THE CRITICAL 2D STOCHASTIC HEAT FLOW

FRANCESCO CARAVENNA, RONGFENG SUN, AND NIKOS ZYGOURAS

Abstract. We consider directed polymers in random environment in the critical dimension $d = 2$, focusing on the intermediate disorder regime when the model undergoes a phase transition. We prove that, at criticality, the diffusively rescaled random field of partition functions has a *unique scaling limit*: a universal process of random measures on $\mathbb{R}^2$ with logarithmic correlations, which we call the *Critical 2d Stochastic Heat Flow*. It is the natural candidate for the long sought solution of the critical 2d Stochastic Heat Equation with multiplicative space-time white noise.

Contents

1. Introduction and main results

1.1. Overview. The model of *directed polymer in random environment* (DPRE) is by now a fundamental model in statistical physics and probability theory. It is one of the simplest and yet most challenging models for disordered systems, where the effect of *disorder* — which is synonymous with random environment — can be investigated. Originally introduced by Huse and Henley [HH85] in the physics literature to study interfaces of the Ising model with random impurities, over the years, DPRE has become an object of mathematical interest and lies at the heart of two areas of intense research in recent years. On the one hand, it is one of the canonical examples in the Kardar-Parisi-Zhang (KPZ) universality class of interface growth models, which has witnessed tremendous progress over the last two decades in spatial dimension $d = 1$ (see e.g. the surveys [QS15, Cor12, Cor16]); on the other hand,

Date: January 31, 2023.

2010 *Mathematics Subject Classification.* Primary: 82B44; Secondary: 35R60, 60H15, 82D60.

Key words and phrases. Directed Polymer in Random Environment, Stochastic Heat Equation, KPZ Equation, Coarse-Graining, Renormalization, Lindeberg Principle.

it provides a discretisation of the Stochastic Heat Equation (SHE) and (via the Cole-Hopf transformation) of the Kardar-Parisi-Zhang (KPZ) equation, for which a robust solution theory in $d = 1$ has been developed only recently in the larger context of singular stochastic partial differential equations (SPDE) [H13, H14, GIP15, Kup14, GJ14].

Our goal in this paper is to consider DPRE in the *critical spatial dimension* $d = 2$, for which much remains unknown. Our main result shows that, in a *critical window* for the disorder strength, the family of partition functions of DPRE converges to a universal limit, which can be interpreted as the solution of the (classically ill-defined) 2-dimensional SHE. This is the first example of a singular SPDE for which a solution has been constructed in the critical dimension and for critical disorder strength.

In the remainder of the introduction, we first recall the definition of DPRE, its basic properties, and the works leading up to our current result. We then present our main results and discuss their connections with singular SPDEs and related research.

1.2. The model. The first ingredient in the definition of DPRE is a simple symmetric random walk $(S = (S_n)_{n\geqslant 0}, \mathrm{P})$ on $\mathbb{Z}^d$, started at $S_0 = 0$. To specify a different starting time m and position z, we will write $\mathrm{P}(\,\cdot\,|S_m = z)$. The second ingredient is the disorder or random environment, encoded by a family of i.i.d. random variables $(\omega = (\omega(n,z))_{n\in\mathbb{N},z\in\mathbb{Z}^d}, \mathbb{P})$ with zero mean, unit variance and some finite exponential moments:

$$\begin{gathered} \mathbb{E}[\omega] = 0\,, \qquad \mathbb{E}[\omega^2] = 1\,, \\ \exists\,\beta_0 > 0 \quad \text{such that} \qquad \lambda(\beta) := \log \mathbb{E}[e^{\beta\omega}] < \infty \qquad \forall \beta \in [0,\beta_0]\,. \end{gathered} \tag{1.1}$$

Given $N \in \mathbb{N}$, $\beta > 0$, and a realization of ω, the polymer measure of length $N \in \mathbb{N}$ and disorder strength (inverse temperature) β in the random environment ω is given by

$$\mathrm{dP}_N^{\beta,\omega}(S \,|\, S_0 = z) := \frac{1}{Z_N^{\beta,\omega}(z)} e^{\sum_{n=1}^{N-1}\{\beta\omega(n,S_n)-\lambda(\beta)\}}\,\mathrm{dP}(S \,|\, S_0 = z)\,, \tag{1.2}$$

where

$$Z_N^{\beta,\omega}(z) = \mathrm{E}\left[e^{\sum_{n=1}^{N-1}\{\beta\omega(n,S_n)-\lambda(\beta)\}} \,\middle|\, S_0 = z \right] \tag{1.3}$$

is the *partition function*. Note that $\lambda(\beta)$ in the exponent ensures that $\mathbb{E}[Z_N^{\beta,\omega}(z)] = 1$.

In the mathematical literature, DPRE was first studied by Imbrie and Spencer [IS88]. There have been many results since then, although many fundamental questions remain open. We briefly recall what is known and refer to the recent monograph by Comets [Com17] for more details and references.

DPRE exhibits a phase transition between a *weak disorder* phase and a *strong disorder* phase. Using the martingale structure of the partition functions $(Z_N^{\beta,\omega}(0))_{N\in\mathbb{N}}$, first identified by Bolthausen in [Bo89], DPRE is said to be in the weak disorder (or strong disorder) phase if the martingale converges almost surely to a positive limit (or to 0). It was later shown in [CY06] that there is a critical value $\beta_c \geqslant 0$ such that strong disorder holds for $\beta > \beta_c$ and weak disorder holds for $0 \leqslant \beta < \beta_c$, where $\beta_c \in (0,\infty)$ for $d \geqslant 3$ [IS88, Bo89], and $\beta_c = 0$ for $d = 1, 2$ [CH02, CSY03] (see also [L10, BL17, N19]).

In the weak disorder phase, a series of works culminating in [CY06] showed that the random walk under the polymer measure converges to a Brownian motion under diffusive scaling of space and time, as if the disorder is not present. In the strong disorder phase, it is believed that under the polymer measure, the path should be super-diffusive, but this has only been proved for special integrable models in dimension $d = 1$, see [J00, CH16]. Even less

is known in $d \geqslant 2$ due to the lack of integrable models within the same universality class. We mention that the strong disorder phase can alternatively be characterised by the fact that two polymer paths sampled independently in the same random environment have positive overlap, see [CH02, CSY03, V07] and the more recent results [Cha19, BC20a, BC20b, Ba21].

1.3. The case $d = 2$. Henceforth, we will focus on dimension $d = 2$. Surprisingly, even though $\beta_c = 0$, there is still a weak to strong disorder transition, which was identified in [CSZ17b]. More precisely, if we choose $\beta = \beta_N = \hat{\beta}/\sqrt{\log N}$, which is called an *intermediate disorder regime*, then it was shown in [CSZ17b] that below the critical point $\hat{\beta}_c = \sqrt{\pi}$, the partition function $Z_N^{\beta_N,\omega}(0)$ converges in distribution to a log-normal random variable, which is strictly positive, while at and above $\hat{\beta}_c$, it converges to 0 (such a transition does not occur in $d = 1$). This raises many interesting questions about the 2-dimensional DPRE.

There are two main perspectives in the study of the partition functions of DPRE. One is to investigate the fluctuation of a single log-partition function $\log Z_N^{\beta,\omega}(0)$ as $N \to \infty$. In $d = 1$, this is conjectured to converge, under suitable rescaling, to the universal Tracy-Widom distribution whenever $\beta > 0$. Similar universal fluctuations are expected to arise in $d \geqslant 2$ when $\beta > \beta_c$, although only numerical results are available so far [HH12, HH13]. In $d = 2$ and in the intermediate disorder regime $\beta_N = \hat{\beta}/\sqrt{\log N}$ with a sub-critical interaction strength $\hat{\beta} < \hat{\beta}_c$, [CSZ17b] showed that $\log Z_N^{\beta,\omega}(0)$ converges to a universal normal limit independent of the law of ω. The super-critical case $\hat{\beta} \geqslant \hat{\beta}_c$ remains a difficult challenge.

Another perspective, which we take in this paper, is to study the *diffusively rescaled field of partition functions* indexed by all starting points in space-time:

$$\big(\mathcal{U}_N(t,x) := Z_{Nt}^{\beta_N,\omega}(\sqrt{N}x)\big)_{t>0,\,x\in\mathbb{R}^2}\,, \tag{1.4}$$

as well as the diffusively rescaled field of log-partition functions:

$$\big(\mathcal{H}_N(t,x) := \log Z_{Nt}^{\beta_N,\omega}(\sqrt{N}x)\big)_{t>0,\,x\in\mathbb{R}^2}\,. \tag{1.5}$$

The fields $\mathcal{U}_N$ and $\mathcal{H}_N$ provide natural discretizations of the solutions of the two-dimensional *Stochastic Heat Equation (SHE)* and *Kardar-Parisi-Zhang equation (KPZ)* respectively:

$$\partial_t u = \frac{1}{2}\Delta u + \beta\,\dot{W}\,u, \tag{1.6}$$

$$\partial_t h = \frac{1}{2}\Delta h + \frac{1}{2}|\nabla h|^2 + \beta\,\dot{W}\,, \tag{1.7}$$

where $\dot{W} = \dot{W}(t,x)$ denotes space-time white noise. These stochastic PDEs are singular and ill-posed: even the recent breakthrough solution theories of regularity structures [H13, H14] and paracontrolled distributions [GIP15, GP17] only apply in $d = 1$ but not in the critical dimension $d = 2$. Therefore, if $\mathcal{U}_N$ and $\mathcal{H}_N$ admit non-trivial limits, then these limits are natural candidates for the long-sought solutions of SHE and KPZ in $d = 2$.

The study of the random field $\mathcal{U}_N$ was initiated in [CSZ17b], which showed that in the subcritical regime $\hat{\beta} < \hat{\beta}_c$, the centered and rescaled random field $\sqrt{\log N}\big(\mathcal{U}_N(t,x) - 1\big)$ converges to the solution of the so-called Edwards-Wilkinson equation, which is a Gaussian free field at each time t. The study of the random field $\mathcal{H}_N$ was first carried out in [CD20],[†] which showed that $\sqrt{\log N}\big(\mathcal{H}_N(t,x) - \mathbb{E}[\mathcal{H}_N(t,x)]\big)$ is tight in N as a family of distribution-valued random variables for $\hat{\beta}$ sufficiently small; shortly after, [CSZ20] proved convergence to

[†]More precisely, [CD20] and [Gu20] both study the analogue of $\mathcal{H}_N$ defined by mollifying the noise $\dot{W}$ in (1.7) instead of discretizing space and time, while [CSZ17b] considered both types of regularizations.

the solution of the same Edwards-Wilkinson equation as for $\mathcal{U}_N$ for all $\hat{\beta} < \hat{\beta}_c$ (simultaneously, the same result was proved in [Gu20] for $\hat{\beta}$ sufficiently small).

In the much more interesting and delicate critical regime $\hat{\beta} = \hat{\beta}_c$ — there is in fact a critical window of width $O(1/\log N)$ around $\hat{\beta}_c$, see (1.11) below — the random field $\mathcal{U}_N(t,x)$ no longer needs any centering and rescaling. Its limiting correlation structure was first identified in [BC98] through a different regularisation of the $2d$ SHE (1.6) (mollifying the noise $\dot{W}$ instead of discretizing space and time). In [CSZ19b], the third moment of the averaged random field $\mathcal{U}_N(t,\varphi) := \int \mathcal{U}_N(t,x)\,\varphi(x)\,\mathrm{d}x$, for test functions φ, was computed and shown to converge to a finite limit as $N \to \infty$, which implies that all subsequential limits of $\mathcal{U}_N$ have the same correlation structure identified in [BC98] (tightness is trivial since $\mathbb{E}[\mathcal{U}_N] \equiv 1$). Subsequently, [GQT21] identified the limit of all moments of $\mathcal{U}_N(t,\varphi)$ (see also the more recent work [Che21]). However, the *uniqueness of the limit* of $\mathcal{U}_N$ remained elusive and challenging, because the limiting moments identified in [GQT21] and [Che21] grow too fast to uniquely determine the law of the random field.

Our main result settles this question and shows that, in the critical window around $\hat{\beta} = \hat{\beta}_c$, the random field $\mathcal{U}_N$ indeed converges to a *unique universal limit*, which naturally provides a notion of solution of the $2d$ SHE (1.6) for disorder strength β in the critical window. Therefore, we name it the *Critical* $2d$ *Stochastic Heat Flow*.

1.4. Main results. To formulate our main results, we generalize the partition functions in (1.3) by introducing a point-to-point version, where both the starting and ending positions of the random walk are fixed: for $M \leqslant N \in \mathbb{N}_0 = \{0,1,2,\ldots\}$ and $w, z \in \mathbb{Z}^2$ we set

$$Z^{\beta,\omega}_{M,N}(w,z) := \mathrm{E}\left[e^{\sum_{n=M+1}^{N-1}\{\beta\omega(n,S_n)-\lambda(\beta)\}}\, \mathbb{1}_{S_N=z} \,\middle|\, S_M = w \right], \tag{1.8}$$

with the convention $\sum_{n=M+1}^{N-1}\{\ldots\} := 0$ for $N \leqslant M+1$.

To deal with parity issues, for $x \in \mathbb{R}^2$ we denote by $[\![x]\!]$ the closest point $z \in \mathbb{Z}^2_{\mathrm{even}} := \{(z_1,z_2) \in \mathbb{Z}^2 : z_1 + z_2 \text{ even}\}$; for $s \in \mathbb{R}$ we define the even approximation $[\![s]\!] := 2\lfloor s/2 \rfloor \in \mathbb{Z}_{\mathrm{even}} := 2\mathbb{Z}$. We then introduce the process of *diffusively rescaled partition functions*:[†]

$$\mathcal{Z}^{\beta_N}_N = \left(\mathcal{Z}^{\beta_N}_{N;\,s,t}(\mathrm{d}x,\mathrm{d}y) := \frac{N}{4}\, Z^{\beta_N,\omega}_{[\![Ns]\!],[\![Nt]\!]}([\![\sqrt{N}x]\!],[\![\sqrt{N}y]\!])\,\mathrm{d}x\,\mathrm{d}y \right)_{0\leqslant s\leqslant t<\infty} \tag{1.9}$$

where $\mathrm{d}x\,\mathrm{d}y$ denotes the Lebesgue measure on $\mathbb{R}^2 \times \mathbb{R}^2$, and β_N will be defined shortly.

We regard $\mathcal{Z}^{\beta_N}_{N;\,s,t}(\mathrm{d}x,\mathrm{d}y)$ as a random measure on $\mathbb{R}^2 \times \mathbb{R}^2$, where we equip the space of locally finite measures on $\mathbb{R}^2 \times \mathbb{R}^2$ with the topology of vague convergence:

$$\mu_N \to \mu \quad \Longleftrightarrow \quad \int \phi(x,y)\,\mu_N(\mathrm{d}x,\mathrm{d}y) \to \int \phi(x,y)\,\mu(\mathrm{d}x,\mathrm{d}y) \quad \forall \phi \in C_c(\mathbb{R}^2 \times \mathbb{R}^2)\,.$$

Our main result proves weak convergence of the law of $\mathcal{Z}^{\beta}_N$ as $N \to \infty$, when $\beta = \beta_N$ is rescaled in a suitable *critical window*, that we define next. Let us introduce the sequence

$$R_N := \sum_{n=1}^{N} \sum_{z\in\mathbb{Z}^2} \mathrm{P}(S_n = z)^2 = \sum_{n=1}^{N} \mathrm{P}(S_{2n} = 0) = \sum_{n=1}^{N} \left\{ \frac{1}{2^{2n}} \binom{2n}{n} \right\}^2 \sim \frac{\log N}{\pi}\,, \tag{1.10}$$

[†]Note that $\mathbb{E}[Z^{\beta_N,\omega}_{M,N}(w,z)] = \mathrm{P}(S_N = z \,|\, S_M = w) = O(\frac{1}{N-M}) = O(\frac{1}{N})$ for $M/N \leqslant c < 1$, by the local limit theorem, which explains the prefactor N in (1.9). The extra factor $\frac{1}{4}$ is due to periodicity.

which is the expected overlap (number of collisions) between two independent simple symmetric random walks starting from the origin in $\mathbb{Z}^2$ up to time N. Recalling that $\lambda(\cdot)$ is the disorder log-moment generating function, see (1.1), the critical window for $\beta = \beta_N$ is

$$e^{\lambda(2\beta_N)-2\lambda(\beta_N)} - 1 = \frac{1}{R_N}\left(1 + \frac{\vartheta + o(1)}{\log N}\right), \qquad \text{for some fixed } \vartheta \in \mathbb{R}\,. \tag{1.11}$$

Since $\lambda(\beta) \sim \frac{1}{2}\beta^2$ as $\beta \downarrow 0$, see (1.1), we have $\beta_N \sim \hat{\beta}_c/\sqrt{\log N}$ with $\hat{\beta}_c = \sqrt{\pi}$ irrespective of the parameter ϑ, which contributes to the second order asymptotics, see (3.12).

We can now state our main result, which will be proved in Section 9.

Theorem 1.1 (Critical 2d Stochastic Heat Flow). *Fix β_N in the critical window* (1.11)*, for some $\vartheta \in \mathbb{R}$. As $N \to \infty$, the family of random measures $\mathscr{Z}_N^{\beta_N} = (\mathscr{Z}_{N;\,s,t}^{\beta_N}(\mathrm{d}x, \mathrm{d}y))_{0\leqslant s\leqslant t<\infty}$ defined in* (1.9) *converges in finite dimensional distributions to a unique limit*

$$\mathscr{Z}^{\vartheta} = (\mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x, \mathrm{d}y))_{0\leqslant s\leqslant t<\infty}\,,$$

which we call the Critical $2d$ Stochastic Heat Flow*. This limit $\mathscr{Z}^{\vartheta}$ is universal, in that it does not depend on the law of the disorder ω except for the assumptions in* (1.1)*.*

We can infer directly from its construction some basic properties of the Critical $2d$ Stochastic Heat Flow, which we collect in the next result, also proved in Section 9.

Theorem 1.2. *The Critical* $2d$ *Stochastic Heat Flow $\mathscr{Z}^{\vartheta}$ is (space-time) translation invariant in law:*

$$(\mathscr{Z}^{\vartheta}_{s+\mathsf{a},t+\mathsf{a}}(\mathrm{d}(x+\mathsf{b}), \mathrm{d}(y+\mathsf{b})))_{0\leqslant s\leqslant t<\infty} \overset{\text{dist}}{=} (\mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x, \mathrm{d}y))_{0\leqslant s\leqslant t<\infty} \qquad \forall \mathsf{a} \geqslant 0,\ \forall \mathsf{b} \in \mathbb{R}^2\,,$$

and it satisfies the following scaling relation:

$$(\mathscr{Z}^{\vartheta}_{\mathsf{a}s,\mathsf{a}t}(\mathrm{d}(\sqrt{\mathsf{a}}x), \mathrm{d}(\sqrt{\mathsf{a}}y)))_{0\leqslant s\leqslant t<\infty} \overset{\text{dist}}{=} (\mathsf{a}\,\mathscr{Z}^{\vartheta+\log \mathsf{a}}_{s,t}(\mathrm{d}x, \mathrm{d}y))_{0\leqslant s\leqslant t<\infty} \qquad \forall \mathsf{a} > 0\,. \tag{1.12}$$

The first and second moments of $\mathscr{Z}^{\vartheta}$ are given by

$$\begin{aligned} \mathbb{E}[\mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x, \mathrm{d}y)] &= \tfrac{1}{2}\, g_{\frac{1}{2}(t-s)}(y-x)\,\mathrm{d}x\,\mathrm{d}y\,, \\ \mathbb{C}\mathrm{ov}[\mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x, \mathrm{d}y), \mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x', \mathrm{d}y')] &= \tfrac{1}{2}\, K^{\vartheta}_{t-s}(x, x'; y, y')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}x'\,\mathrm{d}y'\,, \end{aligned} \tag{1.13}$$

where g denotes the heat kernel in $\mathbb{R}^2$, see (3.20)*, and K^{ϑ} is an explicit kernel, see* (3.56)*.*

Remark 1.3. *The covariance kernel $K^{\vartheta}_{t-s}(x, x'; y, y')$ was first identified in* [BC98] (*see also* [CSZ19b]) *and is* logarithmically divergent *near the diagonals $x = x'$ or $y = y'$.*

We now briefly explain the proof strategy. As noted before, the moments of $\mathscr{Z}^{\vartheta}$ identified in [GQT21] and [Che21] grow too fast to uniquely characterize the law of $\mathscr{Z}^{\vartheta}$. The bounds given in these works suggest that the n-th moment is at most of order $\exp(\exp(n^2))$, while our recent work [CSZ22] gives a lower bound of $\exp(cn^2)$. Physical arguments on the Delta-Bose gas [Raj99] suggest that the growth should be $\exp(\exp(n))$. It may thus be surprising that we are still able to prove Theorem 1.1 and show that the limit is unique, without criteria to uniquely identify the limit. Another prominent result of this nature, which gave us inspiration, is the work of Kozma [Koz07] on the convergence of the three-dimensional loop erased random walk with dyadic scaling of the lattice $2^{-N}\mathbb{Z}^3$.

The basic strategy is to show that the laws of $(\mathscr{Z}_N^{\beta_N})_{N\in\mathbb{N}}$ form a *Cauchy sequence*, i.e.

$$\mathscr{Z}_N^{\beta_N} \text{ and } \mathscr{Z}_M^{\beta_M} \text{ are close in distribution for large } N, M \in \mathbb{N}\,. \tag{1.14}$$

To accomplish this, we first construct a *coarse-grained model* $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$, for each $\varepsilon \in (0,1)$, which is a function of a family Θ of coarse-grained disorder variables. We then perform a coarse-graining approximation of the partition function on the time-space scale $(\varepsilon N, \sqrt{\varepsilon N})$, which shows that $\mathscr{Z}_N^{\beta_N}$ *can be approximated by the coarse-grained model* $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$ for a specific choice of coarse-grained disorder $\Theta = \Theta_{N,\varepsilon}$ that depends on N and ε, with an approximation error which is small for small ε and large N (shown via second moment bounds). As a consequence, we finally prove (1.14) by showing that the coarse-grained models $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$ with $\Theta = \Theta_{N,\varepsilon}$ and $\Theta = \Theta_{M,\varepsilon}$ are close in distribution, for small $\varepsilon > 0$ and large $N, M \in \mathbb{N}$ (shown via a Lindeberg principle).

