $\mathcal{F}^{(h-1)}_{a,c(h)}$.

(ii) we remove the a-th row of $\mathcal{F}^{(h-1)}$.

Note that these two steps result in completely zeroing out the $c(h)$ pivot column that has been altered when constructing M_2.

From this construction and from Lemma 10.6 we immediately obtain

Lemma 10.7 *Every matrix $\mathcal{F}^{(h),a}$ has linearly independent rows and is derived from the matrix M_3. In particular, every $\mathcal{F}^{(h),a}$ is totally unimodular and thus the same holds for $\mathcal{E}^{(h),a}$, defined to consist of the rows of $\mathcal{F}^{(h),a}$ with index at most h.* □

We will also need the following observation. Recall that in the matrix M_1 the h-th entry was the lowest in the $c(h)$ column and this element was removed in M_2 and in M_3. Therefore the matrix M_3 contains no non-zero entry in the $c(h)$ column below the $(h-1)$-th row. The same is true for the $\mathcal{F}^{(h-1)}$ matrix as well, since it follows from the inductive definition that up to the $(h-1)$ steps only the first $(h-1)$ rows of the $\mathcal{F}$ matrices have been modified. Therefore all elements $a \in A_h$ satisfies $a < h$, in other words

$$\{a \,:\, \mathcal{F}^{(h-1)}_{a,c(h)} \neq 0\} \subset I_3 \cap \{1, 2, \dots, h-1\}. \tag{10.23}$$

The following lemma will be proved in the Appendix.

Lemma 10.8 *For $\lambda^3 \le \eta \le \lambda^2$ and for any non-empty index set A we have*

$$\sup_{|\alpha|,|\beta|\le Y} \int \frac{1}{|\alpha - \overline{\omega}(p) - i\eta|\ |\beta - \omega(r+p) + i\eta|} \prod_{a\in A} \frac{1}{\|\!| r_a + p |\!\|}\, \mathrm{d}\mu(p)$$

$$\le C\eta^{-1/2}\zeta^{d-3} |\log\eta|^2 \sum_{a\in A} \Big(\prod_{a'\in A,\, a'\neq a} \frac{1}{\|\!| r_a - r_{a'} |\!\|} \Big) \frac{1}{\|\!| r |\!\|}\,, \tag{10.24}$$

assuming $|r|, |r_a| \le \lambda^{-1}$. If $A = \emptyset$, $|r| \le \lambda^{-1}$, then we have the following improved bound:

$$\sup_{|\alpha|,|\beta|\le Y} \int \frac{\mathrm{d}\mu(p)}{|\alpha - \overline{\omega}(p) - i\eta|\ |\beta - \omega(r+p) + i\eta|} \le \frac{C\zeta^{d-3}|\log\eta|^2}{\|\!| r |\!\|}\,. \tag{10.25}$$

To estimate $E(I^{(h-1)}, M^{(h-1)}, \mathcal{E}^{(h-1)})$ for $h \in I_3$, we integrate out $p = p_{c(h)}$ in the definition of $E(I^{(h-1)}, M^{(h-1)}, \mathcal{E}^{(h-1)})$ (see (10.18)) by applying (10.24) with $A := A(h, \mathcal{F}^{(h-1)})$,

$$r_a := \frac{1}{\mathcal{E}^{(h-1)}_{a,c(h)}} \Big(\sum_{\substack{j=1 \\ j\neq c(h)}}^{k+1} \mathcal{E}^{(h-1)}_{aj} p_j + v_a \Big)\,, \quad r := \frac{1}{M_{h,c(h)}} \Big(\sum_{\substack{j=1 \\ j\neq c(h)}}^{k+1} M_{hj} p_j + \tilde{u}_h \Big)\,.$$

Note that any $a \in A$ satisfies $a \le h-1$ thanks to (10.23), hence the $\mathcal{E}^{(h-1)}_{aj}$ entries are well defined. The entries $M_{h,c(h)}$ and $\mathcal{E}^{(h-1)}_{a,c(h)}$ are ± 1 by definition. and we also used the symmetry of $\omega(p)$.