We give a more detailed proof outline in Section 2. Let us just highlight here the key proof ingredients:

- **A.** *Coarse-Graining*, which leads to a *coarse-grained model* with the same structure as the original model, demonstrating a degree of self-similarity;
- **B.** *Time-Space Renewal Structure*, which sheds probabilistic light on second moment computations and leads in the continuum limit to the so-called *Dickman subordinator*;
- **C.** *Lindeberg Principle* for multilinear polynomials of *dependent* random variables, which controls the effect of changing Θ in the coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$;
- **D.** *Functional Inequalities for Green's Functions* of multiple random walks on $\mathbb{Z}^2$, which yield sharp higher moment bounds for the coarse-grained model.

This framework is robust enough that it can also be used to show convergence of other approximations of SHE (1.6) to the Critical 2d Stochastic Heat Flow.

Remark 1.4 (Mollified SHE). *The same proof steps* ***A****,* ***B****,* ***C****,* ***D*** *can be carried out for the solution u_δ of the* mollified SHE (1.6)*, where the space-time white noise $\dot{W}$ is mollified spatially on the scale δ and $\beta = \beta_\delta$ is chosen in the corresponding critical window, that is $\beta_\delta^2 = \frac{2\pi}{|\log\delta|} + \frac{\vartheta + o(1)}{(\log\delta)^2}$ (cf.* (3.12)*). A key point is that coarse-graining u_δ on the mesoscopic scale leads to exactly* the same coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$ constructed in this paper*, just with a different family of coarse-grained disorder variables $\Theta = \Theta_{\delta,\varepsilon}$. This means that the solution u_δ of the mollified SHE would converge as $\delta \downarrow 0$ to* the same universal limit $\mathscr{Z}^\vartheta$ in Theorem 1.1. *We will not carry out the details here since the paper is long enough.*

We remark that Clark has proved in [Cla21] an analogue of Theorem 1.1 for DPRE on the hierarchical diamond lattice, which is particularly useful for renormalization analysis and can mimic Euclidean lattices of different dimensions as the lattice parameters vary. Furthermore, in [Cla22, Cla19b], he also constructed the continuum polymer measures and studied their properties. This raises interesting questions as to whether similar results can be proved for DPRE on the Euclidean lattice, where exact renormalization analysis is no longer available. We point out that our work developed in parallel to that of Clark, and our proof strategies share some common features, such as coarse-graining and controlling distributional distances via a Lindeberg principle in our case vs. Stein's method in [Cla21], and showing that the laws of the partition functions form a Cauchy sequence.

Now that we have proved the existence of a unique limit $\mathscr{Z}^\vartheta$ —the Critical 2d Stochstic Heat Flow — the next challenge will be to investigate its properties and characterize its law.

Remark 1.5 (Alternative scaling). *The simple random walk on $\mathbb{Z}^2$ is 2-periodic and each component has variance $\frac{1}{2}$. As a consequence, the diffusively rescaled partition functions $\mathcal{U}_N(t,x)$ in* (1.4) *provide a discretization of a slightly modified SHE* (1.6)*, namely*

$$\partial_t \tilde{u} = \frac{1}{4}\Delta\tilde{u} + \sqrt{2}\,\beta\,\dot{W}\,\tilde{u}$$

(see [CSZ22, Appendix A.3] *for more details). The SHE with the usual parameters in* (1.6) *can be recovered via the change of variable $\mathcal{U}_N(t, \frac{x}{\sqrt{2}})$. Therefore to describe a candidate solution of* (1.6)*, we should consider the rescaled Critical 2d Stochastic Heat Flow given by (recall* (1.12)*)*

$$\widehat{\mathscr{Z}}^\vartheta := \big(\mathscr{Z}^\vartheta_{s,t}(\mathrm{d}\tfrac{x}{\sqrt{2}}, \mathrm{d}\tfrac{y}{\sqrt{2}})\big)_{0\leqslant s\leqslant t<\infty} \overset{d}{=} \big(2\,\mathscr{Z}^{\vartheta+\log 2}_{2s,2t}(\mathrm{d}x,\mathrm{d}y)\big)_{0\leqslant s\leqslant t<\infty},$$

which is also normalized to have mean 1 *rather than* $\frac{1}{2}$ *(see* (1.13)*).*

1.5. Related literature. We next discuss the connection between our work and various results in the literature and point out some future directions of research.

Singular SPDEs. As explained in Section 1.3, the scaling limit $\mathscr{Z}^\vartheta$ in Theorem 1.1 can be interpreted as the solution of the 2-dimensional SHE (1.6) in the critical window. For SHE, dimension $d = 2$ marks the *critical dimension* in the language of singular SPDEs and renormalisation group theory. To define a solution for singular SPDEs, such as SHE and KPZ in (1.6)-(1.7), a standard approach is to mollify the space-time noise $\dot{W}$ in space on the scale of ε, and then try to identify a scaling limit as $\varepsilon \downarrow 0$. Discretizing space-time by considering a lattice model, such as the DPRE that we study in this paper, is just another way of removing the singularity on small scales (also known as ultraviolet cutoff).

All existing solution theories for singular SPDEs, including regularity structures [H13, H14], paracontrolled distributions [GIP15, GP17], the renormalization group approach [Kup14], or energy solutions [GJ14], do not apply at the critical dimension. The only singular SPDEs for which progress has been made in defining its solution at the critical dimension are SHE and KPZ (via the Cole-Hopf transform). The phase transition identified in [CSZ17b] was unexpected, and to the best of our knowledge no such transition has been established for other singular SPDEs in the critical dimension. Theorem 1.1 is thus the first result to define a solution for a singular SPDE at the critical dimension and for critical disorder strength.

In dimension $d = 2$, recently there has also been significant progress in understanding the solution of the *anisotropic version* of the KPZ equation (aKPZ), which differs from (1.7) in that $|\nabla v|^2 = (\partial_{x_1} v)^2 + (\partial_{x_2} v)^2$ therein is replaced by $(\partial_{x_1} v)^2 - (\partial_{x_2} v)^2$. This case is also beyond the reach of existing solution theories, and unlike the isotropic KPZ, it cannot be linearized via the Cole-Hopf transformation. Cannizzaro, Erhard, and Schönbauer [CES21] regularized the aKPZ via a cutoff in Fourier space, instead of discretizing space and time or mollifying the noise on the spatial scale ε (all are ultraviolet cutoffs). They showed that if the non-linear term $(\partial_{x_1} v)^2 - (\partial_{x_2} v)^2$ is rescaled by a factor $\lambda/\sqrt{|\log\varepsilon|}$, then the solution of the regularized aKPZ is tight with non-trivial limit points, which is the anisotropic analogue of [CD20]. Very recently, Cannizzaro, Erhard, and Toninelli [CET21] succeeded in proving that the limit is in fact Gaussian and solves the Edwards-Wilkinson equation, which is the anisotropic analogue of [CSZ20, Gu20]. In contrast to the isotropic case (1.7), there is no phase transition in λ for the aKPZ. The same authors also studied the aKPZ without

scaling the non-linearity, and in a surprising result [CET20a, CET20b], they showed that the solution exhibits logarithmic superdiffusive behaviour.

In the supercitical dimensions $d \geqslant 3$, the transition between the weak and strong disorder phases for the directed polymer is long known [Com17] and has a natural counterpart for SHE and KPZ. In recent years, there have been many studies on the solutions of SHE and KPZ via mollification, namely, analogues of the random fields $\mathcal{U}_N$ and $\mathcal{H}_N$ defined in (1.4)-(1.5). These studies are all in the weak disorder regime and are analgous to results in $d = 2$, see e.g. [MU18, MSZ16, CN21, CNN22, CCM20, GRZ18, DGRZ20, LZ22].

Coarse-graining. The first step in our approach is to construct a coarse-grained model. Coarse-graining has a long history in statistical mechanics and renormalisation theory. In the framework of directed polymer models, coarse-graining has played a crucial role in the studies by Lacoin [L10] and Berger-Lacoin [BL17] on free energy asymptotics, which extended previous works in the literature of pinning models, see [Gi11], from which we single out the fundamental work of Giacomin-Lacoin-Toninelli [GLT10].

In our analysis, we need a family of coarse-grained models which provide a *sharp approximation* of the partition function *at the critical point*, while the works mentioned above used coarse-grained models to provide upper bounds away from the critical point. The need for a sharper approximation creates several challenges, which lead to the refined estimates in Sections 5 and 8 and the development of the enhanced Lindeberg principle in Appendix A.

DPRE on hierarchical lattices. In a series of papers [Cla21, Cla22, Cla19b], Clark successfully treated the directed polymer model on hierarchical diamond lattices at the "critical dimension" and in the critical window of disorder strength, which contains an analogue of Theorem 1.1 and more. Due to their tree-like structure, hierarchical lattices are especially convenient for performing exact renormalization group calculations that are typically intractable on the Euclidean lattice. By tuning suitable parameters (such as the number of branches and the number of segments along each branch), hierarchical lattices can mimic Euclidean lattices with different spatial dimensions. When the branch number equals the segment number, hierarchical lattices mimic $\mathbb{Z}^2$. For DPRE on these lattices, Clark was able to prove in [Cla21] the analogue of Theorem 1.1.

Exploiting the structure of hierarchical lattices, in [Cla22], Clark was able to use the limiting partition functions obtained in [Cla21] to construct a continuum version of the polymer measure and study its properties. Furthermore, in [Cla19b], he identified an interesting conditional Gaussian Multiplicative Chaos (GMC) structure among the continuum polymer measures with different parameter ϑ (similar to ϑ in Theorem 1.1). These results raise interesting questions as to whether similar results can be obtained for DPRE on the Euclidean lattice. In this respect, Theorem 1.1 provides the starting point.

Continuum polymer measure. A continuum version of the DPRE polymer measure in dimension $d = 1$ was constructed in [AKQ14a, AKQ14b], exploiting the continuum limit of the point-to-point partition functions. The same approach was applied in [CSZ16] to pinning models with tail exponent $\alpha \in (\frac{1}{2}, 1)$. An essential feature of these constructions, as well as the one by Clark [Cla22] in the hierarchical setting in the critical regime, is that the continuum partition functions are random *functions* of the polymer endpoints. The same holds for DPRE in dimension $d = 2$ in the *subcritical regime* $\beta_N \sim \hat{\beta}/\sqrt{\log N}$, with $\hat{\beta} < \hat{\beta}_c = \sqrt{\pi}$, where it was recently shown in [Ga21] that the discrete polymer measure, diffusively rescaled, converges to the law of Brownian motion.

The situation for DPRE in dimension $d = 2$ *in the critical window* is radically different, because the continuum partition functions $\mathcal{Z}^{\vartheta}_{z,t}(\mathrm{d}x, \mathrm{d}y)$ given in Theorem 1.1 are only random *measures* and undefined pointwise. The point-to-plane partition function $Z_N^{\beta_N,\omega}$ defined in (1.3) in fact converges to 0 as $N \to \infty$, as shown in [CSZ17b]. For this reason, constructing a continuum version of the polymer measure — or studying the scaling properties of the discrete polymer measure — started from a *fixed* point, remains a significant challenge. However, if we consider discrete polymer measures with the starting point chosen uniformly from a ball on the diffusive scale, then the same proof strategy as that for Theorem 1.1 should be applicable to show that the measures converge to a continuum polymer measure starting from a ball, whose finite dimensional distributions are uniquely determined.

Schrödinger operators with point interactions. When the disorder ω is standard normal, a direct calculation shows that for $k \in \mathbb{N}$, the k-th moment of the polymer partition function in (1.3) is the exponential moment (with parameter β^2) of the total pairwise collision local time up to time N among k independent random walks on $\mathbb{Z}^2$. When $k = 2$, by a classic result of Erdös and Taylor [ET60] (see also [GaSu09]), the collision local time rescaled by $1/\log N$ converges to an exponential random variable with parameter π. In the critical window we consider here, we have $\beta_N = \hat{\beta}_c/\sqrt{\log N}$ with $\hat{\beta}_c = \sqrt{\pi}$, and hence the parameter of the exponential moment matches exactly the parameter of the limiting exponential random variable, making the moment analysis particularly delicate.

Via the Feynman-Kac formula, it can also be seen that the k-th moment of the partition function is the solution of a discrete space-time parabolic Schrödinger equation with a potential supported on the diagonal (point interaction). In the continuum setting, there have been a number of studies on the Schrödinger operator with point interactions (also called Delta-Bose gas) in dimension $d = 2$ [AGH+05, AFH+92, DFT94, DR04]. Using ideas from these studies, especially the works of Dell'Antonio-Figari-Teta [DFT94] and of Dimock-Rajeev [DR04], Gu, Quastel, and Tsai [GQT21] were able to compute asymptotically all moments of the averaged solution of the mollified SHE, which are analogues of the averaged polymer partition functions $\mathcal{Z}^{\beta_N}_{N;\,s,t}(\varphi, \psi) := \iint \varphi(x)\,\psi(y)\,\mathcal{Z}^{\beta_N}_{N;\,s,t}(\mathrm{d}x, \mathrm{d}y)$ in (1.9), with φ and ψ assumed to be in L^2 in [GQT21]. Previously, only the third moment had been obtained in [CSZ19b]. When φ is a delta function, the moments of $\mathcal{Z}^{\beta_N}_{N;\,s,t}(\varphi, \psi)$ diverge as $N \to \infty$, and the asymptotics of the third moment has been investigated in [F16]. But all mixed moments of the form $\mathrm{E}\big[\prod_{i=1}^n \mathcal{Z}^{\beta_N}_{N;\,s,t}(\varphi_i, \psi_i)\big]$ converge if φ_i are chosen to be distinct δ functions, which was shown recently by Chen in [Che21].

As an input to the Lindeberg principle mentioned in the proof sketch for Theorem 1.1, we need to bound the fourth moment of the coarse-grained model, which approximates the original partition function. The results from the Schrödinger operator literature and [GQT21] are not applicable in our setting, because they rely on explicit Fourier calculations. We therefore develop an alternative and more robust approach based on *functional inequalities* for Green's function of multiple random walks on $\mathbb{Z}^2$, see Lemma 6.8. Instead of working with $\varphi, \psi \in L^2$ as in [GQT21], we can work with *weighted L^p–L^q spaces with $\frac{1}{p} + \frac{1}{q} = 1$*. The choice of a weight allows us to consider a wider class of boundary conditions, such as $\psi \equiv 1$ and φ an approximate delta function, and also to control spatial decay when the support of φ and ψ are far apart, all of which are needed in our proof. See Section 6 for more details.

Lindeberg principle. A Lindeberg principle is said to hold when the law of a function Φ of a family of random variables does not change much if the family of random variables is

switched to another family with some matching moments. Lindeberg principles have been very powerful tools in proving universality. The usual formulation such as in [Cha06] requires the family of random variables to be independent (or exchangeable), and Φ needs to have bounded first three derivatives. This is not satisfied when Φ is a multilinear polynomial, whose derivatives are unbounded. This case was addressed in [Rot79, MOO10] when the arguments are independent random variables (see also [CSZ17a]).

In the proof of Theorem 1.1, we need to deal with a multilinear polynomial of *dependent* random variables with a local form of dependence. We formulate an extension of the Lindeberg principle to this setting in Appendix A. Our calculations are inspired by a work of Röllin on Stein's method [Rol13], which is an analogue of [Cha06] for a function Φ (with bounded first three derivatives) of dependent random variables.

1.6. Structure of the paper. The rest of the paper is organized as follows.

- In Section 2, we give a detailed proof outline.
- In Section 3, we introduce some basic notation and tools that we need for the rest of the paper, which includes in particular the polynomial chaos expansion and second moment asymptotics for the partition function.
- In Section 4, we define the coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$ and the coarse-grained disorder $\Theta = \Theta_{N,\varepsilon}$. Then in Section 5, we show that $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta_{N,\varepsilon})$ provides a good L^2 approximation for the diffusively rescaled partition functions $\mathcal{Z}_N$ in (1.9).
- In Sections 6, 7 and 8, we derive key moment bounds for $\mathcal{Z}_N$, $\Theta_{N,\varepsilon}$ and $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\,\cdot\,|\Theta)$.
- In Section 9, we wrap up the proof of our main results: Theorems 1.1 and 1.2.
- In Appendix A, we formulate an enhanced Lindeberg principle for multilinear polynomials of dependent random variables.

Notation. We denote by $C_b(\mathbb{R}^d)$, resp. $C_c(\mathbb{R}^d)$, the space of bounded, resp. compactly supported functions $\varphi : \mathbb{R}^d \to \mathbb{R}$. The usual L^p norms will be denoted by $\|\varphi\|_p$ for functions $\varphi : \mathbb{R}^d \to \mathbb{R}$ and by $\|X\|_{L^p}$ for random variables X. For notational simplicity, we will use c, C, C', C'' to denote generic constants, whose values may change from place to place.

2. Proof outline

We elaborate in more detail our proof strategy for Theorem 1.1, especially the coarse-graining procedure. After reading the proof strategy, to see how the pieces fit together more precisely, we encourage the reader to go directly to Section 9.1 to read the proof of Theorems 1.1 and 1.2. The proof is contingent on some earlier results, such as Theorems 4.7 and 8.1, but otherwise is mostly self-contained.

Recalling (1.9), we just consider a single averaged partition function

$$\mathcal{Z}_N := \mathcal{Z}_{N,0,1}^{\beta_N}(\varphi,\psi) = \int_{\mathbb{R}^2\times\mathbb{R}^2} \varphi(x)\,\psi(y)\,\mathcal{Z}_{N,0,1}^{\beta_N}(\mathrm{d}x,\mathrm{d}y)\,,$$

for some $\varphi \in C_c(\mathbb{R}^2)$, $\psi \in C_b(\mathbb{R}^2)$, and $\beta_N = \beta_N(\vartheta)$ chosen as in (1.11) for some fixed $\vartheta \in \mathbb{R}$. To prove that $\mathcal{Z}_N$ converges in distribution to a limit as claimed in Theorem 1.1, we will show that the laws of $(\mathcal{Z}_N)_{N\in\mathbb{N}}$ form a Cauchy sequence.

The starting point of our analysis is a polynomial chaos expansion for $\mathcal{Z}_N$, which will be recalled in more detail in Section 3.3. In short, by introducing the i.i.d. random variables

$$\xi_N(n,z) := e^{\beta_N \omega(n,z) - \lambda(\beta_N)} - 1, \qquad (n,z) \in \mathbb{N} \times \mathbb{Z}^2,$$

which have mean 0 and variance σ_N^2 as in (1.11), we can expand $\mathcal{Z}_N$ as a multilinear polynomial in the ξ_N's as follows:

$$\mathcal{Z}_N \;=\; q^N_{0,N}(\varphi,\psi) + \\ \frac{1}{N}\sum_{r=1}^{\infty} \sum_{\substack{z_1,\dots,z_r\in\mathbb{Z}^2 \\ 0<n_1<\dots<n_r<N}} q^N_{0,n_1}(\varphi,z_1)\,\xi_N(n_1,z_1)\Big\{\prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)\xi_N(n_j,z_j)\Big\} q^N_{n_r,N}(z_r,\psi)\,, \quad (2.1)$$

where $q_{m,n}(x,y) := \mathrm{P}(S_n = y | S_m = x)$ is the random walk transition kernel, and $q^N_{m,n}(\varphi,z_1)$, $q^N_{m,n}(z_r,\psi)$, $q^N_{m,n}(\varphi,\psi)$ are the averages of $q_{m,n}(x,y)$ w.r.t. $\varphi(x/\sqrt{N})$, $\psi(y/\sqrt{N})$, or both (see (3.16)-(3.18)).

Each term in the sum in (2.1) contains a sequence of disorder variables $(\xi_N(n_j,z_j))_{1\leqslant j\leqslant r}$ linked by random walk transition kernels, and different terms in the sum are L^2-orthogonal. We will see that when it comes to second moment calculations, the sequence of points $(n_1,z_1),\dots,(n_r,z_r)$ can be interpreted as a *time-space renewal configuration*.

Before explaining our proof strategy and ingredients, we first give a heuristic calculation that already shows universality, namely that as $N\to\infty$, the limiting law of $\mathcal{Z}_N$ in (2.1) (if a unique limit exists) does not depend on the law of the i.i.d. random variables $\xi_N(\cdot,\cdot)$ provided the first two moments are unchanged. The heuristic is based on a Lindeberg principle, which will help to illustrate some key ideas in our proof.

A. Heuristic Calculation. Let us write $\mathcal{Z}_N(\xi_N)$ to emphasise the dependence on the i.i.d. family $\xi_N(\cdot,\cdot)$, and let $\mathcal{Z}_N(\eta_N)$ be defined similarly with ξ_N replaced by an i.i.d. family η_N with matching first two moments and finite third moment. To show that $\mathcal{Z}_N(\xi_N)$ and $\mathcal{Z}_N(\eta_N)$ are close in law, it suffices to show that for any $f:\mathbb{R}\to\mathbb{R}$ with bounded first three derivatives,

$$\lim_{N\to\infty} |f(\mathcal{Z}_N(\xi_N)) - f(\mathcal{Z}_N(\eta_N))| = 0. \quad (2.2)$$

This difference can be bounded by a Lindeberg principle. In particular, we can apply Theorem A.4 to the case of i.i.d. random variables (the sums in (A.9)-(A.10) will only contain indices $k=l=m$ due to the i.i.d. assumption) to get the bound

$$|f(\mathcal{Z}_N(\xi_N)) - f(\mathcal{Z}_N(\eta_N))| \leqslant C\|f'''\|_\infty \sum_{1\leqslant n\leqslant N, z\in\mathbb{Z}^2} \int_0^1 \mathrm{E}[|\partial_{(n,z)}\mathcal{Z}(\xi_N^{(t)})|^3]\mathrm{d}t, \quad (2.3)$$

where $\xi_N^{(t)} := \sqrt{t}\,\xi_N + \sqrt{1-t}\,\eta_N$ interpolates between η_N and ξ_N, and $\partial_{(n,z)}\mathcal{Z}(\xi_N)$ denotes partial derivative w.r.t. $\xi_N(n,z)$. Since $\mathcal{Z}(\xi_N^{(t)})$ is a multilinear polynomial in $\xi_N^{(t)}(\cdot,\cdot)$, it is easily seen from (2.1) that

$$\partial_{(n,z)}\mathcal{Z}(\xi_N^{(t)}) = \frac{1}{N}\mathcal{Z}(\varphi,(n,z))\mathcal{Z}((n,z),\psi),$$

where $\mathcal{Z}((n,z),\psi)$ is the point-to-plane partition function starting from the point (n,z) and terminating at time N with boundary condition ψ, and $\mathcal{Z}(\varphi,(n,z))$ is the plane-to-point partition function with initial boundary condition φ and terminating at the point (n,z). Since φ has compact support, only (n,z) on the diffusive scale (n of order N and $z\in\mathbb{Z}^2$ of order $\sqrt{N}$)

contribute to the sum in (2.3), and there are N^2 such terms. This sum is more than compensated by the factor $\frac{1}{N^3}$ from $\mathrm{E}[|\partial_{\xi_N(n,z)}\mathcal{Z}(\xi_N)|^3] = \frac{1}{N^3}\mathrm{E}[|\mathcal{Z}(\varphi,(n,z))|^3]\mathrm{E}[|\mathcal{Z}((n,z),\psi)|^3]$, where we used the independence between $\mathcal{Z}(\varphi,(n,z))$ and $\mathcal{Z}((n,z),\psi)$. To deduce (2.2), it suffices to show that the moment of the point-to-plane partition function $\mathrm{E}[|\mathcal{Z}((n,z),\psi)|^3] \ll N$ as $N \to \infty$, which holds by Remark 6.5 below.

This heuristic can be made rigorous using the results we establish in Section 6. But this argument will not show that $\mathcal{Z}_N(\xi_N)$ has a unique limit in law. For that, we need to define coarse-grained models and compare $\mathcal{Z}_N(\xi_N)$, for different N, with the same coarse-grained model. We outline the proof strategy below, which contains many of the same ideas in the heuristic above, but in a more complicated setting.

A. Coarse-graining. As a first step, for each $\varepsilon \in (0,1)$, we approximate $\mathcal{Z}_N$ in L^2 by a coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, which is a multi-linear polynomial in suitable coarse-grained disorder variables $\Theta_{N,\varepsilon}$ and depends on N only through $\Theta_{N,\varepsilon}$. The details will be given in Section 4. Here we give a sketch.

We partition $\mathbb{N}\times\mathbb{Z}^2$ into *mesoscopic* time-space boxes

$$\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a}) \;:=\; \underbrace{((\mathsf{i}-1)\varepsilon N, \mathsf{i}\varepsilon N]}_{\mathcal{T}_{\varepsilon N}(\mathsf{i})} \times \underbrace{((\mathsf{a}-(1,1))\sqrt{\varepsilon N}, \mathsf{a}\sqrt{\varepsilon N}]}_{\mathcal{S}_{\varepsilon N}(\mathsf{a})} \;\cap\; \mathbb{Z}^3_{\mathrm{even}}\,, \tag{2.4}$$

where $(\mathsf{i},\mathsf{a}) \in \mathbb{N}\times\mathbb{Z}^2$ is the mesoscopic time-space index of $\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})$, which has temporal width εN and spatial side length $\sqrt{\varepsilon N}$, and $(\mathsf{a}-\mathsf{b},\mathsf{a}] = (\mathsf{a}_1-\mathsf{b}_1,\mathsf{a}_1]\times(\mathsf{a}_2-\mathsf{b}_2,\mathsf{a}_2]$ for squares in $\mathbb{R}^2$. We then decompose the sum in (2.1) according to the sequence of mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\dots,\mathcal{T}_{\varepsilon N}(\mathsf{i}_k)$ visited by the renewal configuration $(n_1,z_1),\dots,(n_r,z_r)$. For each $\mathcal{T}_{\varepsilon N}(\mathsf{i}_j)$, we then further decompose according to the first and last mesoscopic spatial boxes $\mathcal{S}_{\varepsilon N}(\mathsf{a}_j),\mathcal{S}_{\varepsilon N}(\mathsf{a}'_j)$ visited in this time interval. This replaces the microscopic sum over $(n_1,z_1),\dots,(n_r,z_r)$ in (2.1) by a mesoscopic sum over time-space renewal configurations $(\mathsf{i}_1;\mathsf{a}_1,\mathsf{a}'_1),\dots,(\mathsf{i}_k;\mathsf{a}_k,\mathsf{a}'_k)$, which specify the sequence of mesoscopic boxes $\mathcal{B}_{\varepsilon N}(\mathsf{i}_j,\mathsf{a}_j)$ and $\mathcal{B}_{\varepsilon N}(\mathsf{i}_j,\mathsf{a}'_j)$ visited. See Figure 1 for an illustration.

Ideally, we would like to replace each random walk kernel $q_{n,m}(x,y)$ in (2.1) that connects two consecutive visited mesoscopic boxes $\mathcal{B}_{\varepsilon N}(\mathsf{i}_j,\mathsf{a}'_j) \ni (n,x)$ and $\mathcal{B}_{\varepsilon N}(\mathsf{i}_{j+1},\mathsf{a}_{j+1}) \ni (m,y)$ by a corresponding heat kernel. Namely, by the local limit theorem (3.21), replace $q_{n,m}(x,y)$ by

$$2\,g_{\frac{1}{2}(\mathsf{i}_{j+1}-\mathsf{i}_j)\varepsilon N}((\mathsf{a}_{j+1}-\mathsf{a}'_j)\sqrt{\varepsilon N}) = \frac{2}{\varepsilon N} g_{\frac{1}{2}(\mathsf{i}_{j+1}-\mathsf{i}_j)}(\mathsf{a}_{j+1}-\mathsf{a}'_j)\,,$$

where the factor 2 is due to periodicity. With such replacements, given a mesoscopic renewal configuration $(\mathsf{i}_1;\mathsf{a}_1,\mathsf{a}'_1),\dots,(\mathsf{i}_k;\mathsf{a}_k,\mathsf{a}'_k)$, as we sum over compatible microscopic renewal configurations $(n_1,z_1),\dots,(n_r,z_r)$ in (2.1), the contributions of $\xi_N(n,z)$ from each interval $\mathcal{T}_{\varepsilon N}(\mathsf{i}_j)$ would decouple, leading to a product of coarse-grained disorder variables of the form

$$\Theta_{N,\varepsilon}(\mathsf{i}_j;\mathsf{a}_j,\mathsf{a}'_j) := \frac{2}{\varepsilon N}\sum_{r=1}^{\infty} \sum_{\substack{(n_1,z_1),\dots,(n_r,z_r)\in\mathbb{Z}^3_{\mathrm{even}}\\ z_1\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_j), z_r\in\mathcal{S}_{\varepsilon N}(\mathsf{a}'_j)\\ n_1<\dots<n_r,\, n_i\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_j)}} \xi_N(n_1,z_1)\prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)\xi_N(n_j,z_j)\,, \tag{2.5}$$

with consecutive coarse-grained disorder variables $\Theta_{N,\varepsilon}(\mathsf{i}_j;\mathsf{a}_j,\mathsf{a}'_j)$ and $\Theta_{N,\varepsilon}(\mathsf{i}_{j+1};\mathsf{a}_{j+1},\mathsf{a}'_{j+1})$ linked by the heat kernel $g_{\frac{1}{2}(\mathsf{i}_{j+1}-\mathsf{i}_j)}(\mathsf{a}_{j+1}-\mathsf{a}'_j)$ (we absorbed the factor $\frac{2}{\varepsilon N}$ into (2.5)). This would give our desired coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$.

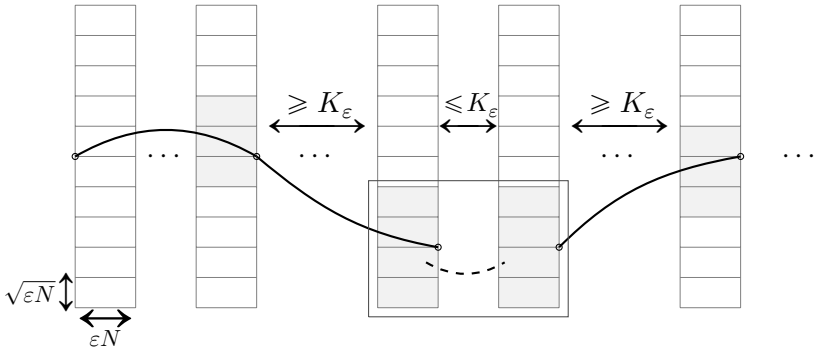


FIGURE 1. An illustration of the chaos expansion for the coarse-grained model (2.7). The solid laces represent heat kernels linking consecutively visited mesoscopic time-space boxes. The grey blocks represent the regions defining the coarse-grained disorder variables $\Theta_{N,\varepsilon}$. The double block in the middle represents a coarse-grained disorder variable $\Theta_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ visiting two mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ and $\mathcal{T}_{\varepsilon N}(\mathsf{i}')$ with $|\mathsf{i}'-\mathsf{i}|\leqslant K_\varepsilon=(\log\frac{1}{\varepsilon})^6$ and cannot be decoupled.

Unfortunately, this ideal procedure *does not produce a sharp approximation of the partition function* $\mathcal{Z}_N$ *in* (2.1). Indeed, the kernel replacement

$$q_{n,m}(x,y)\rightsquigarrow\frac{2}{\varepsilon N}g_{\frac{1}{2}(\mathsf{i}_{j+1}-\mathsf{i}_j)}(\mathsf{a}_{j+1}-\mathsf{a}'_j)\tag{2.6}$$

induces an L^2-error, and this error is small (in the sense that it vanishes as $\varepsilon\downarrow 0$, uniformly in large N) *only if* $\mathsf{i}_{j+1}-\mathsf{i}_j$ *is sufficiently large* (we will choose it to be larger than $K_\varepsilon=(\log\frac{1}{\varepsilon})^6$) and $|\mathsf{a}_{j+1}-\mathsf{a}'_j|$ *is not too large on the diffusive scale* (we will choose it to be smaller than $M_\varepsilon\sqrt{\mathsf{i}_{j+1}-\mathsf{i}_j}$ with $M_\varepsilon=\log\log\frac{1}{\varepsilon}$). We address this issue as follows.

The first crucial observation is that, modulo a small L^2 error, microscopic renewal configurations (n_1,z_1), …, (n_r,z_r) in (2.1) *cannot visit three or more mesoscopic time intervals* $\mathcal{T}_{\varepsilon N}(\mathsf{i}_j)$, $\mathcal{T}_{\varepsilon N}(\mathsf{i}_{j+1})$, *and* $\mathcal{T}_{\varepsilon N}(\mathsf{i}_{j+2})$ *with both* $\mathsf{i}_{j+1}-\mathsf{i}_j\leqslant K_\varepsilon$ *and* $\mathsf{i}_{j+2}-\mathsf{i}_{j+1}\leqslant K_\varepsilon$ (see Lemma 5.1 below). Furthermore, with a small L^2 error, we can also enforce a diffusive truncation $|\mathsf{a}_{j+1}-\mathsf{a}'_j|\leqslant M_\varepsilon\sqrt{\mathsf{i}_{j+1}-\mathsf{i}_j}$ (see Lemma 5.6 below). We will then make the random walk/heat kernel replacement (2.6) *only between mesoscopic boxes* $\mathcal{B}_{\varepsilon N}(\mathsf{i}_j,\mathsf{a}'_j)\ni(n,x)$ *and* $\mathcal{B}_{\varepsilon N}(\mathsf{i}_{j+1},\mathsf{a}_{j+1})\ni(m,y)$ *that satisfy the constraint* $\mathsf{i}_{j+1}-\mathsf{i}_j>K_\varepsilon$.

After such kernel replacements, what are left between the heat kernels decouple and appear as a product of two types of coarse-grained disorder variables:

- one type is as given in (2.5), which visits a single mesoscopic time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i})$;
- another type visits two mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ and $\mathcal{T}_{\varepsilon N}(\mathsf{i}')$, with $\mathsf{i}'-\mathsf{i}\leqslant K_\varepsilon$: we denote it by $\Theta_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ with $\vec{\mathsf{i}}=(\mathsf{i},\mathsf{i}')$ and $\vec{\mathsf{a}}=(\mathsf{a},\mathsf{a}')$, where a identifies the first mesoscopic spatial box visited in the time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i})$, while a' identifies the last mesoscopic spatial box visited in the time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i}')$ (see (4.11)).

This leads to the actual coarse-grained model we will work with:

$$\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta) := \tfrac{1}{2}\, g_{\frac{1}{2}}(\varphi,\psi) + \\ \frac{\varepsilon}{2} \sum_{r=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\\(\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)}} g_{\frac{1}{2}\mathsf{i}_1}(\varphi_\varepsilon,\mathsf{a}_1)\Theta(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1)\Bigg\{\prod_{j=2}^{r} g_{\frac{1}{2}(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\Theta(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j)\Bigg\} g_{\frac{1}{2}(\frac{1}{\varepsilon}-\mathsf{i}'_r)}(\mathsf{a}'_r,\psi_\varepsilon), \tag{2.7}$$

where φ_ε and ψ_ε are averaged versions of φ and ψ on the spatial scale $\sqrt{\varepsilon}$, while $g_{\mathsf{i}/2}(\varphi_\varepsilon,\mathsf{a})$, $g_{\mathsf{i}/2}(\mathsf{a}',\psi_\varepsilon)$, $g_{\mathsf{i}/2}(\varphi_\varepsilon,\psi_\varepsilon)$ are averages of the heat kernel $g_{\mathsf{i}/2}(\mathsf{a}-\mathsf{a}')$ w.r.t. φ_ε, ψ_ε, or both.

In the sum in (2.7), we have hidden the various constraints on the mesoscopic time-space variables for simplicity (see (4.8) for the complete definition). Also note that in (2.7) we denote by $\Theta = (\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}}))$ a generic family of coarse-grained disorder variables; in order to approximate the averaged partition function $\mathscr{Z}_N$, we simply set $\Theta = \Theta_{N,\varepsilon}$.

Remark 2.1 (Self-similarity). *The coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta)$ in (2.7) has the same form as the original partition function $\mathscr{Z}_N$ in (2.1), with $1/\varepsilon$ in place of N, $\Theta_{N,\varepsilon}$ in place of ξ_N, and the heat kernel $g_{\mathsf{i}/2}$ in place of the random walk kernel q_n. This shows a remarkable degree of self-similarity: coarse-graining retains the structure of the model.*

B. Time-Space Renewal Structure. Once we have defined precisely the coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, see Section 4, we need to show that it indeed provides a good L^2 approximations of the original partition function $\mathscr{Z}_N$, in the following sense:

$$\lim_{\varepsilon\downarrow 0}\ \limsup_{N\to\infty}\ \big\|\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon}) - \mathscr{Z}_N\big\|_{L^2}^2 \,=\, 0\,. \tag{2.8}$$

This approximations will be carried out in Section 5, where we rely crucially on the time-space renewal interpretation of the sum in (2.1), which in the continuum limit with $N\to\infty$ leads to the so-called Dickman subordinator [CSZ19a]. This will be reviewed in Section 3.5.

C. Lindeberg Principle. In view of (2.8), given $\varepsilon > 0$ small, we can approximate $\mathscr{Z}_N$ by $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, where the L^2 error is uniform in large N and tends to 0 as $\varepsilon\downarrow 0$. To prove that the laws of $(\mathscr{Z}_N)_{N\in\mathbb{N}}$ form a Cauchy sequence, it then suffices to show that given $\varepsilon > 0$ we can bound the distributional distance between $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{M,\varepsilon})$ and $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$ uniformly in $M \geqslant N$ large, and furthermore, this bound can be made arbitrarily small by choosing $\varepsilon > 0$ sufficiently small. This would then complete the proof that $\mathscr{Z}_N$ converges in distribution to a unique limit.

The control of the distributional distance is carried out via a Lindeberg principle for the coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, which is a multilinear polynomial in the family of coarse-grained disorder variables $\Theta_{N,\varepsilon} = \{\Theta_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})\}$. We note that $\Theta_{N,\varepsilon}((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}'))$ and $\Theta_{N,\varepsilon}((\mathsf{j},\mathsf{j}'),(\mathsf{b},\mathsf{b}'))$ have non-trivial dependence if (i,a) or $(\mathsf{i}',\mathsf{a}')$ coincides with either (j,b) or $(\mathsf{j}',\mathsf{b}')$. We thus need a Lindeberg principle for multilinear polynomials of *dependent* random variables, which we formulate in Appendix A and is of independent interest.

D. Functional Inequalities for Green's Functions. To successfully apply the Lindeberg principle, we need to control the second and fourth moments of the coarse-grained disorder variables $\Theta_{N,\varepsilon}$. We also need to control the *influence* of each $\Theta_{N,\varepsilon}$, which

boils down to bounding the fourth moment of the coarse-grained model $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, with the choice of boundary conditions $\psi\equiv 1$ and $\varphi(x)=\frac{1}{\varepsilon}\mathbb{1}_{|x|\leqslant\sqrt{\varepsilon}}$.

The moment bounds on $\Theta_{N,\varepsilon}$ and $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$ are technically the most delicate parts of the paper, especially since we need to allow $\varphi(x)=\frac{1}{\varepsilon}\mathbb{1}_{|x|\leqslant\sqrt{\varepsilon}}$ and $\psi\equiv 1$. Since the structure of $\Theta_{N,\varepsilon}$ is similar to an averaged partition function, we will first derive general moment bounds on the averaged partition function $\mathscr{Z}_N$ in Section 6. The fourth moment bound on $\Theta_{N,\varepsilon}$ then follows as a corollary in Section 7.

The approach we develop is different from the methods employed in [GQT21] to bound the moments of the averaged solution of the mollified SHE. Our approach is based on functional inequalities for the Green's function of random walks (see Lemma 6.8) and it is robust enough to be applied also to the coarse-grained model defined in (2.7), which will be carried out Section 8.

3. Notation and tools

In this section, we introduce some basic notation and tools, including the polynomial chaos expansion for the partition function, random walk estimates, the renewal interpretation for the second moment of partition functions and the Dickman subordinator that arises in the continuum limit.

3.1. Random walk and disorder. As in Section 1.2, let $(S=(S_n)_{n\geqslant 0},\mathrm{P})$ be the simple symmetric random walk on $\mathbb{Z}^2$, whose transition kernel we denote by

$$q_n(z):=\mathrm{P}(S_n=z)\,,\qquad q_{m,n}(x,z):=q_{n-m}(z-x)=\mathrm{P}(S_n=z\,|\,S_m=x)\,. \tag{3.1}$$

Let $(\omega=(\omega(n,z))_{n\in\mathbb{N},z\in\mathbb{Z}^2},\mathbb{P})$ be the disorder, given by a family of i.i.d. random variables with zero mean, unit variance and locally finite exponential moments, see (1.1).

The expected overlap between two independent walks is (see [CSZ19a, Proposition 3.2])

$$\begin{aligned} R_N&:=\sum_{n=1}^N\sum_{z\in\mathbb{Z}^2}q_n(z)^2=\sum_{n=1}^N q_{2n}(0)=\frac{\log N}{\pi}+\frac{\alpha}{\pi}+o(1)\\ \text{with}\qquad \alpha&:=\gamma+\log 16-\pi\simeq 0.208\,,\qquad \gamma:=-\textstyle\int_0^\infty e^{-u}\log u\,\mathrm{d}u\simeq 0.577\,. \end{aligned} \tag{3.2}$$

Note that R_N is the expected number of collisions up to time N between two independent copies of the random walk S when both start from the origin. Also note that γ is the Euler-Mascheroni constant. We further define

$$u(n):=\sum_{x\in\mathbb{Z}^2}q_n(x)^2=q_{2n}(0)\sim\frac{1}{\pi}\cdot\frac{1}{n}\qquad\text{as } n\to\infty\,, \tag{3.3}$$

where the asymptotic behavior follows by the local limit theorem, see (3.21) below.

In order to deal with the periodicity of simple random walk, we set

$$\mathbb{Z}^d_{\mathrm{even}}:=\{z=(z_1,\dots,z_d)\in\mathbb{Z}^d:\ z_1+\ldots+z_d\text{ is even}\}\,. \tag{3.4}$$

Given $x\in\mathbb{R}^d$ with $d\geqslant 2$, we denote by $[\![x]\!]$ the point in $\mathbb{Z}^d_{\mathrm{even}}$ closest to x (fix any convention to break the tie if $[\![x]\!]$ is not unique). More explicitly, we have

$$[\![x]\!]=v\in\mathbb{Z}^d_{\mathrm{even}}\qquad\Longleftrightarrow\qquad x\in B(v):=\big\{x\in\mathbb{R}^d:\ |x_1-v_1|+\ldots+|x_d-v_d|<1\big\}\,. \tag{3.5}$$

For $s \in \mathbb{R}$ it is convenient to define the even approximation $[\![s]\!] \in 2\mathbb{Z}$ by

$$[\![s]\!] := 2\left\lfloor \frac{s}{2} \right\rfloor. \tag{3.6}$$

3.2. Partition functions at criticality. The point-to-point partition functions $Z^{\beta}_{M,N}(w,z)$ were defined in (1.8). We mainly consider the case $M = 0$, for which we write

$$Z^{\beta}_{N}(w,z) = \mathrm{E}\left[e^{\sum_{n=1}^{N-1}\{\beta\omega(n,S_n)-\lambda(\beta)\}} \mathbb{1}_{S_N=z} \,\middle|\, S_0 = w \right]. \tag{3.7}$$

The field of diffusively rescaled partition functions $\mathcal{Z}^{\beta}_{N;s,t}(\mathrm{d}x,\mathrm{d}y)$ was introduced in (1.9). In the special case $s = 0$ we simply write:

$$\mathcal{Z}^{\beta}_{N,t}(\mathrm{d}x,\mathrm{d}y) := \frac{N}{4}\, Z^{\beta}_{[\![Nt]\!]}([\![\sqrt{N}x]\!],[\![\sqrt{N}y]\!])\,\mathrm{d}x\,\mathrm{d}y\,,$$

where we recall that $\mathrm{d}x\,\mathrm{d}y$ denotes the Lebesgue measure on $\mathbb{R}^2 \times \mathbb{R}^2$. We next define *averaged partition functions* $\mathcal{Z}^{\beta}_{N,t}(\varphi,\psi)$ for suitable $\varphi,\psi : \mathbb{R}^2 \to \mathbb{R}$:

$$\begin{aligned} \mathcal{Z}^{\beta}_{N,t}(\varphi,\psi) &:= \iint\limits_{\mathbb{R}^2\times\mathbb{R}^2} \varphi(x)\,\mathcal{Z}^{\beta}_{N,t}(\mathrm{d}x,\,\mathrm{d}y)\,\psi(y) \\ &= \frac{1}{4N} \iint\limits_{\mathbb{R}^2\times\mathbb{R}^2} \varphi(\tfrac{x}{\sqrt{N}})\; Z^{\beta}_{[\![Nt]\!]}([\![x]\!],[\![y]\!])\; \psi(\tfrac{y}{\sqrt{N}})\,\mathrm{d}x\,\mathrm{d}y\,. \end{aligned} \tag{3.8}$$

We can rewrite the integrals in (3.8) as sums. For a locally integrable function $\varphi : \mathbb{R}^2 \to \mathbb{R}$, we define $\varphi_N : \mathbb{Z}^2_{\mathrm{even}} \to \mathbb{R}$ as the average of $\varphi(\frac{\cdot}{\sqrt{N}})$ over cells $B(v) \subseteq \mathbb{R}^2$, see (3.5):

$$\varphi_N(v) := \frac{1}{|B(v)|} \int\limits_{B(v)} \varphi\big(\tfrac{x}{\sqrt{N}}\big)\,\mathrm{d}x = \frac{1}{2} \int\limits_{\{|x_1-v_1|+|x_2-v_2|<1\}} \varphi\big(\tfrac{x}{\sqrt{N}}\big)\,\mathrm{d}x\,. \tag{3.9}$$

If we similarly define $\psi_N : \mathbb{Z}^2_{\mathrm{even}} \to \mathbb{R}$ given $\psi : \mathbb{R}^2 \to \mathbb{R}$, we can rewrite the second line of (3.8) as a sum over the points $v = [\![x]\!], w = [\![y]\!] \in \mathbb{Z}^2_{\mathrm{even}}$ as follows:

$$\mathcal{Z}^{\beta}_{N,t}(\varphi,\psi) \;=\; \frac{1}{N} \sum_{v,w\,\in\,\mathbb{Z}^2_{\mathrm{even}}} \varphi_N(v)\; Z^{\beta}_{[\![Nt]\!]}(v,w)\; \psi_N(w)\,. \tag{3.10}$$

Remark 3.1 (Parity issue). *Let $\mathbb{Z}^d_{\mathrm{odd}} := \mathbb{Z}^d \backslash \mathbb{Z}^d_{\mathrm{even}}$. If in* (3.10) *we sum over $v, w \in \mathbb{Z}^2_{\mathrm{odd}}$, we obtain an alternative "odd version" of the averaged partition function, which is* independent of the "even version" *because two simple random walks started at even vs. odd sites can never meet. This explains why we enforce a parity restriction in* (3.10).