In the special case $c(h) = 1$, the propagator $|\alpha - \overline{\omega}(p_{c(h)}) - i\eta|^{-1}$ is not present (it has been removed by the $\mathrm{d}\alpha$-integration) and the measure $\mathrm{d}\mu(p_1)$ is changed to $|\widehat{\psi}_0(p_1)\widehat{\psi}_0(p_1+\tilde{u}_1)|\mathrm{d}\mu(p_1)$. However, the estimates (10.24), (10.25) still hold with these modifications on their left hand side, the constant may depend on $\widehat{\psi}_0$.

Focusing on the p dependence of the denominators, we notice that the denominators on the right hand side of (10.24), without the η regularizations, correspond exactly to the procedure how $\mathcal{F}^{(h),a}$ was obtained from $\mathcal{F}^{(h-1)}$. The dummy variables $\tilde{\mathbf{u}}, \mathbf{v}$ are redefined but their precise new form is irrelevant since we eventually take the supremum over all of them. This gives the estimate

$$E(I^{(h-1)}, M^{(h-1)}, \mathcal{E}^{(h-1)}) \le C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2 \sum_{a\in A_h} E(I^{(h)}, M^{(h)}, \mathcal{E}^{(h),a}) \tag{10.26}$$

for $h \in I_3$.

Using inequalities (10.19), (10.20), (10.26), we shall subsequently eliminate all denominators with β in (10.18) and reduce the integral to contain only α-denominators and point singularities. To integrate out the remaining p_j variables successively, we define, with a little abuse of notations,

$$E(J,\emptyset,\mathcal{E}) := \lambda^{2k} \sup_{\mathbf{v},\tilde{u}_1} \int \mathrm{d}\mu(p_1) |\widehat{\psi}_0(p_1)| |\widehat{\psi}_0(p_1+\tilde{u}_1)| \\ \times \sup_{\alpha} \int \prod_{\substack{j\in J\\ j\neq 1}} \frac{\mathrm{d}\mu(p_j)}{|\alpha - \overline{\omega}(p_j) - i\eta|} \prod_{\mu=1}^{\nu} \frac{1}{\|[\mathcal{E}\mathbf{p}+\mathbf{v}]_\mu\|} \, \mathcal{U}_J(\mathbf{p}) \tag{10.27}$$

for any $J \subset I^{(1)}$ and for any matrix $\mathcal{E}$ of dimensions $\nu \times |J|$. It may happen that after elimininting all β-denominators, no α denominator is left.

Lemma 10.9 *For any index set $J \subset I^{(1)}$ and any totally unimodular $\nu \times |J|$ matrix $\mathcal{E}$ with rank$(\mathcal{E}) = \nu$, we have*

$$E(J,\emptyset,\mathcal{E}) \le C\lambda^{2k} \big[C\zeta^{d-2} |\log\eta|\big]^{|J|} \nu! \tag{10.28}$$

Proof. We first estimate $\sup_{\mathbf{p}} |\mathcal{U}_J(\mathbf{p})| \le C^{|J|}$. Notice that the bound (10.28) is trivial if $|J| = 0$ or $J = \{1\}$. Otherwise, let $j_1, j_2, \ldots, j_{|J|}$ be the elements of $J \setminus \{1\}$. We integrate out $p_{j_1}, p_{j_2}, \ldots$ one by one. If the corresponding j_ℓ-th column of $\mathcal{E}$ is empty, then we pick up only a $C|\log\eta|$ factor. Otherwise, we use the inequality

$$\sup_{|\alpha|\le Y} \int \frac{1}{|\alpha - \overline{\omega}(p) - i\eta|} \prod_{a\in A} \frac{1}{\|r_a + p\|} \, \mathrm{d}\mu(p) \; \le C\zeta^{d-2} |\log\eta| \sum_{a\in A} \Big(\prod_{\substack{a'\in A\\ a'\neq a}} \frac{1}{\|r_a - r_{a'}\|} \Big) \,, \tag{10.29}$$