Finally, we recall the *critical window* of the disorder strength (inverse temperature) that was introduced in (1.11). Given the definition (3.2) of R_N, for some fixed $\vartheta \in \mathbb{R}$, we choose $\beta = \beta_N = \beta_N(\vartheta)$ such that

$$\sigma_N^2 := \mathbb{V}\mathrm{ar}[e^{\beta_N\omega-\lambda(\beta_N)}] = e^{\lambda(2\beta_N)-2\lambda(\beta_N)} - 1 = \frac{1}{R_N}\left(1 + \frac{\vartheta + o(1)}{\log N}\right). \tag{3.11}$$

We can spell out this condition more explicitly in terms of β_N (see [CSZ19a, Appendix A.4]):

$$\begin{aligned}\beta_N^2 &= \frac{1}{R_N} - \frac{\kappa_3}{(R_N)^{3/2}} + \frac{\vartheta/\pi + (\frac{3}{2}\kappa_3^2 - \frac{7}{12}\kappa_4 - \frac{1}{2})}{(R_N)^2} + o\Big(\frac{1}{(R_N)^2}\Big) \\ &= \frac{\pi}{\log N} - \frac{\kappa_3\,\pi^{3/2}}{(\log N)^{3/2}} + \frac{\pi(\vartheta-\alpha) + \pi^2(\frac{3}{2}\kappa_3^2 - \frac{7}{12}\kappa_4 - \frac{1}{2})}{(\log N)^2} + o\Big(\frac{1}{(\log N)^2}\Big)\,,\end{aligned} \tag{3.12}$$

where κ_3, κ_4 are the disorder cumulants, i.e. $\lambda(\beta) = \frac{1}{2}\beta^2 + \frac{\kappa_3}{3!}\beta^3 + \frac{\kappa_4}{4!}\beta^4 + o(\beta^4)$ as $\beta \downarrow 0$, and $\alpha \simeq 0.208$ is as in (3.2). *Henceforth we always set $\beta = \beta_N$.*

3.3. Polynomial chaos expansion. We now recall the *polynomial chaos expansion* of the partition function. This is based on the following product expansion, valid for any set A and any family of real numbers $(h_n)_{n\in A}$ labelled by A:

$$e^{\sum_{n\in A} h_n} \;=\; \prod_{n\in A}\big(1 + (e^{h_n} - 1)\big) \;=\; 1 + \sum_{\varnothing \neq B \subseteq A}\;\prod_{n\in B}(e^{h_n} - 1)\,. \tag{3.13}$$

If we apply (3.13) to the partition function $Z_{d,f}^{\beta_N}(x,y)$ in (1.8), by (3.1) we obtain

$$\begin{aligned}&Z_{d,f}^{\beta_N}(x,y) \;-\; q_{d,f}(x,y) \\ &\quad= \mathrm{E}\Big[\Big(e^{\sum_{n=d+1}^{f-1}\sum_{z\in\mathbb{Z}^2}(\beta_N\omega(n,z) - \lambda(\beta_N))\mathbb{1}_{S_n = z}} - 1\Big)\,\mathbb{1}_{S_f = y}\,\Big|\,S_d = x\Big] \\ &\quad= \sum_{r=1}^{\infty}\;\sum_{\substack{d<n_1<\ldots<n_r<f \\ z_1,\ldots,z_r\in\mathbb{Z}^2}}\mathrm{E}\Big[\Big\{\prod_{j=1}^{r}\Big(e^{(\beta_N\omega(n_j,z_j) - \lambda(\beta_N))\mathbb{1}_{S_{n_j} = z_j}} - 1\Big)\Big\}\,\mathbb{1}_{S_f = y}\,\Big|\,S_d = x\Big]\,.\end{aligned}$$

Recalling (3.11), we introduce a family $(\xi_N(n,z))_{(n,z)\in\mathbb{Z}^2}$ of i.i.d. random variables by

$$\begin{aligned}&\xi_N(n,z) := e^{\beta_N\omega(n,z) - \lambda(\beta_N)} - 1 \\ &\text{so that}\qquad \mathbb{E}[\xi_N(n,z)] = 0\,,\qquad \mathbb{V}\mathrm{ar}[\xi_N(n,z)] = \sigma_N^2\,.\end{aligned} \tag{3.14}$$

These variables allow us to write

$$e^{(\beta_N\omega(n,z) - \lambda(\beta_N))\mathbb{1}_{S_n = z}} - 1 = (e^{\beta_N\omega(n,z) - \lambda(\beta_N)} - 1)\,\mathbb{1}_{S_n = z} = \xi_N(n,z)\,\mathbb{1}_{S_n = z}\,,$$

hence, by the Markov property for the random walk with kernel q, we get

$$\begin{aligned}Z_{d,f}^{\beta_N}(x,y) \;=\; q_{d,f}(x,y) + \sum_{r=1}^{\infty}\;&\sum_{\substack{d<n_1<\ldots<n_r<f \\ z_1,\ldots,z_r\in\mathbb{Z}^2}} \\ &q_{d,n_1}(x,z_1)\,\xi_N(n_1,z_1)\Big\{\prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)\,\xi_N(n_j,z_j)\Big\}\,q_{n_j,f}(z_j,y)\,,\end{aligned} \tag{3.15}$$

where $\prod_{j=2}^{r}(\ldots) := 1$ if $r = 1$. We have expressed the point-to-point partition function as a multilinear polynomial (*polynomial chaos*) in the independent random variables $\xi_N(n,z)$.

A similar polynomial chaos representation holds for the *averaged* partition function $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)$ given in (3.10). To simplify notation, it is convenient to define an averaged version of the random walk transition kernel $q_{m,n}(x,y)$. Given suitable $\varphi, \psi : \mathbb{R}^2 \to \mathbb{R}$, a

time horizon $M \in (0,\infty)$, and two points $(m,w),(n,z) \in \mathbb{Z}^3_{\text{even}}$, recalling φ_N and ψ_N from (3.9), we define

$$q^N_{0,m}(\varphi,w) := \sum_{v \in \mathbb{Z}^2_{\text{even}}} \varphi_N(v)\, q_{0,m}(v,w)\,, \tag{3.16}$$

$$q^N_{n,M}(z,\psi) := \sum_{w \in \mathbb{Z}^2_{\text{even}}} q_{n,[\![M]\!]}(z,w)\, \psi_N(w)\,, \tag{3.17}$$

$$q^N_{0,M}(\varphi,\psi) = \frac{1}{N} \sum_{v,w \in \mathbb{Z}^2_{\text{even}}} \varphi_N(v)\, q_{0,[\![M]\!]}(v,w)\, \psi_N(w)\,. \tag{3.18}$$

Then (3.15) yields the following polynomial chaos expansion for $\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi)$ in (3.10):

$$\begin{aligned} \mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi) \;=\; q^N_{0,Nt}(\varphi,\psi) \;+\; \frac{1}{N} \sum_{r=1}^{\infty} \sum_{\substack{0<n_1<\ldots<n_r<Nt \\ z_1,\ldots,z_r \in \mathbb{Z}^2}} \\ q^N_{0,n_1}(\varphi,z_1)\, \xi_N(n_1,z_1) \cdot \Big\{ \prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)\, \xi_N(n_j,z_j) \Big\} \cdot q^N_{n_r,[\![Nt]\!]}(z_r,\psi)\,. \end{aligned} \tag{3.19}$$

As will be explained later, when it comes to second moment calculations, the time-space points $(n_1,z_1),\ldots,(n_r,z_r)$ in the sum can be interpreted as a time-space renewal configuration.

3.4. Random walk estimates.

Let $g_t : \mathbb{R}^2 \to (0,\infty)$ denote the heat kernel on $\mathbb{R}^2$:

$$g_t(x) := \frac{1}{2\pi t}\, e^{-\frac{|x|^2}{2t}}\,, \qquad g_t(x,y) := g_t(y-x)\,, \tag{3.20}$$

where, unless otherwise specified, we denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^d$.

The asymptotic behavior of the random walk transition kernel $q_n(x) = \mathrm{P}(S_n = x)$ is given by the local central limit theorem: as $n \to \infty$ we have, uniformly for $x \in \mathbb{Z}^2$,

$$\begin{aligned} q_n(x) \;&=\; \big(g_{\frac{n}{2}}(x) + O\big(\tfrac{1}{n^2}\big)\big)\, 2\, \mathbb{1}_{(n,x) \in \mathbb{Z}^3_{\text{even}}} \\ &=\; g_{\frac{n}{2}}(x)\, e^{O\left(\frac{1}{n}\right) + O\left(\frac{|x|^4}{n^3}\right)}\, 2\, \mathbb{1}_{(n,x) \in \mathbb{Z}^3_{\text{even}}}\,, \end{aligned} \tag{3.21}$$

where the two lines are two different variants of the local central limit theorem for the simple symmetric random walk on $\mathbb{Z}^2$ given by Theorems 2.3.5 and 2.3.11 in [LaLi10]. We recall that $\mathbb{Z}^d_{\text{even}}$ is defined in (3.4), the multiplicative factor 2 comes from the periodicity of the simple random walk $S_n = (S^{(1)}_n, S^{(2)}_n)$ on $\mathbb{Z}^2$, while the factor $\frac{1}{2}$ in the time argument of the heat kernel comes from the fact that $\mathrm{E}[S^{(i)}_n S^{(j)}_n] = \frac{n}{2}\, \mathbb{1}_{i=j}$. We also note that

$$g_{\frac{n}{2}}(x) = \tfrac{1}{N}\, g_{\frac{1}{2}\frac{n}{N}}\big(\tfrac{x}{\sqrt{N}}\big) \qquad \forall n, N \in \mathbb{N},\ \forall x \in \mathbb{Z}^2\,. \tag{3.22}$$

Similar to the averaged random walk kernels $q^N_{\cdot,\cdot}$ defined in (3.16)-(3.18), given $\varphi \in L^1(\mathbb{R}^2)$, $\psi \in L^\infty(\mathbb{R}^2)$, $t > 0$, and $a, b \in \mathbb{R}^2$, we define the averaged heat kernels

$$g_t(\varphi, a) := \int_{\mathbb{R}^2} \varphi(x)\, g_t(a - x)\, \mathrm{d}x\,, \tag{3.23}$$

$$g_t(b, \psi) := \int_{\mathbb{R}^2} g_t(y - b)\, \psi(y)\, \mathrm{d}y\,, \tag{3.24}$$

$$g_t(\varphi, \psi) := \int_{\mathbb{R}^2 \times \mathbb{R}^2} \varphi(x)\, g_t(y - x)\, \psi(y)\, \mathrm{d}x\, \mathrm{d}y\,. \tag{3.25}$$

Recall $q^N_{0,Nt}(\varphi, \psi)$ from (3.18). By the local limit theorem (3.21), recalling (3.9) and (3.22), we have

$$\forall t > 0 : \qquad \lim_{N\to\infty} q^N_{0,Nt}(\varphi, \psi) \;=\; \tfrac{1}{2}\, g_{\frac{t}{2}}(\varphi, \psi)\,, \tag{3.26}$$

where the prefactor $\frac{1}{2}$ is due to periodicity.

We will also need the following lemma, which allows us to replace a random walk transition kernel by a heat kernel even if the time-space increments are perturbed.

Lemma 3.2. *Let $q_n(\cdot)$ be the transition kernel of the simple symmetric random walk on $\mathbb{Z}^2$, see (3.1), and let $g_t(\cdot)$ be the heat kernel on $\mathbb{R}^2$, see (3.20). Then there exists $C \in (0, \infty)$ such that, for all $n \in \mathbb{N}$ and for all $x \in \mathbb{Z}^2$ with $|x| \leqslant n^{\frac{3}{4}}$, we have*

$$q_n(x) \leqslant C\, g_{\frac{n}{2}}(x). \tag{3.27}$$

Let $\varrho_1, \varrho_2 > 0$ and set $C := 2e\, \varrho_1\, \varrho_2$. Then, given an arbitrary $m \in \mathbb{N}$, for all $n_1, n_2 \in \mathbb{N}$ with $n_1 \geqslant m$ and $\frac{n_2}{n_1} \in [1/\varrho_1, \varrho_2]$, and for all $x_1, x_2 \in \mathbb{R}^2$ with $|x_1 - x_2| \leqslant \sqrt{m}$, we have

$$g_{\frac{n_1}{2}}(x_1) \leqslant C g_{\varrho_1 n_2}(x_2) = \frac{C}{m} g_{\frac{\varrho_1 n_2}{m}}\Big(\frac{x_2}{\sqrt{m}}\Big). \tag{3.28}$$

Proof. Let us prove (3.27): by the second variant of the local limit theorem in (3.21),

$$q_{n_1}(x_1) = \mathbb{1}_{\{(n_1,x_1)\in\mathbb{Z}^3_{\text{even}}\}}\, 2\, g_{\frac{n_1}{2}}(x_1)\, \exp\Big\{O\Big(\frac{1}{n_1}\Big) + O\Big(\frac{|x_1|^4}{n_1^3}\Big)\Big\} \leqslant C\, g_{\frac{n_1}{2}}(x_1)\,.$$

We next prove (3.28): by the assumption $\frac{n_2}{n_1} \in [1/\varrho_1, \varrho_2]$, we have

$$\frac{g_{\frac{n_1}{2}}(x_1)}{g_{\varrho_1 n_2}(x_2)} = \frac{2\varrho_1 n_2}{n_1} \exp\Big\{\frac{|x_2|^2}{2\varrho_1 n_2} - \frac{|x_1|^2}{n_1}\Big\} \leqslant 2\varrho_1\varrho_2\, \exp\Big\{\frac{|x_2|^2}{2n_1} - \frac{|x_1|^2}{n_1}\Big\} \leqslant 2\varrho_1\varrho_2\, e\,,$$

where the last inequality holds because $|x_2|^2 \leqslant 2(|x_1|^2 + |x_2 - x_1|^2) \leqslant 2|x_1|^2 + 2m$ and $n_1 \geqslant m$ by assumption. □

3.5. Renewal estimates and Dickman subordinator. We next present the time-space renewal process underlying the second moment calculations for the partition function. Under diffusive scaling, this leads to the so-called Dickman subordinator in the continuum limit. This approach was developed in [CSZ19a, CSZ19b].

We first define a slight modification of the partition function $Z^\beta_{d,f}(x, y)$ in (1.8), where we "attach" disorder variables $\xi_N(n, z)$, see (3.14), at the boundary points (d, x) and (f, y)

(which may coincide, if $d = f$):

$$X^{\beta_N}_{d,f}(x,y) := \begin{cases} \xi_N(d,x)\,\mathbb{1}_{\{y=x\}} & \text{if } f = d \\ \xi_N(d,x)\,Z^{\beta_N}_{d,f}(x,y)\,\xi_N(f,y) & \text{if } f \geqslant d+1 \end{cases}. \tag{3.29}$$

Such quantities will appear as basic building blocks in our proofs. Note that $\mathbb{E}[X^{\beta_N}_{d,f}(x,y)] = 0$. The second moment of $X^{\beta_N}_{d,f}(x,y)$ can be computed explicitly by the polynomial chaos expansion (3.15) and it can be expressed as follows:

$$\mathbb{E}\big[X^{\beta_N}_{d,f}(x,y)^2\big] = \sigma_N^2\, U_N(f-d, y-x)\,, \tag{3.30}$$

where we recall that $\sigma_N^2 = \mathbb{V}\mathrm{ar}(\xi_N(a,x))$, and for $n \in \mathbb{N}_0 = \{0,1,2,\ldots\}$ and $x \in \mathbb{Z}^2$ we define

$$U_N(n,x) := \begin{cases} \mathbb{1}_{\{x=0\}} & \text{if } n = 0, \\ \sigma_N^2\, q(n,x)^2 + \displaystyle\sum_{r=1}^{\infty} (\sigma_N^2)^{r+1} \sum_{\substack{0<n_1<\ldots<n_r<n \\ z_1,\ldots,z_r\in\mathbb{Z}^2}} q_{0,n_1}(0,z_1)^2 \times & \\ \qquad \times \Big\{ \displaystyle\prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)^2 \Big\}\, q_{n_r,n}(z_r,x)^2 & \text{if } n \geqslant 1. \end{cases} \tag{3.31}$$

The quantity $U_N(n,x)$, which plays an important role throughout this paper, admits a probabilistic interpretation as a *renewal function*. More precisely, let $(\tau_r^{(N)}, S_r^{(N)})_{r\geqslant 0}$ denote the random walk (time-space renewal process) on $\mathbb{N}_0 \times \mathbb{Z}^2$ starting at $(0,0)$ and with one-step distribution

$$\mathrm{P}\big(\tau_1^{(N)} = n,\, S_1^{(N)} = x\big) = \frac{q_n(x)^2}{R_N}\,\mathbb{1}_{\{1,\ldots,N\}}(n)\,, \tag{3.32}$$

where R_N is the random walk overlap defined in (3.2). Then we can write, recalling (3.11),

$$\begin{aligned} U_N(n,x) &= \sum_{r=1}^{\infty} (\lambda_N)^r\, \mathrm{P}\big(\tau_r^{(N)} = n,\, S_r^{(N)} = x\big) \\ \text{where} \quad \lambda_N &:= \sigma_N^2\, R_N = 1 + \frac{\vartheta + o(1)}{\log N}\,. \end{aligned} \tag{3.33}$$

When $\lambda_N = 1$, we see that $U_N(n,x)$ is just the renewal function of $(\tau_r^{(N)}, S_r^{(N)})_{r\geqslant 0}$. When $\lambda_N \neq 1$, we can think of $U_N(n,x)$ as an exponentially weighted renewal function, weighted according to the number of renewals. Note that the first component $\tau^{(N)} = (\tau_r^{(N)})_{r\geqslant 0}$ is a renewal process with one-step distribution

$$\mathrm{P}\big(\tau_1^{(N)} = n\big) = \frac{u(n)}{R_N}\,\mathbb{1}_{\{1,\ldots,N\}}(n)\,, \tag{3.34}$$

where $u(n) = \sum_x q_n(x)^2$ is defined in (3.3). Correspondingly, we can define

$$U_N(n) := \sum_{x\in\mathbb{Z}^2} U_N(n,x) = \sum_{r=1}^{\infty} (\lambda_N)^r\, \mathrm{P}\big(\tau_r^{(N)} = n\big)\,. \tag{3.35}$$

The asymptotic behaviors of $U_N(n,x)$ and $U_N(n)$ were obtained in [CSZ19a], exploiting the fact that $\tau^{(N)}$ is in the domain of attraction of the so-called *Dickman subordinator*, defined as the pure jump Lévy process with Lévy measure $\frac{1}{x}\,\mathbb{1}_{(0,1)}(x)\,\mathrm{d}x$. More precisely, we

have the following convergence result, which is an extension of [CSZ19a, Proposition 2.2] from finite dimensional distribution convergence to process level convergence.

Lemma 3.3. *Let $(\tau_r^{(N)}, S_r^{(N)})_{r\geqslant 0}$ be the space-time random walk defined in* (3.32)*. Let $(Y_s)_{s\geqslant 0}$ be the so-called Dickman subordinator* [CSZ19a]*, i.e. the pure jump Lévy process with Lévy measure $\frac{1}{t}\mathbb{1}_{(0,1)}(t)\mathrm{d}t$, and let $V_s := \frac{1}{2}W_{Y_s}$ where W is an independent Brownian motion. Then we have the convergence in distribution*

$$\left(\frac{\tau^{(N)}_{\lfloor s\log N\rfloor}}{N}, \frac{S^{(N)}_{\lfloor s\log N\rfloor}}{\sqrt{N}}\right)_{s\geqslant 0} \underset{N\to\infty}{\Longrightarrow} (\boldsymbol{Y}_s)_{s\geqslant 0} := (Y_s, V_s)_{s\geqslant 0}, \tag{3.36}$$

on the space of càdlàg paths equipped with the Skorohod topology.

Proof. Denote $\boldsymbol{Y}_s^{(N)} = (Y_s^{(N)}, V_s^{(N)}) := \left(\frac{\tau^{(N)}_{\lfloor s\log N\rfloor}}{N}, \frac{S^{(N)}_{\lfloor s\log N\rfloor}}{\sqrt{N}}\right)$. The convergence of finite dimensional distributions was already proved in [CSZ19a, Proposition 2.2]. We prove tightness by verifying Aldous' tightness criterion [Kal97, Theorem 14.11], namely that for any bounded sequence of stopping times τ_N with respect to $(\boldsymbol{Y}_s^{(N)})_{s\geqslant 0}$ and any positive constants $h_N \downarrow 0$, we have $\boldsymbol{Y}^{(N)}_{\tau_N+h_N} - \boldsymbol{Y}^{(N)}_{\tau_N} \to 0$ in probability as $N\to\infty$. This follows immediately from the fact that the increments of $\boldsymbol{Y}^{(N)}$ are i.i.d. and $\boldsymbol{Y}^{(N)}_{h_N} \to (0,0)$ in probability as $N\to\infty$. □

For $\vartheta \in (0,\infty)$, we define the exponentially weighted Green's function for $\boldsymbol{Y} = (\boldsymbol{Y}_s)_{s\geqslant 0}$:

$$\widehat{G}_\vartheta(t,x) = \int_0^\infty e^{\vartheta s}\boldsymbol{f}_s(t,x)\mathrm{d}s, \tag{3.37}$$

where $\boldsymbol{f}_s(\cdot,\cdot)$ is the density of the law of $\boldsymbol{Y}_s$ on $[0,\infty)\times\mathbb{R}^2$, given that $\boldsymbol{Y}_0 = (0,0)$ (we take notation from (3.36)). It was shown in [CSZ19a] that

$$\widehat{G}_\vartheta(t,x) := \widehat{G}_\vartheta(t)\, g_{\frac{t}{4}}(x)\,, \tag{3.38}$$

where $g_\cdot(\cdot)$ is the heat kernel, see (3.20), and $\widehat{G}_\vartheta(t) := \int_{\mathbb{R}^2}\widehat{G}_\vartheta(t,x)\mathrm{d}x$ is closely related to the so-called Dickman function in number theory. For $t\leqslant 1$, it can be computed explicitly as

$$\widehat{G}_\vartheta(t) = G_\vartheta(t) := \int_0^\infty \frac{e^{(\vartheta-\gamma)s}\, s\, t^{s-1}}{\Gamma(s+1)}\,\mathrm{d}s\,, \tag{3.39}$$

with γ as in (3.2) (see [CSZ19a][†]). We will also denote $G_\vartheta(t,x) := G_\vartheta(t)\, g_{\frac{t}{4}}(x)$. Note that for $t\leqslant 1$, $G_\vartheta(t,x)$ and $G_\vartheta(t)$ are the continuum analogues of $U_N(n,x)$ and $U_N(n)$, respectively. It is therefore no surprise that the asymptotics of U_N will be expressed in terms of G_ϑ, which we record below for later use.

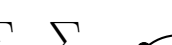
In light of (3.30), it is convenient to define

$$\overline{U}_N(n,x) := \sigma_N^2\, U_N(n,x)\,, \qquad \overline{U}_N(n) := \sigma_N^2\, U_N(n) = \sum_{x\in\mathbb{Z}^2}\overline{U}_N(n,x)\,. \tag{3.40}$$

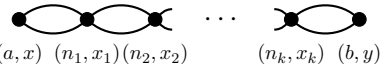
[†]In [CSZ19a], there was no separate notation $\widehat{G}_\vartheta$ for the weighted Green's function, which might cause some confusion.

Recalling (3.31), we can give a graphical representation for $\overline{U}_N(b-a, y-x)$ as follows:

$$\overline{U}_N(b-a,y-x) \equiv \qquad (3.41)$$



$$:= \sum_{k\geqslant 1} \sum_{\substack{n_1,\dots,n_k \\ x_1,\dots,x_k}}$$



where in the second line we assign weights $q_{n'-n}(x'-x)$ to any solid line going from (n,x) to (n',x') and we assign weight σ_N^2 to every solid dot.

Recall that $\sigma_N^2 \sim \frac{\pi}{\log N}$, see (3.11) and (3.2). We now rephrase some results from [CSZ19a]. Fix $T>0$.

- By [CSZ19a, Theorem 1.4], for any fixed $\delta>0$, as $N\to\infty$ we have
$$\overline{U}_N(n) = \frac{\pi}{N}\big(G_\vartheta\big(\tfrac{n}{N}\big)+o(1)\big) \qquad \text{uniformlyfor} \;\; \delta N \leqslant n \leqslant TN\,, \qquad (3.42)$$
and moreover there is $C<\infty$ such that
$$\overline{U}_N(n) \leqslant \frac{C}{N} G_\vartheta\big(\tfrac{n}{N}\big) \qquad \forall 0<n\leqslant TN\,. \qquad (3.43)$$
- By [CSZ19a, Theorem 2.3 and 3.7], for any fixed $\delta>0$, as $N\to\infty$ we have
$$\overline{U}_N(n,x) = \frac{\pi}{N^2}\big(G_\vartheta\big(\tfrac{n}{N},\tfrac{x}{\sqrt{N}}\big)+o(1)\big)\, 2\,\mathbb{1}_{(n,x)\in\mathbb{Z}^3_{\text{even}}} \quad \text{uniformly for} \;\; \delta N\leqslant n\leqslant TN \;\; \text{and} \;\; |x|\leqslant \tfrac{1}{\delta}\sqrt{N}\,. \qquad (3.44)$$
The prefactor 2 is due to periodicity and, moreover, there is $C<\infty$ such that
$$\overline{U}_N(n,x) \leqslant \frac{C}{N}\frac{1}{n} G_\vartheta\big(\tfrac{n}{N}\big) \qquad \forall 0<n\leqslant TN\,,\; \forall x\in\mathbb{Z}^2\,. \qquad (3.45)$$
- By [CSZ19a, Proposition 1.6], for $t\in(0,1]$ the function $G_\vartheta(t)$ is C^∞ and strictly positive, and as $t\downarrow 0$ it has the following asymptotic behavior:
$$G_\vartheta(t) = \frac{1}{t(\log\frac{1}{t})^2}\bigg\{1+\frac{2\vartheta}{\log\frac{1}{t}}+O\bigg(\frac{1}{(\log\frac{1}{t})^2}\bigg)\bigg\}, \qquad (3.46)$$
hence as $t\downarrow 0$
$$\int_0^t G_\vartheta(s)\mathrm{d}s = \frac{1}{\log\frac{1}{t}}\bigg\{1+\frac{\vartheta}{\log\frac{1}{t}}+O\bigg(\frac{1}{(\log\frac{1}{t})^2}\bigg)\bigg\}. \qquad (3.47)$$

Remark 3.4. *In the proof of* (3.42)-(3.45)*, the case $T>1$ has to be treated differently from $T=1$. In* [CSZ19a]*, the case $T>1$ was reduced to $T=1$ through a renewal decomposition and recursion (see* [CSZ19a, Section 7]*). Alternatively, we can reduce the case $T>1$ to $T=1$ by first setting $\widetilde{N}:=TN$, $\tilde{\vartheta}:=\vartheta+\log T+o(1)$ so that $\sigma_N^2=\sigma_N^2(\vartheta)=\sigma_{\widetilde{N}}^2(\tilde{\vartheta})$ by their definitions in* (3.11)*, and then applying* (3.42)-(3.45) *with N replaced by $\widetilde{N}$, using the observation that $\frac{1}{T}G_{\vartheta+\log T}(\frac{t}{T})=G_\vartheta(t)$.*

We will also need the following bound to complement (3.44).

Lemma 3.5. *There exists $c\in(0,\infty)$ such that for all $\lambda\geqslant 0$ and $0\leqslant n\leqslant N$,*
$$\sum_{x\in\mathbb{Z}^2} \overline{U}_N(n,x)\, e^{\lambda|x|} \;\leqslant\; c\, e^{c\lambda^2 n}\, \overline{U}_N(n)\,. \qquad (3.48)$$

Note that by the Markov inequality and optimisation over $\lambda > 0$, (3.48) implies that the probability kernel $\overline{U}_N(n,\cdot)/\overline{U}_N(n)$ has Gaussian decay on the spatial scale $\sqrt{n}$.

Proof. Recall the definition of $U_N(n,x)$ from (3.33). Conditioned on $\tau_1^{(N)},\dots,\tau_r^{(N)}$ with $\tau_r^{(N)} = n$, we can write $S_r^{(N)} = \zeta_1 + \cdots + \zeta_r$ for independent ζ_i's with

$$\mathrm{P}(\zeta_i = x) = \frac{q_{n_i}(x)^2}{\sum_{y\in\mathbb{Z}^2} q_{n_i}(y)^2}\,,$$

where $n_i := \tau_i^{(N)} - \tau_{i-1}^{(N)}$ and $x \in \mathbb{Z}^2$. For each i, denote by $\zeta_{i,1}$ and $\zeta_{i,2}$ the two components of $\zeta_i \in \mathbb{Z}^2$. Then we note that there exists $c > 0$ such that for any $\lambda \geqslant 0$, $n_i \in \mathbb{N}$,

$$\mathrm{E}[e^{\pm\lambda\zeta_{i,j}}] \leqslant e^{c\lambda^2 n_i}, \qquad j = 1,2\,. \tag{3.49}$$

This can be seen by Taylor expanding the exponential and using that $\mathrm{E}[\zeta_{i,\cdot}] = 0$ by symmetry, $|\,\mathrm{E}[\zeta_{i,\cdot}^{2k+1}]| \leqslant \frac{1}{2}(\mathrm{E}[\xi_{i,\cdot}^{2k}] + \mathrm{E}[\zeta_{i,\cdot}^{2k+2}])$ by Young's inequality, as well as $\mathrm{E}[\zeta_{i,\cdot}^{2k}] \leqslant (Cn_i)^k(2k-1)!!$ for some $C > 0$ uniformly in $n_i, k \in \mathbb{N}$. The bound on $\mathrm{E}[\zeta_{i,\cdot}^{2k}]$ holds because by (3.21),

$$\mathrm{P}(\zeta_i = x) = \frac{q_{n_i}(x)^2}{q_{2n_i}(0)} \leqslant \left\{\frac{\sup_{x\in\mathbb{Z}^2} q_{n_i}(x)}{q_{2n_i}(0)}\right\} q_{n_i}(x) \leqslant C'\, q_{n_i}(x)\,,$$

where q_{n_i} has the same Gaussian tail decay as the heat kernel $g_{n_i/2}$. Using $e^{|x|} \leqslant e^x + e^{-x}$, this then implies

$$\begin{aligned}
\mathrm{E}\left[e^{\lambda|S_r^{(N)}|}\big|\tau_\cdot^{(N)}\right] = \mathrm{E}\left[e^{\lambda|\sum_{i=1}^r \zeta_i|}\right] &\leqslant \mathrm{E}\left[e^{2\lambda|\sum_{i=1}^r \zeta_{i,1}|}\right]^{\frac{1}{2}} \mathrm{E}\left[e^{2\lambda|\sum_{i=1}^r \zeta_{i,2}|}\right]^{\frac{1}{2}} \\
&\leqslant \prod_{j=1,2}\left(\mathrm{E}\left[e^{2\lambda\sum_{i=1}^r \zeta_{i,j}}\right] + \mathrm{E}\left[e^{-2\lambda\sum_{i=1}^r \zeta_{i,j}}\right]\right)^{\frac{1}{2}} \\
&= \prod_{j=1,2}\left(\prod_{i=1}^r \mathrm{E}\left[e^{2\lambda\zeta_{i,j}}\right] + \prod_{i=1}^r \mathrm{E}\left[e^{-2\lambda\zeta_{i,j}}\right]\right)^{\frac{1}{2}} \\
&\leqslant 2e^{4c\lambda^2 n}.
\end{aligned}$$

The bound (3.48) then follows readily from the definitions of $U_N(n,x)$ and $U_N(n)$ in (3.33) and (3.35), recalling that $\overline{U}_N(n,x)$ and $\overline{U}_N(n)$ are defined in (3.40). □

3.6. Second moment of averaged partition function. Using $X_{d,f}^{\beta_N}(x,y)$ as introduced in (3.29), and recalling (3.15), we can now rewrite the chaos expansion for the averaged partition function $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)$ in (3.19) as follows:

$$\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi) \;=\; q_{0,Nt}^N(\varphi,\psi) + \frac{1}{N}\sum_{\substack{0<d\leqslant f<Nt \\ x,y\in\mathbb{Z}^2}} q_{0,d}^N(\varphi,x)\, X_{d,f}^{\beta_N}(x,y)\, q_{f,Nt}^N(y,\psi)\,, \tag{3.50}$$

so that by (3.30) and the fact that $\overline{U}_N := \sigma_N^2 U_N$, we have

$$\mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)^2\big] \;=\; q_{0,Nt}^N(\varphi,\psi)^2 + \frac{1}{N^2}\sum_{\substack{x,y\in\mathbb{Z}^2 \\ 0<d\leqslant f<Nt}} q_{0,d}^N(\varphi,x)^2\, \overline{U}_N(f-d,y-x)\, q_{f,Nt}^N(y,\psi)^2. \tag{3.51}$$

We now compute the limit of $\mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)^2\big]$ as $N\to\infty$. This was first obtained for the Stochastic Heat Equation in [BC98] in the special case $\psi\equiv 1$; see also [CSZ19b, Theorems 1.2 and 1.7] for an alternative derivation, that also includes directed polymers.

Proposition 3.6 (First and second moments). *Recall $G_\vartheta(t)$ from (3.39) for all $t>0$. For $\varphi:\mathbb{R}^2\to\mathbb{R}$, define*

$$\|\varphi\|_{\mathcal{G}_t}^2 := \iint\limits_{\mathbb{R}^2\times\mathbb{R}^2} \varphi(z)\,\mathcal{G}_t(z'-z)\,\varphi(z')\,\mathrm{d}z\,\mathrm{d}z',^{\dagger} \qquad \text{where} \quad \mathcal{G}_t(x) := \int_0^t g_s(x)\,\mathrm{d}s\,. \tag{3.52}$$

Then for all φ with $\|\varphi\|_{\mathcal{G}_t}<\infty$ and all $\psi\in L^\infty(\mathbb{R}^2)$, we have

$$\lim_{N\to\infty}\mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)\big] \;=\; \tfrac{1}{2}\,g_{\frac{t}{2}}(\varphi,\psi)\,, \tag{3.53}$$

$$\lim_{N\to\infty}\mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)^2\big] \;=\; \tfrac{1}{4}\,g_{\frac{t}{2}}(\varphi,\psi)^2 \;+\; \tfrac{1}{2}\mathscr{V}_t^{\vartheta}(\varphi,\psi)\,, \tag{3.54}$$

where

$$\begin{aligned}\mathscr{V}_t^{\vartheta}(\varphi,\psi) \;&=\; \iiiint\limits_{(\mathbb{R}^2)^4} \varphi(z)\,\varphi(z')\,K_t^{\vartheta}(z,z';w,w')\,\psi(w)\,\psi(w')\,\mathrm{d}z\,\mathrm{d}z'\,\mathrm{d}w\,\mathrm{d}w' \\ &\leqslant\; \pi\,\|\psi\|_\infty^2\,\|\varphi\|_{\mathcal{G}_t}^2\int_0^t G_\vartheta(u)\,\mathrm{d}u\,,\end{aligned} \tag{3.55}$$

and the kernel K_t^{ϑ} is defined by

$$\begin{aligned}K_t^{\vartheta}(z,z';w,w') \;:=\; \pi\; g_{\frac{t}{4}}\big(\tfrac{w+w'}{2}-\tfrac{z+z'}{2}\big) \\ \times \iint\limits_{0<s<u<t} g_s(z'-z)\,G_\vartheta(u-s)\,g_{t-u}(w'-w)\,\mathrm{d}s\,\mathrm{d}u\,.\end{aligned} \tag{3.56}$$

Proof. The first moment convergence (3.53) holds because by $\mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)\big]=q_{0,Nt}^N(\varphi,\psi)$, see (3.50), in view of the asymptotic relation (3.26).

For the second moment computation (3.54) we exploit (3.51), where the first term in the r.h.s. converges to $\frac{1}{4}\,g_{t/2}(\varphi,\psi)^2$ by (3.26), which matches the first term in the r.h.s. of (3.54). It remains to show that the sum in (3.51) converges to the term $\frac{1}{2}\mathscr{V}_t^{\vartheta}(\varphi,\psi)$ in (3.54).

Recall the definition of $q_\cdot^N$ in (3.16)-(3.17). By the local limit theorem (3.21) and in view of (3.22), we see that for any $\varepsilon>0$, uniformly for $m>\varepsilon N$ and $w\in\mathbb{Z}^2$, we have as $N\to\infty$

$$q_{0,m}^N(\varphi,w) \;=\; \Big(g_{\frac{1}{2}\frac{m}{N}}\big(\varphi,\tfrac{w}{\sqrt{N}}\big) \;+\; o(1)\Big)\,\mathbb{1}_{(m,w)\in\mathbb{Z}^3_{\mathrm{even}}}\,,$$

and similarly, uniformly for $n\leqslant(1-\varepsilon)Nt$ and $z\in\mathbb{Z}^2$,

$$q_{n,Nt}^N(z,\psi) \;=\; \Big(g_{\frac{1}{2}(t-\frac{n}{N})}\big(\tfrac{z}{\sqrt{N}},\psi\big) \;+\; o(1)\Big)\,\mathbb{1}_{(n,z)\in\mathbb{Z}^3_{\mathrm{even}}}\,.$$

†The positivity of $\|\varphi\|_{\mathcal{G}_t}^2$ can be seen via Fourier transform.

Applying the asymptotic relation (3.44) for $\overline{U}_N(f-d, y-x)$, we see that the sum in (3.51) is a Riemann sum that converges as $N \to \infty$ to the multiple integral[†]

$$\frac{1}{2}\mathscr{V}_t^{\vartheta}(\varphi,\psi) \;:=\; \frac{\pi}{2} \iiiint\limits_{\substack{0<s<u<t\\ a,b\in\mathbb{R}^2}} g_{\frac{s}{2}}(\varphi,a)^2\, G_{\vartheta}(u-s,b-a)\, g_{\frac{t-u}{2}}(b,\psi)^2 \mathrm{d}s\,\mathrm{d}u\,\mathrm{d}a\,\mathrm{d}b\,, \tag{3.57}$$

where the prefactor $\frac{1}{2}$ results from combining the periodicity factor 2 in (3.44) with the volume factor $\frac{1}{2}\cdot\frac{1}{2}$ which originates from the restrictions $(d,x),(f,y)\in\mathbb{Z}^3_{\mathrm{even}}$ in (3.51). Then it follows by (3.23), (3.24) and (3.38) that the equality in (3.55) holds with

$$\begin{aligned} K_t^{\vartheta}(z,z';w,w') \;=\; \pi \iiiint\limits_{\substack{0<s<u<t\\ a,b\in\mathbb{R}^2}} &\big\{g_{\frac{s}{2}}(a-z)\, g_{\frac{s}{2}}(a-z')\big\}\, G_{\vartheta}(u-s)\, g_{\frac{u-s}{4}}(b-a) \\ &\times \big\{g_{\frac{t-u}{2}}(w-b)\, g_{\frac{t-u}{2}}(w'-b)\big\}\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}a\,\mathrm{d}b\,. \end{aligned}$$

We can simplify both brackets via the identity $g_t(x)\, g_t(y) = g_{2t}(x-y)\, g_{\frac{t}{2}}(\frac{x+y}{2})$, see (3.20). Performing the integrals over $a,b\in\mathbb{R}^2$ we then obtain (3.56).

The bound in (3.55) follows by bounding ψ with $\|\psi\|_\infty$ and then successively integrating out w, w', followed by u and s in (3.56). □

Remark 3.7 (Point-to-plane partition function). *For $\psi(w) = \mathbb{1}(w) \equiv 1$, we can view $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\mathbb{1})$ as the point-to-plane partition function $Z_N^{\beta_N}(z)$ in* (1.3) *averaged over its starting point z. By* (3.53)-(3.56)*,*

$$\lim_{N\to\infty} \mathbb{E}\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\mathbb{1})\big] = \frac{1}{2}\, g_{\frac{t}{2}}(\varphi,1) = \frac{1}{2}\int_{\mathbb{R}^2} \varphi(z)\,\mathrm{d}z\,,$$

$$\lim_{N\to\infty} \mathbb{V}\mathrm{ar}\,\big[\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\mathbb{1})\big] = \frac{1}{2}\mathscr{V}_t^{\vartheta}(\varphi,1) = \frac{1}{2}\iint_{(\mathbb{R}^2)^2} \varphi(z)\,\varphi(z')\, K_t^{\vartheta}(z-z')\mathrm{d}z\mathrm{d}z'\,,$$

where we set

$$K_t^{\vartheta}(x) := \pi \iint\limits_{0<s<u<t} g_s(x)\, G_{\vartheta}(u-s)\,\mathrm{d}s\,\mathrm{d}u\,.$$

We note that both the asymptotic mean and the asymptotic variance of $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\mathbb{1})$ are half of those obtained in [CSZ19b, eq. (1.19)-(1.20)]*. This is because here we have defined $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)$ as a sum over $\mathbb{Z}^2_{\mathrm{even}}$, see* (3.10)*, while in* [CSZ19b]*, the sum is over both $\mathbb{Z}^2_{\mathrm{odd}}$ and $\mathbb{Z}^2_{\mathrm{even}}$, which give rise to two i.i.d. limits as $N\to\infty$ by the parity of the simple random walk on $\mathbb{Z}^2$.*

4. Coarse-graining

In this section, we give the details of how to coarse-grain the averaged partition function and what is the precise definition of the coarse-grained model, which were outlined in Section 2. The main result is Theorem 4.7, which shows that the averaged partition function $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)$, see (3.10), can be approximated in L^2 by the coarse-grained model.

[†]The contributions to the sum in (3.51) given by $m \leqslant \varepsilon N$ and $n > (1-\varepsilon)Nt$ are small when $\varepsilon > 0$ is small, uniformly in large N, as can be checked using the uniform bound (3.27).

4.1. Preparation. The starting point is the polynomial chaos expansion (3.19) for the averaged partition function $\mathcal{Z}_{N,t}^{\beta_N}(\varphi,\psi)$, which is a multilinear polynomial in the disorder variables $\xi_N(n,z)$. We will call the sequence of time-space points $(n_1,z_1),\dots,(n_r,z_r)\in\mathbb{N}\times\mathbb{Z}^2$ in the sum in (3.19) a *microscopic (time-space) renewal configuration*. We assume that the disorder strength is chosen to be $\beta_N=\beta_N(\vartheta)$ as defined in (3.11)-(3.12). For simplicity, we assume the time horizon to be tN with $t=1$.