whose proof is similar to that of Lemma 10.8 and it follows directly from (A.1), (A.3) and (A.7). Notice that the remaining integrand is of the same form, but the index set J is reduced by one and the $\mathcal{E}$ matrix has changed. After the first step, with the choice of $p := p_{j_1}$, $A := \{a \,:\, \mathcal{E}_{aj_1} \neq 0\}$ and $r_a = (\mathcal{E}_{a,j_1})^{-1} \Big(\sum_{j\neq j_1} \mathcal{E}_{aj} p_j + v_a \Big)$ in (10.29), we obtain

$$E(J,\emptyset,\mathcal{E}) \le C\zeta^{d-2} |\log\eta| \sum_{a\in A} E(J',\emptyset,\mathcal{E}'^{,a}) \,,$$

where $J' := J \setminus \{j_1\}$ and $\mathcal{E}'^{,a}$ is obtained from $\mathcal{E}$ by following a construction analogous to step (i) and (ii) in Case 3 above. In other words, we first Gauss-eliminate the j_1 column by the element $\mathcal{E}_{a,j_1}$ and finally we remove the a-th row. As before, it follows from the construction that $\mathcal{E}'^{,a}$ is totally unimodular and of full rank. The number of terms in the summation is bounded by ν, the number of rows in $\mathcal{E}$. Clearly each $\mathcal{E}'^{,a}$ has $\nu - 1$ rows. We continue this procedure until $|J| = 0$ or $J = \{1\}$ and we arrive at (10.28). □

Armed with these, we can proceed to the proof of Theorem 8.4 with estimating (10.2). We start with the identity

$$E(M(\sigma)) = E(I^{(1)}, M^{(1)}, \mathcal{E}^{(1)}) \,. \tag{10.30}$$

We recall that $I^{(1)} = I$, $M^{(1)}$ is the matrix $M(\sigma)$ with the first row removed and $\mathcal{E}^{(1)}$ is the trivial matrix with no rows.

Introduce $b := |I_3 \cap I_\ell|$. We perform the integration procedure above. The $\mathrm{d}\alpha\mathrm{d}\beta$ integration picks up a factor $C|\log\eta|^2$. Inequality (10.19) in Case 1 is used $(k - r - \ell + b)$ times. Case 2 integrates out altogether $\ell - b$ ladder indices, i.e. the sum of the τ's in the consecutive inequalities (10.20) is $\ell - b$. Each time we apply Case 2, there is an index, $h+\tau$, that belongs to $I_\ell^c \cup I_3$. Therefore the additional factor $C\zeta^d$ appears at most $|I_\ell^c \cup I_3| \le k + 1 - \ell + r$ times (recall $|I_3| = r$). Finally, Case 3 occurs r times. The number of terms in (10.26) is the cardinality of A_h and it is bounded by k. Altogether we obtain

$$E(M(\sigma)) \le C|\log\eta|^2(\eta^{-1})^{k-r-\ell+b} \cdot \lambda^{-2(\ell-b)}(C\zeta^d)^{k+1-\ell+r} \cdot (C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2 k)^r$$
$$\times \sup\Big\{E(I^{(k+1)}, \emptyset, \mathcal{E}) \ : \ \mathcal{E} \text{ full row-rank and totally unimodular with } \nu \text{ rows}\Big\} . \tag{10.31}$$

Now we apply Lemma 10.9 with

$$J := I^{(k+1)} = I^{(1)} \setminus \{c(h) \ : \ h \in I_\ell \cup I_3\}$$

and with the matrix $\mathcal{E}$ that is obtained by removing all the zero columns from $\mathcal{E}^{(k+1)}$ that correspond to the integrated out p_i's, $i \in I^{(1)} \setminus J$. In particular, $\mathrm{rank}(\mathcal{E}) = \mathrm{rank}(\mathcal{E}^{(k+1)})$ is the number of rows in $\mathcal{E}$.