Given $\varepsilon\in(0,1)$ and $N\in\mathbb{N}$, we partition discrete time-space $\{1,\dots,N\}\times\mathbb{Z}^2$ into *mesoscopic boxes*

$$\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a}) \;:=\; \underbrace{\big((\mathsf{i}-1)\varepsilon N,\mathsf{i}\varepsilon N\big]}_{\mathcal{T}_{\varepsilon N}(\mathsf{i})}\times\underbrace{\big((\mathsf{a}-(1,1))\sqrt{\varepsilon N},\mathsf{a}\sqrt{\varepsilon N}\big]}_{\mathcal{S}_{\varepsilon N}(\mathsf{a})}\;\cap\;\mathbb{Z}^3_{\text{even}}\,, \tag{4.1}$$

where $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ is mesoscopic time interval and $\mathcal{S}_{\varepsilon N}(\mathsf{a})$ a mesoscopic spatial square.[†] These boxes are indexed by *mesoscopic variables*

$$(\mathsf{i},\mathsf{a})\in\{1,\dots,\lfloor\tfrac{1}{\varepsilon}\rfloor\}\times\mathbb{Z}^2\,.$$

Recall from Section 2 that to carry out the coarse-graining, we need to organize the chaos expansion (3.19) according to which mesoscopic boxes $\mathcal{B}_{\varepsilon N}$ are visited by the microscopic renewal configuration $(n_1,z_1),\dots,(n_r,z_r)$. To perform the kernel replacement (2.6), which allows each summand in the chaos expansion (3.19) to factorize into a product of coarse-grained disorder variables $\Theta_{N,\varepsilon}$ connected by heat kernels, we will impose some constraints on the set of visited mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\cdot)$ and spatial boxes $\mathcal{S}_{\varepsilon N}(\cdot)$, which will be shown to have negligible costs in L^2. We first introduce the necessary notation.

Let us fix two thresholds

$$K_\varepsilon:=(\log\tfrac{1}{\varepsilon})^6\,,\qquad M_\varepsilon:=\log\log\tfrac{1}{\varepsilon}\,. \tag{4.2}$$

We will require that the visited mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\dots,\mathcal{T}_{\varepsilon N}(\mathsf{i}_k)$ belong to

$$\begin{aligned}\mathcal{A}_\varepsilon^{(\text{no triple})} \;:=\; \bigcup_{k\in\mathbb{N}}\Big\{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathbb{N}^k : K_\varepsilon\leqslant\mathsf{i}_1<\mathsf{i}_2<\ldots<\mathsf{i}_k\leqslant\lfloor\tfrac{1}{\varepsilon}\rfloor-K_\varepsilon \;\text{ such that}\\ \text{if }\;\mathsf{i}_{j+1}-\mathsf{i}_j<K_\varepsilon\,,\;\text{then }\;\mathsf{i}_{j+2}-\mathsf{i}_{j+1}\geqslant K_\varepsilon\;\Big\}\,.\end{aligned} \tag{4.3}$$

We call this the *no-triple condition*, since it forbids three consecutive mesoscopic time indices $\mathsf{i}_j,\mathsf{i}_{j+1},\mathsf{i}_{j+2}$ with both $\mathsf{i}_{j+1}-\mathsf{i}_j<K_\varepsilon$ and $\mathsf{i}_{j+2}-\mathsf{i}_{j+1}<K_\varepsilon$. We can then partition $(\mathsf{i}_1,\dots,\mathsf{i}_k)$ into time blocks such that $\mathsf{i}_j,\mathsf{i}_{j+1}$ belong to the same block whenever $\mathsf{i}_{j+1}-\mathsf{i}_j<K_\varepsilon$.

Definition 4.1 (Time block). *We call a* time block *any pair* $\vec{\mathsf{i}}=(\mathsf{i},\mathsf{i}')\in\mathbb{N}\times\mathbb{N}$ *with* $\mathsf{i}\leqslant\mathsf{i}'$. *The width of a time block is*

$$|\vec{\mathsf{i}}|:=\mathsf{i}'-\mathsf{i}+1\,.$$

The (non symmetric) "distance" between two time blocks $\vec{\mathsf{i}},\vec{\mathsf{m}}$ *is defined by*

$$\operatorname{dist}(\vec{\mathsf{i}},\vec{\mathsf{m}}):=\mathsf{m}-\mathsf{i}'\qquad\textit{for }\;\vec{\mathsf{i}}=(\mathsf{i},\mathsf{i}')\;\textit{ and }\;\vec{\mathsf{m}}=(\mathsf{m},\mathsf{m}')\,,$$

and we write "$\vec{\mathsf{i}}<\vec{\mathsf{m}}$*" to mean that "*$\vec{\mathsf{i}}$ *precedes* $\vec{\mathsf{m}}$*":*

$$\vec{\mathsf{i}}<\vec{\mathsf{m}}\qquad\Longleftrightarrow\qquad\operatorname{dist}(\vec{\mathsf{i}},\vec{\mathsf{m}})>0\;\;\textit{i.e.}\;\;\mathsf{i}'<\mathsf{m}\,.$$

[†]We use the notation $(\mathsf{a}-\mathsf{b},\mathsf{a}]=(\mathsf{a}_1-\mathsf{b}_1,\mathsf{a}_1]\times(\mathsf{a}_2-\mathsf{b}_2,\mathsf{a}_2]$ for squares in $\mathbb{R}^2$.

With the partitioning of the indices $(\mathsf{i}_1, \dots, \mathsf{i}_k)$ of the visited mesoscopic time intervals into consecutive time blocks as defined above, which we denote by $\vec{\mathsf{i}}_1 = (\mathsf{i}_1, \mathsf{i}_1')$, $\dots$, $\vec{\mathsf{i}}_r = (\mathsf{i}_r, \mathsf{i}_r')$ with possibly $\mathsf{i}_\ell = \mathsf{i}_\ell'$, the constraint $\mathcal{A}_\varepsilon^{(\text{no triple})}$ then becomes the following:

$$\vec{\mathcal{A}}_\varepsilon^{(\text{no triple})} := \bigcup_{r\in\mathbb{N}} \Big\{ \text{time blocks} \quad K_\varepsilon \leqslant \vec{\mathsf{i}}_1 < \dots < \vec{\mathsf{i}}_r \leqslant \lfloor \tfrac{1}{\varepsilon} \rfloor - K_\varepsilon \quad \text{such that} \\ |\vec{\mathsf{i}}_j| \leqslant K_\varepsilon \ \ \forall j = 1, \dots, r\,, \quad \operatorname{dist}(\vec{\mathsf{i}}_{j-1}\,,\, \vec{\mathsf{i}}_j) \geqslant K_\varepsilon \ \ \forall j = 2, \dots, r \Big\}. \tag{4.4}$$

If the time horizon is Nt with $t \neq 1$, then in (4.3) and (4.4) we just replace the upper bound $\vec{\mathsf{i}}_r \leqslant \lfloor \frac{1}{\varepsilon} \rfloor - K_\varepsilon$ by $\vec{\mathsf{i}}_r \leqslant \lfloor \frac{t}{\varepsilon} \rfloor - K_\varepsilon$.

Given a time block $\vec{\mathsf{i}} = (\mathsf{i}, \mathsf{i}')$ with $\mathsf{i}' - \mathsf{i} + 1 \leqslant K_\varepsilon$ (possibly $\mathsf{i} = \mathsf{i}'$), which identifies two mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ and $\mathcal{T}_{\varepsilon N}(\mathsf{i}')$ visited by the microscopic renewal configuration $(n_1, z_1), \dots, (n_r, z_r)$ from (3.19) and no intervals in-between is visited, we can identify the first and last mesoscopic spatial boxes visited in the time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ and $\mathcal{T}_{\varepsilon N}(\mathsf{i}')$, respectively. We call this pair of mesoscopic spatial indices a space block.

Definition 4.2 (Space block). *We call a* space block *any pair* $\vec{\mathsf{a}} = (\mathsf{a}, \mathsf{a}') \in \mathbb{Z}^2 \times \mathbb{Z}^2$. *The width of a space block is*

$$|\vec{\mathsf{a}}| := |\mathsf{a}' - \mathsf{a}|\,,$$

with $|\cdot|$ *being the Euclidean norm. The (non symmetric) "distance" between two space blocks* $\vec{\mathsf{a}}, \vec{\mathsf{b}}$ *is*

$$\operatorname{dist}(\vec{\mathsf{a}}\,,\, \vec{\mathsf{b}}) := |\mathsf{b} - \mathsf{a}'| \qquad \textit{for} \ \ \vec{\mathsf{a}} = (\mathsf{a}, \mathsf{a}') \ \ \textit{and} \ \ \vec{\mathsf{b}} = (\mathsf{b}, \mathsf{b}')\,.$$

Putting the time block and space block together, we have the following.

Definition 4.3 (Time-space block). *We call a* time-space block *any pair* $(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ *where* $\vec{\mathsf{i}}$ *is a time block and* $\vec{\mathsf{a}}$ *is a space block. We also define*

$$\mathbb{T}_\varepsilon := \Big\{ \textit{time-space blocks} \ \ (\vec{\mathsf{i}}, \vec{\mathsf{a}}) \ \ \textit{with} \ \ |\vec{\mathsf{i}}| \leqslant K_\varepsilon \ \ \textit{and} \ \ |\vec{\mathsf{a}}| \leqslant M_\varepsilon \sqrt{|\vec{\mathsf{i}}|} \Big\}. \tag{4.5}$$

In (3.19), we will restrict to $(n_1, z_1), \dots, (n_r, z_r)$ (interpreted as a time-space renewal configuration) that satisfy condition (4.3), so that they determine a sequence of mesoscopic time-space blocks $\{(\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1), \dots, (\vec{\mathsf{i}}_r, \vec{\mathsf{a}}_r)\} \in \vec{\mathcal{A}}_\varepsilon^{(\text{no triple})}$. This would give the main contribution in (3.19). We now impose further constraints on the spatial components that still capture the main contribution.

Given two "boundary variables" $\mathsf{b}, \mathsf{c} \in \mathbb{Z}^2$ and a sequence of time blocks $(\vec{\mathsf{i}}_1, \dots, \vec{\mathsf{i}}_r)$, we denote by $\vec{\mathcal{A}}_{\varepsilon;\,\mathsf{b},\mathsf{c}}^{(\text{diff})} = \vec{\mathcal{A}}_{\varepsilon;\,\mathsf{b},\mathsf{c}}^{(\text{diff})}(\vec{\mathsf{i}}_1, \dots, \vec{\mathsf{i}}_r)$ the following subset of space blocks $(\vec{\mathsf{a}}_1, \dots, \vec{\mathsf{a}}_r)$, where we impose *diffusive constraints* on their widths and distances:

$$\vec{\mathcal{A}}_{\varepsilon;\,\mathsf{b},\mathsf{c}}^{(\text{diff})} := \Big\{ \text{space blocks} \ \ \vec{\mathsf{a}}_1, \dots, \vec{\mathsf{a}}_r \quad \text{such that} \ \ |\vec{\mathsf{a}}_j| \leqslant M_\varepsilon \sqrt{|\vec{\mathsf{i}}_j|} \quad \forall j = 1, \dots, r\,, \\ \operatorname{dist}(\vec{\mathsf{a}}_{j-1}, \vec{\mathsf{a}}_j) \leqslant M_\varepsilon \sqrt{\operatorname{dist}(\vec{\mathsf{i}}_{j-1}\,,\, \vec{\mathsf{i}}_j)} \quad \forall j = 2, \dots, r\,, \\ |\mathsf{a}_1 - \mathsf{b}| \leqslant M_\varepsilon \sqrt{\mathsf{i}_1} \quad \text{and} \quad |\mathsf{c} - \mathsf{a}_r'| \leqslant M_\varepsilon \sqrt{\lfloor \tfrac{1}{\varepsilon} \rfloor - \mathsf{i}_r'} \Big\}. \tag{4.6}$$

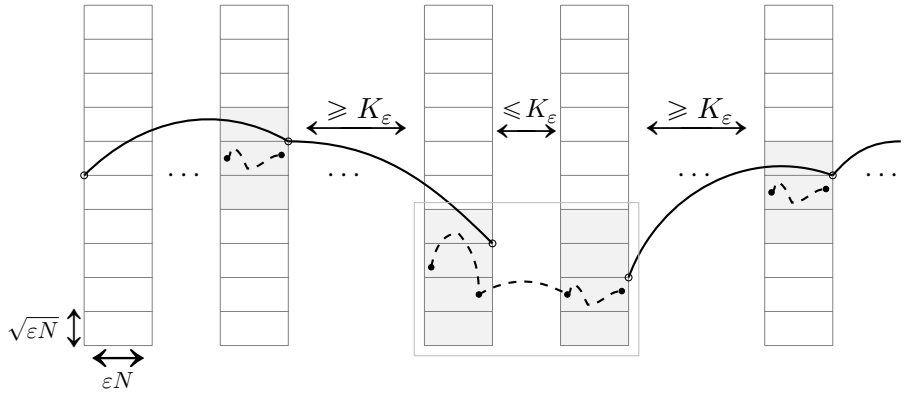


FIGURE 2. An illustration of the coarse-graining procedure. The solid lines represent the heat kernels after the kernel replacement (2.6), which connect adjacent coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ consisting of sums over the dashed lines in each visited time-space block $\mathcal{B}_{\varepsilon N}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ (see (4.11)). The solid and the dashed lines satisfy the diffusive constraint given in $\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\,\mathsf{b},\mathsf{c}}$ and (4.11), respectively.

Given a sequence of mesoscopic time-space blocks $(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1)$, …, $(\vec{\mathsf{i}}_r,\vec{\mathsf{a}}_r)$ determined by the microscopic renewal configuration $(n_1,z_1),\dots,(n_r,z_r)$ from (3.19), which satisfies the constraints $\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}$ and $\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\,\mathsf{b},\mathsf{c}}$, we will perform the kernel replacement (2.6), which leads to a factorization of each summand in (3.19) as the product of coarse-grained disorder variables $\Theta_{N,\varepsilon}(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_r)$, $1\leqslant j\leqslant r$, connected by the heat kernels $g_{\frac12(\mathsf{i}_{j+1}-\mathsf{i}_j)}(\mathsf{a}_{j+1}-\mathsf{a}'_j)$. See Figure 2.

4.2. Coarse-grained model. We are now ready to give the precise definition of the *coarse-grained model* $\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta)$ given earlier in (2.7), which depends on $\varepsilon\in(0,1)$ and is a multilinear polynomial of a given family of random variables $\Theta=\{\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}})\}$ indexed by time-space blocks $(\vec{\mathsf{i}},\vec{\mathsf{a}})$.

Definition 4.4 (Coarse-grained model). *Fix $\varepsilon\in(0,1)$ and a family of random variables $\Theta=(\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}}))_{(\vec{\mathsf{i}},\vec{\mathsf{a}})\in\mathbb{T}_\varepsilon}$ indexed by the set $\mathbb{T}_\varepsilon$ of time-space blocks defined in* (4.5)*. Fix two locally integrable functions $\varphi,\psi:\mathbb{R}^2\to\mathbb{R}$ and define $\varphi_\varepsilon,\psi_\varepsilon:\mathbb{Z}^2\to\mathbb{R}$ as follows:*

$$\varphi_\varepsilon(\mathsf{b}):=\int_{(\mathsf{b}-(1,1),\mathsf{b}]}\varphi(\sqrt{\varepsilon}x)\,\mathrm{d}x\,,\qquad \psi_\varepsilon(\mathsf{c}):=\int_{(\mathsf{c}-(1,1),\mathsf{c}]}\psi(\sqrt{\varepsilon}y)\,\mathrm{d}y\qquad \text{for } \mathsf{b},\mathsf{c}\in\mathbb{Z}^2\,. \tag{4.7}$$

Recall the heat kernel $g_t(\cdot)$ from (3.20)*. Then the* coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta)$ *is*

$$\begin{aligned}\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta):=\frac12\,g_{\frac12}(\varphi,\psi)\;+\;\frac{\varepsilon}{2}\sum_{r=1}^{(\log\frac1\varepsilon)^2}\sum_{\mathsf{b},\mathsf{c}\in\mathbb{Z}^2}\sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}\\(\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\,b,c}}}\varphi_\varepsilon(\mathsf{b})g_{\frac12\mathsf{i}_1}(\mathsf{a}_1-\mathsf{b})\Theta(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1)\\ \times\left\{\prod_{j=2}^{r}g_{\frac12(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\,\Theta(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j)\right\}g_{\frac12(\lfloor\frac1\varepsilon\rfloor-\mathsf{i}'_r)}(\mathsf{c}-\mathsf{a}'_r)\psi_\varepsilon(\mathsf{c})\,.\end{aligned} \tag{4.8}$$

Note that in (4.8), for technical reasons that will become clear later (to control the error induced by kernel replacements - see Section 5.3), we also imposed the constraint that the number of time-space blocks cannot exceed $(\log \frac{1}{\varepsilon})^2$. This coarse-grained model has the same structure as the original averaged partition function $\mathscr{Z}_{N,1}^{\beta_N}(\varphi,\psi)$ in (3.19), with $1/\varepsilon$ replacing N, Θ replacing ξ_N, and the heat kernels replacing the random walk kernels. Note that when φ has compact support, (4.8) is a sum over finitely many terms.

Remark 4.5. *To approximate the averaged partition function $\mathscr{Z}_{N,t}^{\beta_N}(\varphi,\psi)$ with $t \neq 1$, we define a corresponding coarse-grained model $\mathscr{Z}_{\varepsilon,t}^{(\mathrm{cg})}(\varphi,\psi|\Theta)$ which is obtained from (4.8) simply replacing $g_{\frac12}(\varphi,\psi)$ by $g_{\frac t2}(\varphi,\psi)$ and $g_{\frac12(\lfloor\frac{1}{\varepsilon}\rfloor-\mathsf{i}'_r)}$ by $g_{\frac12(\lfloor\frac{t}{\varepsilon}\rfloor-\mathsf{i}'_r)}$, as well as modifying accordingly $\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon}^{(\text{no triple})}$ and $\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\,b,c}^{(\mathrm{diff})}$ (replacing $\lfloor\frac{1}{\varepsilon}\rfloor$ by $\lfloor\frac{t}{\varepsilon}\rfloor$ therein).*

4.3. Coarse-grained disorder variables. We now identify the coarse-grained disorder variables $\Theta_{N,\varepsilon}^{(\mathrm{cg})}$ so that the averaged partition function $\mathscr{Z}_{N,t}^{\beta_N}(\varphi,\psi)$ can be approximated in L^2 by the coarse-grained model $\mathscr{Z}_{\varepsilon,t}^{(\mathrm{cg})}(\varphi,\psi|\Theta)$ with $\Theta=\Theta_{N,\varepsilon}^{(\mathrm{cg})}$.

Recall the point-to-point partition function $Z_{d,f}^{\beta_N}(x,y)$ with its chaos expansion as in (3.15). Assuming $f-d\leqslant \varepsilon N$, we introduce a diffusive truncation as follows, the effect of which will be negligible in L^2, but it ensures that the coarse-grained disorder variable $\Theta_{N,\varepsilon}^{(\mathrm{cg})}$ will only depend on $\xi_N(n,z)$ in a localized time-space window. In (3.15), let $\mathsf{a}=\mathsf{a}(x)\in\mathbb{Z}^2$ be such that $x\in\mathcal{S}_{\varepsilon N}(\mathsf{a})$ (recall (4.1)). We then restrict y and all space variables z_j in (3.15) to those mesoscopic boxes $\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}})$ with $|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon=\log\log\frac{1}{\varepsilon}$ as in (4.2), and define

$$Z_{d,f}^{(\mathrm{diff})}(x,y) := \begin{cases} 0 & \text{if } y\notin \bigcup\limits_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}}),\\[2ex] q_{d,f}(x,y)+\sum\limits_{r=1}^{\infty}\sum\limits_{\substack{d<n_1<\ldots<n_r<f\\ z_1,\ldots,z_r\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}})}} & \\ \quad q_{d,n_1}(x,z_1)\,\xi_N(n_1,z_1)\times & \text{if } y\in \bigcup\limits_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}}).\\ \quad\Big\{\prod\limits_{j=2}^{r}q_{n_{j-1},n_j}(z_{j-1},z_j)\,\xi_N(n_j,z_j)\Big\}\,q_{n_j,f}(z_j,y) & \end{cases} \tag{4.9}$$

Similar to the definition of $X_{d,f}^{\beta_N}(x,y)$ in (3.29), we define

$$X_{d,f}^{(\mathrm{diff})}(x,y) := \begin{cases}\xi_N(d,x) & \text{if } f=d\\ \xi_N(d,x)\,Z_{d,f}^{(\mathrm{diff})}(x,y)\,\xi_N(f,y) & \text{if } f\geqslant d+1\end{cases}. \tag{4.10}$$

Note that we omit the dependence of $Z_{d,f}^{(\mathrm{diff})}(x,y)$ and $X_{d,f}^{(\mathrm{diff})}(x,y)$ on N,ε.

The coarse-grained disorder variables $\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ are defined as follows (see Figure 2).

Definition 4.6 (Coarse-grained disorder variable). *Given $N\in\mathbb{N}$, $\varepsilon\in(0,1)$ and a time-space block $(\vec{\mathsf{i}},\vec{\mathsf{a}})$, with $\vec{\mathsf{i}}=(\mathsf{i},\mathsf{i}')$ and $\vec{\mathsf{a}}=(\mathsf{a},\mathsf{a}')$, the associated* coarse-grained disorder

variable $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ *is defined by*

$$\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}}) := \begin{cases} \dfrac{2}{\varepsilon N} \displaystyle\sum_{\substack{(d,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})\\(f,y)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a}')\\ \text{with } d\leqslant f}} X^{(\mathrm{diff})}_{d,f}(x,y) & \text{if } |\vec{\mathsf{i}}|=1\,, \\ \dfrac{2}{\varepsilon N} \displaystyle\sum_{\substack{(d,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})\\(f',y')\in\mathcal{B}_{\varepsilon N}(\mathsf{i}',\mathsf{a}')}} \sum_{\substack{\mathsf{b}:\,|\mathsf{b}-\mathsf{a}|\leqslant M_\varepsilon\\ \mathsf{b}':\,|\mathsf{b}'-\mathsf{a}'|\leqslant M_\varepsilon\\ \text{such that}\\ |\mathsf{b}'-\mathsf{b}|\leqslant M_\varepsilon\sqrt{\mathsf{i}'-\mathsf{i}}}} \sum_{\substack{(f,y)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{b})\\(d',x')\in\mathcal{B}_{\varepsilon N}(\mathsf{i}',\mathsf{b}')\\ \text{such that}\\ d\leqslant f,\,d'\leqslant f'}} & \text{if } |\vec{\mathsf{i}}|>1\,. \\ X^{(\mathrm{diff})}_{d,f}(x,y)\; q_{f,d'}(y,x')\; X^{(\mathrm{diff})}_{d',f'}(x',y') & \end{cases} \tag{4.11}$$

In the special case $|\vec{\mathsf{i}}|=1$*, i.e.,* $\mathsf{i}=\mathsf{i}'$*, we will also write* $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{i};\mathsf{a},\mathsf{a}')$ *in place of* $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$.

We point out that the prefactor 2 in (4.11) is due to periodicity, because the sums are restricted to $\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})\subseteq\mathbb{Z}^3_{\mathrm{even}}$, see (4.1).

4.4. Coarse-graining approximation. We can finally state the key result of this section, which approximates the averaged partition function $\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi)$ in L^2 by the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon,t}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$, with an error which is much smaller than $\mathbb{V}\mathrm{ar}(\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi))$ in (3.55) when N is large and ε is small. Recall $\|\cdot\|_{\mathcal{G}_t}$ from (3.52).

Theorem 4.7 (Coarse-graining). *Let* $\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi)$ *be the averaged partition function in* (3.8)*, where* $\beta_N=\beta_N(\vartheta)$ *satisfies* (3.11) *for some fixed* $\vartheta\in\mathbb{R}$*. Let* $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon,t}(\varphi,\psi|\Theta)$ *be the coarse-grained model from* (4.8)*, with* $K_\varepsilon=(\log\frac{1}{\varepsilon})^6$ *and* $M_\varepsilon=\log\log\frac{1}{\varepsilon}$ *as in* (4.2)*, and let* $\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}})=\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ *be the coarse-grained disorder variables from Definition 4.6. Then, for any* $T\in(0,\infty)$*, there exists* $\mathsf{C}=\mathsf{C}(T)<\infty$ *such that, for* $\varepsilon>0$ *small enough, we have*

$$\limsup_{N\to\infty}\; \big\|\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi)\;-\;\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon,t}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})\big\|^2_{L^2}\;\leqslant\;\mathsf{C}\left(\|\varphi\|^2_{\mathcal{G}_{K_\varepsilon\varepsilon}}+\frac{\|\varphi\|^2_{\mathcal{G}_T}}{\sqrt{\log\frac{1}{\varepsilon}}}\right)\|\psi\|^2_\infty\,, \tag{4.12}$$

uniformly in $t\in[0,T]$*,* $\psi\in L^\infty(\mathbb{R}^2)$ *and* $\varphi:\mathbb{R}^2\to\mathbb{R}$ *with* $\|\varphi\|_{\mathcal{G}_T}<\infty$.

Note that the r.h.s. of (4.12) tends to 0 as $\varepsilon\downarrow 0$, because $K_\varepsilon\,\varepsilon\to 0$. The whole of Section 5 is devoted to the proof of Theorem 4.7.

5. Second moment bounds for averaged partition functions

This section is devoted mainly to the proof of Theorem 4.7, which approximates the averaged partition function $\mathcal{Z}^{\beta_N}_{N,t}(\varphi,\psi)$ from (3.10) by the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon,t}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$ from (4.8). We may assume $t=T=1$ without loss of generality. The uniformity in $t\leqslant T$ will be clear from the proof. *Throughout this section we simply write* $\mathcal{Z}_N(\varphi,\psi)$*, omitting the dependence on* $t=1$ *and on* β_N.

The starting point of our proof of Theorem 4.7 is the polynomial chaos expansion (3.19). In the second moment calculations, the time-space renewal representation and the limiting Dickman subordinator presented in Section 3.5 play a crucial role. The proof will be carried

out in three steps, presented in Subsections 5.1-5.3 below: given $\varepsilon > 0$, we introduce two intermediate approximations $\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi)$ and $\mathcal{Z}_{N,\varepsilon}^{(\text{diff})}(\varphi,\psi)$ of the averaged partition function $\mathcal{Z}_N(\varphi,\psi)$, and bound the following in L^2.

- Step 1. We bound $\mathcal{Z}_N(\varphi,\psi) - \mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi)$, see Lemma 5.1 in Section 5.1: this is the cost of imposing the constraints $\mathcal{A}_\varepsilon^{(\text{no triple})}$ and $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$, see (4.3) and (4.4).
- Step 2. We bound $\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi) - \mathcal{Z}_{N,\varepsilon}^{(\text{diff})}(\varphi,\psi)$, see Lemma 5.6 in Section 5.2: this is the cost of imposing diffusive constraints, including $\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\,b,c}^{(\text{diff})}$ in (4.6) and the diffusive truncation in the definition of $\Theta_{N,\varepsilon}^{(\text{cg})}$ in (4.11).
- Step 3. We bound $\mathcal{Z}_{N,\varepsilon}^{(\text{diff})}(\varphi,\psi) - \mathscr{Z}_\varepsilon^{(\text{cg})}(\varphi,\psi|\Theta_{N,\varepsilon})$, see Lemma 5.8 in Section 5.3: this is the cost of the kernel replacement (2.6).

Combining Lemmas 5.1, 5.6 and 5.8 then gives Theorem 4.7.

In the last Subsection 5.4, we will prove a separate second moment estimate for the coarse-grained model, which is needed later in Section 8 for higher moment bounds.

The proof details in this section are technically a bit heavy and could be skipped in a first reading.

5.1. Step 1: Constraints on mesoscopic time variables. In this step, we introduce our first approximation $\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi)$ and show that it is close to $\mathcal{Z}_N(\varphi,\psi)$.

Recall the mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\mathsf{i}) := ((\mathsf{i}-1)\varepsilon N, \mathsf{i}\varepsilon N]$ introduced in (4.1), to which we associate the mesoscopic time index $\mathsf{i} \in \{1,\dots,\frac{1}{\varepsilon}\}$. In the chaos expansion for $\mathcal{Z}_N^{\beta_N}(\varphi,\psi)$ in (3.19), each time index n_j belongs to $\mathcal{T}_{\varepsilon N}(\mathsf{i})$ for some $\mathsf{i} \in \{1,\dots,\frac{1}{\varepsilon}\}$. The first step of coarse-graining is to group the terms in the expansion in (3.19) in terms of the mesoscopic time intervals $\mathcal{T}_{\varepsilon N}(\cdot)$ visited by the sequence of time indices $n_1,\dots,n_r$. Namely, we can rewrite (3.19) as (omitting β_N from $\mathcal{Z}_N^{\beta_N}(\varphi,\psi)$, and expanding $q_{0,n_1}^N(\varphi,z_1)$ and $q_{n_r,N}^N(z_r,\psi)$ according to their definitions in (3.16)-(3.17)):

$$\begin{aligned}\mathcal{Z}_N(\varphi,\psi) \;&=\; q_{0,N}^N(\varphi,\psi)\\ &\quad+\;\frac{1}{N}\sum_{v,w\in\mathbb{Z}^2_{\text{even}}}\varphi_N(v)\sum_{k=1}^{\infty}\sum_{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}}\sum_{\substack{d_1\leqslant f_1\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\,\dots,\,d_k\leqslant f_k\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_k)\\ x_1,\,y_1,\,\dots,x_k,\,y_k\in\mathbb{Z}^2}}\\ &\quad q_{0,d_1}(v,x_1)\,X_{d_1,f_1}(x_1,y_1)\left\{\prod_{j=2}^{k}q_{f_{j-1},d_j}(y_{j-1},x_j)\,X_{d_j,f_j}(x_j,y_j)\right\}q_{f_k,N}(y_k,w)\,\psi_N(w)\,,\end{aligned}\tag{5.1}$$

where φ_N, ψ_N were defined in (3.9), and $X_{d,f}(x,y)$ was defined in (3.29).

Recall from (4.2) that $K_\varepsilon = (\log\frac{1}{\varepsilon})^6$. We will show that in (5.1), the dominant contribution (in L^2) comes from mesoscopic time variables $(\mathsf{i}_1,\dots,\mathsf{i}_k)$ which contains *no consecutive triples* $\mathsf{i}_j,\mathsf{i}_{j+1},\mathsf{i}_{j+2}$ *with both* $\mathsf{i}_{j+1}-\mathsf{i}_j < K_\varepsilon$ *and* $\mathsf{i}_{j+2}-\mathsf{i}_{j+1} < K_\varepsilon$. This is encoded in $\mathcal{A}_\varepsilon^{(\text{no triple})}$ from

(4.3), which we recall here

$$\mathcal{A}_\varepsilon^{(\text{no triple})} := \bigcup_{k\in\mathbb{N}} \Big\{ (\mathsf{i}_1,\dots,\mathsf{i}_k) \in \mathbb{N}^k : K_\varepsilon \leqslant \mathsf{i}_1 < \mathsf{i}_2 < \ldots < \mathsf{i}_k \leqslant \lfloor \tfrac{1}{\varepsilon} \rfloor - K_\varepsilon \ \text{ such that } \\ \text{if } \ \mathsf{i}_{j+1} - \mathsf{i}_j < K_\varepsilon \,, \ \text{then } \ \mathsf{i}_{j+2} - \mathsf{i}_{j+1} \geqslant K_\varepsilon \ \Big\}. \tag{5.2}$$

We will further restrict the sum in (5.1) to $(\mathsf{i}_1,\dots,\mathsf{i}_k)$ with $k \leqslant (\log \frac{1}{\varepsilon})^2$, which leads to the following first approximation of $\mathcal{Z}_N(\varphi,\psi)$:

$$\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi) := q_{0,N}^N(\varphi,\psi) \ + \ \frac{1}{N} \sum_{v,w\in\mathbb{Z}^2_{\text{even}}} \varphi_N(v) \\ \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}_\varepsilon^{(\text{no triple})}} \sum_{\substack{d_1\leqslant f_1\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\dots,d_k\leqslant f_k\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_k) \\ x_1,y_1,\dots,x_k,y_k\in\mathbb{Z}^2}} \\ q_{0,d_1}(v,x_1)\, X_{d_1,f_1}(x_1,y_1) \Bigg\{ \prod_{j=2}^{k} q_{f_{j-1},d_j}(y_{j-1},x_j)\, X_{d_j,f_j}(x_j,y_j) \Bigg\} q_{f_k,N}(y_k,w)\, \psi_N(w)\,. \tag{5.3}$$

The main result of this subsection is the following approximation, which constitutes part of the bound in (4.12). The proof is a bit lengthy, but it contains many important ingredients, including a key renewal interpretation of second moment bounds.

Lemma 5.1 (No close triples). *Recall from* (4.2) *that* $K_\varepsilon = (\log\frac{1}{\varepsilon})^6$ *and recall* $\|\cdot\|_{\mathcal{G}_t}$ *from* (3.52). *There exists* $\mathsf{C}\in(0,\infty)$ *such that for* $\varepsilon>0$ *small enough, we have: for all* φ *with* $\|\varphi\|_{\mathcal{G}_1}^2 < \infty$ *and* $\psi\in L^\infty(\mathbb{R}^2)$,

$$\limsup_{N\to\infty \text{ with } N\in 2\mathbb{N}} \big\| (\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})} - \mathcal{Z}_N)(\varphi,\psi) \big\|_{L^2}^2 \ \leqslant \ \mathsf{C} \left(\|\varphi\|_{\mathcal{G}_{K_\varepsilon\varepsilon}}^2 + \frac{(\log K_\varepsilon)^2}{\log\frac{1}{\varepsilon}} \|\varphi\|_{\mathcal{G}_1}^2 \right) \|\psi\|_\infty^2 . \tag{5.4}$$

Proof. The random variables $X_{d,f}(x,y)$ depend on the disorder variables $\xi_N(n,x)$ for $d\leqslant n\leqslant f$, see (3.29). They are centered and orthogonal in L^2 and, by (3.30) and (3.40),

$$\mathbb{E}[X_{d,f}(x,y)\, X_{d',f'}(x',y')] = \mathbb{1}_{\{(d,f,x,y)=(d',f',x',y')\}}\, \overline{U}_N(f-d,y-x)\,.$$

Since the sum which defines $\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}(\varphi,\psi)$ *is a subset of that of* $\mathcal{Z}_N(\varphi,\psi)$, cf. (5.1) and (5.3), it follows that we can write

$$\begin{aligned}
&\big\|\big(\mathcal{Z}_{N,\varepsilon}^{(\text{no triple})}-\mathcal{Z}_N\big)(\varphi,\psi)\big\|_{L^2}^2\\
&\quad=\frac{1}{N^2}\Bigg\{\underbrace{\sum_{k>(\log\frac{1}{\varepsilon})^2}\ \sum_{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}}}_{\mathrm{I}_{N,\varepsilon}}+\underbrace{\sum_{k=1}^{(\log\frac{1}{\varepsilon})^2}\ \sum_{(\mathsf{i}_1,\ldots,\mathsf{i}_k)\in(\mathcal{A}_\varepsilon^{(\text{no triple})})^c}}_{\mathrm{II}_{N,\varepsilon}}\Bigg\}\\
&\quad\times\sum_{\substack{d_1\leqslant f_1\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\ldots,d_k\leqslant f_k\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_k)\\ x_1,y_1,\ldots,x_k,y_k\in\mathbb{Z}^2}}\Bigg(\sum_{v\in\mathbb{Z}^2_{\text{even}}}\varphi_N(v)\,q_{0,d_1}(v,x_1)\Bigg)^2\,\overline{U}_N(f_1-d_1,y_1-x_1)\\
&\quad\times\Bigg\{\prod_{j=2}^{k}q_{f_{j-1},d_j}(y_{j-1},x_j)^2\,\overline{U}_N(f_j-d_j,y_j-x_j)\Bigg\}\Bigg(\sum_{w\in\mathbb{Z}^2_{\text{even}}}q_{f_k,N}(y_k,w)\,\psi_N(w)\Bigg)^2,
\end{aligned}\tag{5.5}$$

where $\mathrm{I}_{N,\varepsilon}$ and $\mathrm{II}_{N,\varepsilon}$ are the contributions of $\{k>(\log\frac{1}{\varepsilon})^2\}$ and $\{(\mathsf{i}_1,\ldots,\mathsf{i}_k)\in(\mathcal{A}_\varepsilon^{(\text{no triple})})^c\}$. We split the proof in two parts, where we show that for some $\mathsf{C}<\infty$ we have:

$$\limsup_{N\to\infty\text{ with }N\in2\mathbb{N}}\ \mathrm{I}_{N,\varepsilon}\ \leqslant\ \mathsf{C}\,\frac{1}{\log\frac{1}{\varepsilon}}\,\|\varphi\|_{\mathcal{G}_1}^2\,\|\psi\|_\infty^2\,,\tag{5.6}$$

$$\limsup_{N\to\infty\text{ with }N\in2\mathbb{N}}\ \mathrm{II}_{N,\varepsilon}\ \leqslant\ \mathsf{C}\left(\|\varphi\|_{\mathcal{G}_{K_\varepsilon\varepsilon}}^2+\frac{(\log K_\varepsilon)^2}{\log\frac{1}{\varepsilon}}\|\varphi\|_{\mathcal{G}_1}^2\right)\|\psi\|_\infty^2\,.\tag{5.7}$$

Remark 5.2. *Let us sketch a probabilistic interpretation of* (5.5)*. From* (3.33)*, we recall that the expansion for* $\overline{U}_N(f-d,y-x)$ *has a time-space renewal interpretation, and from* (3.51) *the expansion of* $\mathbb{E}\big[\mathcal{Z}_N(\varphi,\psi)^2\big]$ *consists of a mixture of* $\overline{U}_N(f-d,y-x)$*, with weight* $\frac{1}{N^2}q_{0,d}^N(\varphi,x)^2q_{f,Nt}^N(y,\psi)^2$*. We can therefore write*

$$\mathbb{E}\big[\mathcal{Z}_N(\varphi,\psi)^2\big]=\sum_{\substack{\mathcal{S}=\{(n_1,z_1),\ldots,(n_r,z_r)\}\\ 1\leqslant n_1<\cdots n_r\leqslant N,\,z_1,\ldots,z_r\in\mathbb{Z}^2}}\mathsf{M}_N^{\varphi,\psi}(\mathcal{S}),$$

where denoting $\mathcal{S}=\{(n_1,z_1),\ldots,(n_r,z_r)\}$ *with* $r=|\mathcal{S}|$*, we define*

$$\begin{aligned}
\mathsf{M}_N^{\varphi,\psi}(\mathcal{S}):=\frac{\sigma_N^{2r}R_N^{2(r-1)}}{N^2}&\Bigg(\sum_{v\in\mathbb{Z}^2_{\text{even}}}\varphi_N(v)\,q_{0,n_1}(v,z_1)\Bigg)^2\Bigg(\sum_{w\in\mathbb{Z}^2_{\text{even}}}q_{n_r,N}(z_r,w)\,\psi_N(w)\Bigg)^2\\
&\times\mathrm{P}\Big(\big(\tau_i^{(N)},S_i^{(N)}\big)=(n_i,z_i)\,\forall\,1\leqslant i\leqslant r\,\Big|\,(\tau_1^{(N)},S_1^{(N)})=(n_1,z_1)\Big).
\end{aligned}\tag{5.8}$$

The measure $\mathsf{M}_N^{\varphi,\psi}(\cdot)$ *is called a* spectral measure *since* $\mathsf{M}_N^{\varphi,\psi}(\mathcal{S})$ *equals the square of the coefficient of* $\prod_{(n_i,z_i)\in\mathcal{S}}\xi_N(n_i,z_i)$ *in the chaos expansion* (3.19) *for* $\mathcal{Z}_N(\varphi,\psi)$*, where different terms in the expansion are orthogonal in* L^2*, similar to a Fourier decomposition. For more on spectral measure, see e.g.* [GaSt12].

The r.h.s. of (5.5) *can then be written as*

$$\big\|\big(Z_{N,\varepsilon}^{(\text{no triple})}-Z_N\big)(\varphi,\psi)\big\|_{L^2}^2\ =\ \mathsf{M}_N^{\varphi,\psi}\Big(|\mathcal{I}(\mathcal{S})|>(\log\tfrac{1}{\varepsilon})^2\ \text{ or }\ \mathcal{I}(\mathcal{S})\in(\mathcal{A}_\varepsilon^{(\text{no triple})})^c\Big)\,,\tag{5.9}$$

where given $\mathcal{S} \subseteq \{1, \dots, N\} \times \mathbb{Z}^2$,

$$\mathcal{I}(\mathcal{S}) := \big\{ \mathsf{i} \in \{1, \dots, \tfrac{1}{\varepsilon}\} : \ \mathcal{S} \cap \big(\mathcal{T}_{\varepsilon N}(\mathsf{i}) \times \mathbb{Z}^2 \big) \neq \varnothing \big\} .$$

Thanks to Lemma 3.3, it can be shown that as $N \to \infty$, $\mathsf{M}_N^{\varphi,\psi}$ *converges to a similarly re-weighted measure for the continuum time-space renewal process introduced in Lemma 3.3, whose time component is the Dickman subordinator* $(Y_s)_{s \geqslant 0}$ *with exponentially weighted Green's function* G_ϑ*, see* (3.37) *and* (3.39).

Second Moment Bound via Renewal. We first explain the common steps in bounding $\mathrm{I}_{N,\varepsilon}$ and $\mathrm{II}_{N,\varepsilon}$ from (5.5), which also applies to the variance of $\mathcal{Z}_N(\varphi, \psi)$ and the mean squared error of later approximations. The common feature is that they all have the same expansion as in (5.5), except the summation constraints are different.

Consider $\mathrm{II}_{N,\varepsilon}$ from (5.5). We first sum out the space variables in (5.5). Recall (3.9) and note that

$$\|\psi_N\|_{\ell^\infty} := \sup_{w \in \mathbb{Z}^2} |\psi_N(w)| \leqslant \|\psi\|_\infty , \tag{5.10}$$

so that in (5.5) we can bound

$$\Big(\sum_{w \in \mathbb{Z}^2_{\mathrm{even}}} q_{f_k,N}(y_k, w)\, \psi_N(w) \Big)^2 \leqslant \|\psi_N\|_{\ell^\infty}^2 \leqslant \|\psi\|_\infty^2 . \tag{5.11}$$

We can plug this bound into (5.5) and sum over the space variables in reverse order, from y_k, x_k until y_2, x_2, y_1, thus replacing $\overline{U}_N(f_j - d_j, y_j - x_j)$ by $\overline{U}_N(f_j - d_j)$ and $q_{f_{j-1},d_j}(y_{j-1}, x_j)^2$ by $u(d_j - f_{j-1})$, see (3.40) and (3.3). Finally, we sum over x_1 and observe that

$$\sum_{x_1 \in \mathbb{Z}^2} \Big(\sum_{v \in \mathbb{Z}^2_{\mathrm{even}}} \varphi_N(v)\, q_{0,d_1}(v, x_1) \Big)^2 = \sum_{v, v' \in \mathbb{Z}^2_{\mathrm{even}}} \varphi_N(v) \varphi_N(v')\, q_{2d_1}(v - v') =: N\, \Phi_N\Big(\frac{d_1}{N}\Big) , \tag{5.12}$$

where we introduced the function Φ_N. Substituting these bounds into (5.5) then gives

$$\begin{aligned} \mathrm{II}_{N,\varepsilon} \leqslant \frac{C\, \|\psi\|_\infty^2}{N} \sum_{k=1}^{(\log \frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1, \dots, \mathsf{i}_k) \in (\mathcal{A}_\varepsilon^{(\text{no triple})})^c} \ \sum_{d_1 \leqslant f_1 \in \mathcal{T}_{\varepsilon N}(\mathsf{i}_1), \dots, d_k \leqslant f_k \in \mathcal{T}_{\varepsilon N}(\mathsf{i}_k)} \\ \Phi_N\Big(\frac{d_1}{N}\Big)\, \overline{U}_N(f_1 - d_1) \Big\{ \prod_{j=2}^{k} u(d_j - f_{j-1})\, \overline{U}_N(f_j - d_j) \Big\} . \end{aligned} \tag{5.13}$$

A similar estimate can be derived for $\mathrm{I}_{N,\varepsilon}$, with a corresponding summation constraint.

We now compute the limit as $N \to \infty$ of the r.h.s. of (5.13). Recalling φ_N from (3.9), the local limit theorem (3.21), and (3.22), if $d_1/N \to s \in (0,1)$, then for Φ_N in (5.12) we have

$$\begin{aligned} \Phi_N\Big(\frac{d_1}{N}\Big) := \ & \frac{1}{N} \sum_{v, v' \in \mathbb{Z}^2_{\mathrm{even}}} \varphi_N(v)\, \varphi_N(v')\, q_{2d_1}(v - v') \\ & \xrightarrow[N \to \infty]{} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \varphi(z) \varphi(z') g_s(z' - z)\, \mathrm{d}z \mathrm{d}z' =: \Phi(s), \end{aligned} \tag{5.14}$$

where we note that, by the definition of $\|\varphi\|_{\mathcal{G}_t}$ from (3.52),

$$\int_0^t \Phi(s) \mathrm{d}s = \|\varphi\|_{\mathcal{G}_t}^2 . \tag{5.15}$$

We will use the following result, which says that as $N \to \infty$, for each $\mathsf{I} := (\mathsf{i}_1, \dots, \mathsf{i}_k) \subset \{1, \dots, \frac{1}{\varepsilon}\}$, the term in (5.13) converges to a limit that can be interpreted in terms of the Dickman subordinator, as mentioned in Remark 5.2.