To complete the proof of Theorem 8.4, we combine (10.30), (10.31), (10.28) with the bounds

$$|J| \le k + 1 - \ell \, , \quad \text{and} \quad \nu \le k + 1 - \ell \, .$$

The latter inequality follows from the fact that ν is the number of rows in $\mathcal{E}$ that is equal to $|I \setminus (I_\ell \cup I_3)|$. We also estimate $\nu! \le |J|! \le (Ck)^{k+1-\ell}$ and we obtain

$$\begin{aligned} E(M(\sigma)) &\le C|\log\eta|^2\Big(C\lambda^{2-d\kappa-O(\delta)}\eta^{-1}k\Big)^{k+1-\ell}\Big(C\lambda^{2-2d\kappa-O(\delta)}\eta^{-1/2}k\Big)^r(\lambda^2\eta^{-1})^{b-r} \\ &\le C|\log\eta|^2\Big(C\lambda^{-(d+1)\kappa-O(\delta)}\Big)^{k+1-\ell}\Big(C\lambda^{1-(2d+\frac{3}{2})\kappa-O(\delta)}\Big)^r \, , \end{aligned}$$

where in the second inequality we used (7.25), $b \le r$ and that δ is sufficiently small. Using $r \ge \frac{1}{3}(k + 1 - \ell)$ from (10.16) and recalling that $\deg(\sigma) = k + 1 - \ell$, we obtain (10.2). □

A Proof of Lemma 10.8.

First we reduce the multiple point singularities to one point singularity using the pointwise estimate

$$\prod_{a\in A} \frac{1}{\|\!| p - r_a |\!\|} \le \sum_{a\in A} \frac{1}{\|\!| p - r_a |\!\|} \prod_{a'\in A, a'\ne a} \frac{1}{\|\!| r_a - r_{a'} |\!\|} \, , \tag{A.1}$$

that can be proven by a simple induction on $|A|$ using the triangle inequality for $\|\!|\cdot|\!\|$. Then the proof of (10.24) will follow from

$$\sup_{\alpha,\beta,r} \int \frac{\mathrm{d}\mu(p)}{|\alpha - \omega(p) + i\eta| \, |\beta - \omega(p+q) + i\eta|} \frac{1}{\|\!| p - r |\!\|} \le \frac{C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2}{\|\!| q |\!\|} \, . \tag{A.2}$$

We can replace $\omega(p)$ with $e(p)$ by using a straightforward resolvent expansion:

$$\frac{1}{|\alpha-\omega(p)+i\eta|}\le\frac{1}{|\alpha-\lambda^2\Theta(\alpha)-e(p)+i\eta|}\left[1+\frac{C\lambda^2|\alpha-e(p)|^{1/2}}{|\alpha-\omega(p)+i\eta|}\right]\le\frac{C}{|\widetilde{\alpha}-e(p)+i\eta|} \tag{A.3}$$

with $\widetilde{\alpha}=\alpha-\lambda^2\,\mathrm{Re}\,\Theta(\alpha)$. We used the boundedness and the Hölder continuity of Θ (3.5). Therefore the proof of Lemma 10.8 is reduced to

Lemma A.1 *For any* $|q|\le\lambda^{-1}$

$$I_1:=\int\frac{\mathrm{d}\mu(p)}{|\alpha-e(p)+i\eta|\,|\beta-e(p+q)+i\eta|}\le\frac{C\zeta^{d-3}|\log\eta|^2}{\|\!|q|\!\|} \tag{A.4}$$

$$I_2:=\int\frac{\mathrm{d}\mu(p)}{|\alpha-e(p)+i\eta|\,|\beta-e(p+q)+i\eta|}\,\frac{1}{\|\!|p-r|\!\|}\le\frac{C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2}{\|\!|q|\!\|} \tag{A.5}$$

uniformly in r,α,β.

Proof of Lemma A.1. The bound on I_1 follows from a direct calculation and $|q|\le\lambda^{-1}$

$$I_1\le\int_0^\zeta\frac{u^{d-1}\mathrm{d}u}{|\alpha-u^2/2+i\eta|}\int_{-1}^1\frac{\mathrm{d}c}{\big|2\beta-(u^2+q^2)-2|q|uc\big|+\eta}\le\frac{C\zeta^{d-3}|\log\eta|^2}{|q|}\,. \tag{A.6}$$

If $\eta\le|q|$, then we use $\|\!|q|\!\|\le|q|$ to obtain (A.4). If $|q|\le\eta$, then we use Schwarz inequality to separate the denominators and use (3.14) to conclude the proof of (A.4).

To prove (A.5), we first establish the following bound uniformly in α:

$$J:=\int\frac{\mathrm{d}\mu(p)}{|\alpha-e(p)+i\eta|}\,\frac{1}{\|\!|p-r|\!\|}\le C\zeta^{d-2}|\log\eta| \tag{A.7}$$

that follows by a direct calculation

$$J\le C\int_0^\zeta\frac{u^{d-1}\mathrm{d}u}{|\alpha-u^2/2+i\eta|}\left[1+\int_{-1}^1\frac{\mathrm{d}c}{\big|u^2+r^2-2|r|uc\,\big|^{1/2}}\right]$$

$$\le C(\zeta^{d-3}+\zeta^{d-2})\int_0^{\zeta^2}\frac{\mathrm{d}v}{|\alpha-v+i\eta|}\le C\zeta^{d-2}|\log\eta|$$

with $u=|p|$, $v=u^2/2$ using $\big|u^2+r^2-2|r|uc\,\big|\ge|u|^2|1-c^2|$ for $|c|\le1$.

We can assume that $|q|\ge\eta$, otherwise we can estimate the β-denominator in (A.5) trivially by η^{-1} and we can conclude with (A.7). We then distinguish two regimes. If $\|\!|p-r|\!\|\ge\eta^{1/2}$, then we estimate $\|\!|p-r|\!\|$ trivially and we use (A.4).

Now let $\|\!|p-r|\!\|\le\eta^{1/2}$. We split this regime to two subregimes, $|q|\le2|p|$ and $|q|\le2|p+q|$, the union of whose clearly cover all values of p.

In the regime, where $|q| \leq 2|p|$, we estimate the square root of the β-denominator trivially and use a Schwarz inequality to separate the remaining β denominator from the point singularity. The corresponding contribution can be estimated by

$$C\eta^{-1/2} \int \frac{\mathbf{1}(|q| \leq 2|p|)}{|\alpha - e(p) + i\eta|} \left[\frac{1}{|\beta - e(p+q) + i\eta|} + \frac{\mathbf{1}(\|\!|p-r|\!\| \leq \eta^{1/2})}{\|\!|p-r|\!\|^2} \right] \mathrm{d}\mu(p) \; .$$

The first term was already estimated in (A.4). The second term is bounded by the co-area formula by

$$C\eta^{-1/2} \int_{(|q|/2)^2}^{\zeta^2} \frac{J_a \, \mathrm{d}a}{|\alpha - a + i\eta| \, |a|^{1/2}} \; , \qquad \text{with} \quad J_a := \int_{\Sigma_a} \frac{\mathbf{1}(\|\!|p-r|\!\| \leq \eta^{1/2})}{\|\!|p-r|\!\|^2} \, \mathrm{d}\nu(p) \; ,$$

where $\Sigma_a := \{p \; : \; e(p) = a\}$ and $\mathrm{d}\nu(p)$ being the surface measure. Clearly $J_a \leq |\log \eta|$. and we obtain the estimate $C\eta^{-1/2} |\log \eta|^2 / |q|$.

In the regime where $|q| \leq 2|p+q|$, we shift $p \to p+q$, $r \to r-q$ and interchange the role of the α and β denominators in the above proof. This completes the proof of Lemma A.1. $\square$.

References

[1] M. Aizenman and S. Molchanov, *Localization at large disorder and at extreme energies: an elementary derivation*, Commun. Math. Phys. **157**, 245–278 (1993)

[2] M. Aizenman, R Sims, S. Warzel, *Absolutely continuous spectra of quantum tree graphs with weak disorder.* Preprint. http://xxx.lanl.gov/abs/math-ph/0504039

[3] P. Anderson, *Absences of diffusion in certain random lattices*, Phys. Rev. **109**, 1492–1505 (1958)

[4] C. Boldrighrini, L. Bunimovich and Y. Sinai: *On the Boltzmann equation for the Lorenz gas*, J. Statis. Phys. **32** no. 3, 477-501 (1983)

[5] R. Brown, Philosophical Magazine N. S. **4** 161–173 (1828), and **6** 161–166 (1829)

[6] D. C. Brydges, J. Dimock, T. R. Hurd, *The short distance behaviour of* $(\phi^4)_3$, Comm. Math. Phys. **172** (1995) 143–186

[7] D. C. Brydges, H.-T. Yau, *Grad* ϕ *perturbations of massless Gaussian fields*, Comm. Math. Phys. **129** (1990) 351–392

[8] J. Bourgain, *Random lattice Schrödinger operators with decaying potential: some higher dimensional phenomena.* Lecture Notes in Mathematics, Vol. 1807, 70-99 (2003).

[9] L. Bunimovich, Y. Sinai: *Statistical properties of Lorentz gas with periodic configuration of scatterers.* Commun. Math. Phys. **78** no. 4, 479–497 (1980/81),

[10] T. Chen, *Localization Lengths and Boltzmann Limit for the Anderson Model at Small Disorders in Dimension 3.* Preprint http://xxx.lanl.gov/abs/math-ph/0305051

[11] S. A. Denisov, *Absolutely continuous spectrum of multidimensional Schrödinger operator.* Int. Math. Res. Not. **2004** no. 74, 3963–3982.

[12] D. Dürr, S. Goldstein, J. Lebowitz: *A mechanical model of Brownian motion.* Commun. Math. Phys. **78** (1980/81) no. 4, 507-530.

[13] D. Dürr, S. Goldstein, J. Lebowitz: *Asymptotic motion of a classical particle in random potential in two dimensions: Landau model*, Commun. Math. Phys. **113** (1987) no 2. 209-230.

[14] A. Einstein, *Zur Theorie der Brownschen Bewegung*, Ann. der Physik, **19** 180 (1906).

[15] L. Erdős, *Linear Boltzmann equation as the long time dynamics of an electron weakly coupled to a phonon field.* J. Stat. Phys., **107**, 1043-1128 (2002)

[16] L. Erdős and H.-T. Yau, *Linear Boltzmann equation as the weak coupling limit of the random Schrödinger equation*, Commun. Pure Appl. Math. **LIII**, 667-735, (2000).

[17] L. Erdős, M. Salmhofer and H.-T. Yau, *Towards the quantum Brownian motion.* Submitted to the QMath-9 Conference Proceedings, Giens, 2004.

[18] L. Erdős, M. Salmhofer and H.-T. Yau, *Quantum diffusion of random Schrödinger evolution in scaling limit.* Preprint. http://xxx.lanl.gov/abs/math-ph/0502025

[19] L. Erdős, M. Salmhofer and H.-T. Yau, *Quantum diffusion of the random Schrödinger evolution in the scaling limit II. The recollision diagrams.* Submitted to Commun. Math. Phys.

[20] J. Feldman, H. Knörrer, E. Trubowitz, *A Representation for Fermionic Correlation Functions*, Commun. Math. Phys. **195** :465–493 (1998)

[21] J. Feldman, H. Knörrer, E. Trubowitz, *Convergence of Perturbation Expansions in Fermionic Models, Part 2: Overlapping Loops*, Comm. Math. Phys. **247**, 243–319 (2004)

[22] J. Feldman, J. Magnen, V. Rivasseau, R. Sénéor, Comm. Math. Phys. **98** (1985) 273, **100** (1985) 23, **109** (1987) 437

[23] J. Feldman, J. Magnen, V. Rivasseau, and R. Sénéor, *A Renormalizable Field Theory: The Massive Gross-Neveu Model in Two Dimensions*, Comm. Math. Phys. **103** (1986) 67–103

[24] R. Froese, D. Hasler, W. Spitzer, *Absolutely continuous spectrum for the Anderson model on a tree: a geometric proof of Klein's theorem.* Preprint http://xxx.lanl.gov/math-ph/051150

[25] J. Fröhlich and T. Spencer, *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Commun. Math. Phys. **88**, 151–184 (1983)

[26] J. Feldman, M. Salmhofer, E. Trubowitz, *Perturbation Theory around Non-nested Fermi Surfaces I. Keeping the Fermi Surface Fixed*, Journal of Statistical Physics **84** (1996) 1209–1336

[27] J. Feldman, M. Salmhofer, E. Trubowitz, *Regularity of Interacting Nonspherical Fermi Surfaces: The Full Self–Energy*, Comm. Pure Appl. Math. **52**, 273–324 (1999)

[28] T. G. Ho, L. J. Landau and A. J. Wilkins: *On the weak coupling limit for a Fermi gas in a random potential.* Rev. Math. Phys. **5** (1993), no. 2, 209-298.

[29] G. Gallavotti, *Rigorous theory of the Boltzmann equation in the Lorentz model.* Nota interna n. 358, Physics Department, Universita' di Roma, pp.1–9 (1972) $mp\_arc@math.utexas.edu$, # 93-304.

[30] K. Gawedzki, A. Kupiainen, *Massless Lattice ϕ^4_4 Theory: Rigorous Control of a Renormalizable Asymptotically Free Model*, Comm. Math. Phys. **99** 197–252 (1985)

[31] K. Gawedzki and A. Kupiainen, *Gross-Neveu Model Through Convergent Perturbation Expansions*. Commun. Math. Phys. **102** 1–30 (1985)

[32] I. Ya. Goldsheid, S. A. Molchanov and L. A. Pastur, *A pure point spectrum of the one dimensional Schrödinger operator.* Funct. Anal. Appl. **11**, 1–10 (1997)

[33] H. Kesten, G. Papanicolaou: *A limit theorem for stochastic acceleration.* Comm. Math. Phys. **78** 19-63. (1980/81)

[34] A. Klein, *Absolutely continuous spectrum in the Anderson model on the Bethe lattice*, Math. Res. Lett. **1**, 399–407 (1994)

[35] W. Kirsch, F. Martinelli: *On the essential selfadjointness of stochastic Schrödinger operators*, Duke Math. J. **50**, no.4., 1255–1260 (1983)

[36] T. Komorowski, L. Ryzhik: *Diffusion in a weakly random Hamiltonian flow.* Preprint, http://xxx.lanl.gov/math-ph/0505082

[37] O. Lanford, *On the derivation of the Boltzmann equation.* Astérisque **40**, 117-137 (1976)

[38] P. A. Lee, T. V. Ramakrishnan, *Disordered electronic systems.* Rev. Mod. Phys. **57**, 287–337 (1985)

[39] J. Lukkarinen and H. Spohn, *Kinetic limit for wave propagation in a random medium.* Preprint. http://xxx.lanl.gov/math-ph/0505075

[40] J. B. Perrin, *Mouvement brownien et réalité moléculaire.* Annales de chimie et de physiqe, VIII 18, 5–114 (1909).

[41] I. Rodnianski, W. Schlag, *Classical and quantum scattering for a class of long range random potentials.* Int. Math. Res. Not. **5** 243–300 (2003).

[42] M. Salmhofer, C. Wieczerkowski, *Positivity and Convergence in Fermionic Field Theory*, J. Stat. Phys. **99**, 557–586 (2000)

[43] E. Seiler, Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics. Lecture Notes in Physics, **159.** Springer, (1982).

[44] W. Schlag, C. Shubin, T. Wolff, *Frequency concentration and location lengths for the Anderson model at small disorders.* J. Anal. Math. **88** (2002), 173–220.

[45] H. Spohn: *Derivation of the transport equation for electrons moving through random impurities.* J. Statist. Phys.**17** (1977), no. 6., 385-412.

[46] H. Spohn: *The Lorentz process converges to a random flight process.* Commun. Math. Phys. **60** (1978), no. 3, 277-290.

[47] D. Vollhardt, P. Wölfle, *Diagrammatic, self-consistent treatment of the Anderson localization problem in $d \leq 2$ dimensions.* Phys. Rev. B **22**, 4666-4679 (1980)