Lemma 5.3. *Let Φ_N and Φ be defined as in (5.14). For any fixed $\varepsilon > 0$, $k \in \mathbb{N}$ and $\mathsf{I} := \{\mathsf{i}_1, \dots, \mathsf{i}_k\} \subset \{1, \dots, \frac{1}{\varepsilon}\}$ with $\mathsf{i}_1 < \mathsf{i}_2 < \ldots < \mathsf{i}_k$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{d_1 \leqslant f_1 \in \mathcal{T}_{\varepsilon N}(\mathsf{i}_1), \dots, d_k \leqslant f_k \in \mathcal{T}_{\varepsilon N}(\mathsf{i}_k)} \Phi_N\big(\tfrac{d_1}{N}\big)\, \overline{U}_N(f_1 - d_1) \left\{ \prod_{j=2}^{k} u(d_j - f_{j-1})\, \overline{U}_N(f_j - d_j) \right\} = \mathcal{I}_\varepsilon^\Phi(\mathsf{I})\,, \tag{5.16}$$

with

$$\mathcal{I}_\varepsilon^\Phi(\mathsf{I}) := \int\cdots\int_{a_1 \leqslant b_1 \in \mathcal{T}_\varepsilon(\mathsf{i}_1), \dots, a_k \leqslant b_k \in \mathcal{T}_\varepsilon(\mathsf{i}_k)} \mathrm{d}a_1\, \mathrm{d}b_1 \cdots \mathrm{d}a_k\, \mathrm{d}b_k \; \Phi(a_1)\, G_\vartheta(b_1 - a_1) \left\{ \prod_{j=2}^{k} \frac{1}{a_j - b_{j-1}}\, G_\vartheta(b_j - a_j) \right\}, \tag{5.17}$$

where $\mathcal{T}_\varepsilon(\mathsf{i}) := (\varepsilon(\mathsf{i}-1), \varepsilon\mathsf{i}]$, and G_ϑ (see (3.39) and (3.37)) is the weighted Green's function for the Dickman subordinator with Lévy measure $\frac{1}{t}\mathbb{1}_{(0,1)}(t)\mathrm{d}t$ introduced in Lemma 3.3.

Proof. If we introduce the macroscopic variables $a_i := d_i/N$, $b_i := f_i/N$ in (5.16), the sums converge to corresponding integrals as $N \to \infty$ (for fixed $\varepsilon > 0$), by the asymptotic expansions (3.3), (3.11), (3.42) and (3.43) for $u(\cdot)$, σ_N^2, $\overline{U}_N(\cdot)$ (also recall (3.2)), as well as the local limit theorem (3.21). This gives (5.16)-(5.17). □

We can interpret $\mathcal{I}_\varepsilon^\Phi(\mathsf{I})$ in (5.17) as the weight associated to a Dickman subordinator. More precisely, recall that G_ϑ is the weighted Green's function of the Dickman subordinator Y introduced in Lemma 3.3, and satisfies the following *renewal property* [CSZ19a, eq. (6.14)]:

$$\forall s < \bar{t} < t: \qquad G_\vartheta(t-s) = \iint_{u\in(s,\bar{t}),\, v\in(\bar{t},t)} G_\vartheta(u-s)\, \frac{1}{v-u}\, G_\vartheta(t-v)\, \mathrm{d}u\, \mathrm{d}v\,. \tag{5.18}$$

In (5.17), let us denote by $\mathcal{I}_\varepsilon^{s,t}(\mathsf{i}_1, \dots, \mathsf{i}_k)$ the integral where the extreme variables a_1 and b_k are not integrated out but rather fixed to be s and t respectively, namely,

$$\mathcal{I}_\varepsilon^{s,t}(\mathsf{I}) := \int\cdots\int_{s \leqslant b_1 \in \mathcal{T}_\varepsilon(\mathsf{i}_1), \dots, a_k \leqslant t \in \mathcal{T}_\varepsilon(\mathsf{i}_k)} \mathrm{d}b_1\, \mathrm{d}a_2\, \mathrm{d}b_2 \cdots \mathrm{d}a_{k-1}\, \mathrm{d}b_{k-1}\, \mathrm{d}a_k \; G_\vartheta(b_1 - s) \left\{ \prod_{j=2}^{k-1} \frac{1}{a_j - b_{j-1}}\, G_\vartheta(b_j - a_j) \right\} \frac{1}{a_k - b_{k-1}}\, G_\vartheta(t - a_k)\,. \tag{5.19}$$

This is the weight of renewal configurations that only visit the intervals $\mathcal{T}_\varepsilon(\mathsf{i}_1)$, …, $\mathcal{T}_\varepsilon(\mathsf{i}_k)$, and a_j, b_j are the first and last renewal points in $\mathcal{T}_\varepsilon(\mathsf{i}_j)$, while $\frac{1}{a_j - b_{j-1}}$ comes from the Lévy

measure of the Dickman subordinator. An iterative application of (5.18) then shows that

$$\forall \mathsf{j} \leqslant \mathsf{j}',\ \forall s \in \mathcal{T}_\varepsilon(\mathsf{j}),\ \forall t \in \mathcal{T}_\varepsilon(\mathsf{j}') : \\ \sum_{k=1}^{\infty} \sum_{\mathsf{j}=:\mathsf{i}_1<\mathsf{i}_2<\ldots<\mathsf{i}_{k-1}<\mathsf{i}_k:=\mathsf{j}'} \mathcal{I}_\varepsilon^{s,t}(\mathsf{i}_1,\ldots,\mathsf{i}_k) \;=\; G_\vartheta(t-s)\,, \tag{5.20}$$

which is just a renewal decomposition by summing over the set of possible intervals $\mathcal{T}_\varepsilon(\mathsf{i})$, $\mathsf{j} \leqslant \mathsf{i} \leqslant \mathsf{j}'$, visited by the Dickman subordinator Y, given that s, t are in the range of Y.

Applying Lemma 5.3 to (5.13) then gives

$$\limsup_{N\to\infty} \mathrm{II}_{N,\varepsilon} \;\leqslant\; C\,\|\psi\|_\infty^2 \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\ldots,\mathsf{i}_k)\in(\mathcal{A}_\varepsilon^{(\text{no triple})})^c} \mathcal{I}_\varepsilon^{\Phi}(\mathsf{i}_1,\ldots,\mathsf{i}_k)\,, \tag{5.21}$$

With the same arguments, we obtain a corresponding bound for $\mathrm{I}_{N,\varepsilon}$:

$$\limsup_{N\to\infty} \mathrm{I}_{N,\varepsilon} \;\leqslant\; C\,\|\psi\|_\infty^2 \sum_{k>(\log\frac{1}{\varepsilon})^2} \sum_{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}} \mathcal{I}_\varepsilon^{\Phi}(\mathsf{i}_1,\ldots,\mathsf{i}_k)\,. \tag{5.22}$$

To complete the proof of Lemma 5.1, it remains to derive (5.6) and (5.7) from these bounds. We start with $\mathrm{II}_{N,\varepsilon}$, which is more involved, but we first make a remark.

Remark 5.4 (Variance bound). *If we remove any constraint on k and $(\mathsf{i}_1,\ldots,\mathsf{i}_k)$ from formula* (5.5)*, summing over all $k \in \mathbb{N}$ and $0 < \mathsf{i}_1 < \cdots < \mathsf{i}_k \leqslant \frac{1}{\varepsilon}$, we obtain $\mathbb{V}\mathrm{ar}(\mathcal{Z}_N(\varphi,\psi))$ (recall* (3.19)*). We thus have a simpler analogue of* (5.21) *and* (5.22)*:*

$$\begin{aligned}\limsup_{N\to\infty} \mathbb{V}\mathrm{ar}(\mathcal{Z}_N(\varphi,\psi)) \;&\leqslant\; C\,\|\psi\|_\infty^2 \sum_{k=1}^{\infty} \sum_{0<\mathsf{i}_1<\cdots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}} \mathcal{I}_\varepsilon^{\Phi}(\mathsf{i}_1,\ldots,\mathsf{i}_k) \\ &\leqslant\; C\,\|\psi\|_\infty^2 \iint\limits_{0<s<t<1} \Phi(s) G_\vartheta(t-s)\,\mathrm{d}t\,\mathrm{d}s \\ &\leqslant\; C\,\|\psi\|_\infty^2 \bigg(\int_0^1 G_\vartheta(t)\mathrm{d}t\bigg)\bigg(\int_0^1 \Phi(s)\,\mathrm{d}s\bigg) = C c_\vartheta \|\psi\|_\infty^2 \|\varphi\|_{\mathcal{G}_1}^2,\end{aligned} \tag{5.23}$$

where in the second inequality we applied the renewal decomposition (5.20)*, with s and t being the first and last renewal points, we denoted $c_\vartheta := \int_0^1 G_\vartheta(t)\mathrm{d}t$, see* (3.46)*, and recalled* (5.15)*. Note that this bound is the same as the one in* (3.55) *and does not depend on ε.*

Bound for $\mathrm{II}_{N,\varepsilon}$: proof of (5.7)**.** We start with (5.21). The constraint $(\mathcal{A}_\varepsilon^{(\text{no triple})})^c$ contains $\mathsf{I} = (\mathsf{i}_1,\ldots,\mathsf{i}_k)$ with either $1 \leqslant \mathsf{i}_1 < K_\varepsilon$, or $\mathsf{i}_k > \frac{1}{\varepsilon} - K_\varepsilon$, or $\mathsf{i}_{j+1} - \mathsf{i}_j, \mathsf{i}_{j+2} - \mathsf{i}_{j+1} < K_\varepsilon$ for some j (hence $k \geqslant 3$). We will treat the three cases one by one.

For the first case $1 \leqslant \mathsf{i}_1 < K_\varepsilon$, omitting the factor $\|\psi\|_\infty^2$, its contribution is bounded by

$$\sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{\substack{0<\mathsf{i}_1<K_\varepsilon \\ \mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}}} \mathcal{I}_\varepsilon^{\Phi}(\mathsf{i}_1,\ldots,\mathsf{i}_k) \leqslant \iint\limits_{\substack{0<s<K_\varepsilon\varepsilon \\ s<t<1}} \Phi(s)\,G_\vartheta(t-s)\,\mathrm{d}s\,\mathrm{d}t \leqslant c_\vartheta \|\varphi\|_{\mathcal{G}_{K_\varepsilon\varepsilon}}^2, \tag{5.24}$$

where we applied the renewal decomposition (5.20) as in (5.23), and recalled from (5.15) that $\|\varphi\|_{\mathcal{G}_{K_\varepsilon\varepsilon}}^2 = \int_0^{K_\varepsilon\varepsilon} \Phi(s)\mathrm{d}s$. This gives the first term in (5.7).

For the second case $\mathsf{i}_k > \frac{1}{\varepsilon} - K_\varepsilon$, omitting $\|\psi\|_\infty^2$, its contribution is bounded by

$$\sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{\substack{0<\mathsf{i}_1<\ldots<\mathsf{i}_{k-1}<\mathsf{i}_k \\ \frac{1}{\varepsilon}-K_\varepsilon<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}}} \mathcal{I}_\varepsilon^\Phi(\mathsf{i}_1,\ldots,\mathsf{i}_k) \leqslant \iint_{\substack{0<s<t \\ 1-K_\varepsilon\varepsilon<t<1}} \Phi(s)\, G_\vartheta(t-s)\,\mathrm{d}s\,\mathrm{d}t$$
$$\leqslant C\Big(\int_0^{K_\varepsilon\varepsilon} G(t)\mathrm{d}t\Big)\Big(\int_0^1 \Phi(s)\mathrm{d}s\Big) \leqslant \frac{C\|\varphi\|_{\mathcal{G}_1}^2}{\log\frac{1}{\varepsilon}}\,,$$

where in the second inequality, we used that $\int_a^{a+K_\varepsilon\varepsilon} G_\vartheta(t)\,\mathrm{d}t \leqslant C\int_0^{K_\varepsilon\varepsilon} G_\vartheta(t)\mathrm{d}t \leqslant C/\log\frac{1}{\varepsilon}$ uniformly in ε small enough and $a\in(0,1)$ (by (3.46)-(3.47) and the choice (4.2) of K_ε). This bound is much smaller than the second term in (5.7).

For the third case with $\mathsf{i}_{j+1}-\mathsf{i}_j, \mathsf{i}_{j+2}-\mathsf{i}_{j+1} < K_\varepsilon$ for some j, we need to bound

$$\mathcal{W}_\varepsilon := \sum_{k=3}^{(\log\frac{1}{\varepsilon})^2}\sum_{j=1}^{k-2} \sum_{\substack{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon} \\ \mathsf{i}_{j+1}-\mathsf{i}_j<K_\varepsilon \text{ and } \mathsf{i}_{j+2}-\mathsf{i}_{j+1}<K_\varepsilon}} \mathcal{I}_\varepsilon^\Phi(\mathsf{i}_1,\ldots,\mathsf{i}_k)$$
$$\leqslant \sum_{\substack{0<\mathsf{i}<\mathsf{i}'<\mathsf{i}''\leqslant\frac{1}{\varepsilon} \\ |\mathsf{i}'-\mathsf{i}|<K_\varepsilon \\ |\mathsf{i}''-\mathsf{i}'|<K_\varepsilon}} \iiiint_{\substack{b\in\mathcal{T}_\varepsilon(\mathsf{i}) \\ a'\leqslant b'\in\mathcal{T}_\varepsilon(\mathsf{i}') \\ a''\in\mathcal{T}_\varepsilon(\mathsf{i}'')}} \mathrm{d}b\,\mathrm{d}a'\,\mathrm{d}b'\,\mathrm{d}a''\int_0^b \mathrm{d}s\int_{a''}^1 \mathrm{d}t$$
$$\Phi(s)\, G_\vartheta(b-s)\,\frac{1}{a'-b}\,G_\vartheta(b'-a')\,\frac{1}{a''-b'}\,G_\vartheta(t-a'')\,.$$

where we again applied the renewal decomposition (5.20). Bounding the integral of $G_\vartheta(t-a'')$ over t by $c_\vartheta = \int_0^1 G_\vartheta(u)\,\mathrm{d}u$, we obtain

$$\mathcal{W}_\varepsilon \leqslant \sum_{\mathsf{i}=1}^{\frac{1}{\varepsilon}}\sum_{\mathsf{i}'=\mathsf{i}+1}^{\mathsf{i}+K_\varepsilon}\sum_{\mathsf{i}''=\mathsf{i}'+1}^{\mathsf{i}'+K_\varepsilon} \iiiint_{\substack{b\in\mathcal{T}_\varepsilon(\mathsf{i}) \\ a'\leqslant b'\in\mathcal{T}_\varepsilon(\mathsf{i}') \\ a''\in\mathcal{T}_\varepsilon(\mathsf{i}'')}} \mathrm{d}b\,\mathrm{d}a'\,\mathrm{d}b'\,\mathrm{d}a''\int_0^b c_\vartheta\Phi(s)\frac{G_\vartheta(b-s)}{a'-b}\frac{G_\vartheta(b'-a')}{a''-b'}\,\mathrm{d}s\,. \tag{5.25}$$

Note that if we restrict the sum to $2\leqslant \mathsf{i}'-\mathsf{i}, \mathsf{i}''-\mathsf{i}' \leqslant K_\varepsilon$, then using (3.47), it is not difficult to see that the integrals can be bounded by $C\frac{(\log K_\varepsilon)^2}{\log\frac{1}{\varepsilon}}\int_0^1\Phi(s)\mathrm{d}s$. Complications only arise when $\mathsf{i}'=\mathsf{i}+1$ or $\mathsf{i}''=\mathsf{i}'+1$.

We will proceed in three steps. The following bound will be used repeatedly:

$$\forall\delta\in(0,\tfrac12)\,,\ \forall z\in[\delta,\infty): \qquad \int_0^\delta G_\vartheta(x)\,\log\Big(1+\frac{z}{\delta-x}\Big)\,\mathrm{d}x \leqslant C\,\frac{\log(1+\frac{z}{\delta})}{\log\frac{1}{\delta}}\,. \tag{5.26}$$

Indeed, splitting the integral over $(0,\frac{\delta}{2})$ and $(\frac{\delta}{2},\delta)$ and exploiting (3.46), we note that:

- for $x<\frac{\delta}{2}$ we have $\log(1+\frac{z}{\delta-x}) \leqslant \log(1+\frac{z}{\delta/2}) \leqslant C\,\log(1+\frac{z}{\delta})$ and $\int_0^{\delta/2} G_\vartheta(x)\,\mathrm{d}x \leqslant \frac{C'}{\log\frac{1}{\delta}}$;
- for $x>\frac{\delta}{2}$ we can bound $G_\vartheta(x)\leqslant\frac{C}{\delta(\log\frac{1}{\delta})^2}$ and, by the change of variable $t:=\frac{z}{\delta-x}$,

$$\int_0^\delta \log\Big(1+\frac{z}{\delta-x}\Big)\,\mathrm{d}x = z\int_{\frac{z}{\delta}}^\infty \frac{\log(1+t)}{t^2}\,\mathrm{d}t \leqslant C'\,\delta\,\log\Big(1+\frac{z}{\delta}\Big)\,. \tag{5.27}$$

We now continue to bound the r.h.s. of (5.25).

Step 1. Given $\mathsf{i}' \in \mathbb{N}$ and $a' \in \mathcal{T}_\varepsilon(\mathsf{i}')$, we bound the integrals over a'' and b' in (5.25):

$$\sum_{\mathsf{i}''=\mathsf{i}'+1}^{\mathsf{i}'+K_\varepsilon} \iint\limits_{\substack{b' \in \mathcal{T}_\varepsilon(\mathsf{i}'):\, b'>a' \\ a'' \in \mathcal{T}_\varepsilon(\mathsf{i}'')}} \mathrm{d}a''\, \mathrm{d}b'\, G_\vartheta(b'-a')\, \frac{1}{a''-b'} \;=\; \int_{a'}^{\varepsilon \mathsf{i}'} \mathrm{d}b'\, G_\vartheta(b'-a')\, \log \frac{\varepsilon(\mathsf{i}'+K_\varepsilon)-b'}{\varepsilon \mathsf{i}'-b'}$$

$$= \Big[\int_0^\delta \mathrm{d}x\, G_\vartheta(x)\, \log\Big(1+\frac{\varepsilon K_\varepsilon}{\delta - x}\Big) \Big]\Big|_{\delta := \varepsilon \mathsf{i}'-a'} \leqslant C\, \frac{\log(1+\frac{\varepsilon K_\varepsilon}{\varepsilon \mathsf{i}'-a'})}{\log \frac{1}{\varepsilon \mathsf{i}'-a'}}\,,$$

where we used (5.26) and changed variable $x := b' - a'$. Plugging it into (5.25), we obtain

$$\mathcal{W}_\varepsilon \leqslant C \sum_{\mathsf{i}=1}^{\frac{1}{\varepsilon}} \sum_{\mathsf{i}'=\mathsf{i}+1}^{\mathsf{i}+K_\varepsilon} \iint\limits_{\substack{b \in \mathcal{T}_\varepsilon(\mathsf{i}) \\ a' \in \mathcal{T}_\varepsilon(\mathsf{i}')}} \mathrm{d}b\, \mathrm{d}a' \int_0^b \mathrm{d}s\; \Phi(s)\, G_\vartheta(b-s)\, \frac{1}{a'-b}\, \frac{\log(1+\frac{\varepsilon K_\varepsilon}{\varepsilon \mathsf{i}'-a'})}{\log \frac{1}{\varepsilon \mathsf{i}'-a'}}\,. \tag{5.28}$$

Step 2. Given $\mathsf{i} \in \mathbb{N}$ and $b \in \mathcal{T}_\varepsilon(\mathsf{i})$, we focus on the integral over a' in (5.28):

$$\sum_{\mathsf{i}'=\mathsf{i}+1}^{\mathsf{i}+K_\varepsilon} \int_{\varepsilon(\mathsf{i}'-1)}^{\varepsilon \mathsf{i}'} \mathrm{d}a'\, \frac{1}{a'-b}\, \frac{\log(1+\frac{\varepsilon K_\varepsilon}{\varepsilon \mathsf{i}'-a'})}{\log \frac{1}{\varepsilon \mathsf{i}'-a'}} = \sum_{\mathsf{i}'=\mathsf{i}+1}^{\mathsf{i}+K_\varepsilon} \int_0^\varepsilon \frac{1}{\varepsilon \mathsf{i}'-b-x}\, \frac{\log(1+\frac{\varepsilon K_\varepsilon}{x})}{\log \frac{1}{x}}\, \mathrm{d}x\,, \tag{5.29}$$

by the change of variables $x := \varepsilon \mathsf{i}' - a'$. We first bound the sum from $\mathsf{i}' = \mathsf{i}+2$ onward, for which we note that for $x \in (0, \varepsilon)$,

$$\sum_{\mathsf{i}'=\mathsf{i}+2}^{\mathsf{i}+K_\varepsilon} \frac{1}{\varepsilon \mathsf{i}'-b-x} \leqslant \frac{1}{\varepsilon} \log \frac{\varepsilon(\mathsf{i}+K_\varepsilon+1)-b-x}{\varepsilon(\mathsf{i}+1)-b-x} = \frac{1}{\varepsilon} \log\Big(1+\frac{\varepsilon K_\varepsilon}{\varepsilon(\mathsf{i}+1)-b-x}\Big)$$

$$\leqslant \frac{1}{\varepsilon} \log\Big(1+\frac{\varepsilon K_\varepsilon}{\varepsilon \mathsf{i}-b}\Big)\,.$$

Moreover, by the change of variables $x = \varepsilon\, e^{-t}$,

$$\int_0^\varepsilon \frac{\log(1+\frac{\varepsilon K_\varepsilon}{x})}{\log \frac{1}{x}}\, \mathrm{d}x \;=\; \varepsilon \int_0^\infty e^{-t}\, \frac{\log(1+K_\varepsilon e^t)}{t+\log \frac{1}{\varepsilon}}\, \mathrm{d}t \;\leqslant\; \frac{\varepsilon}{\log \frac{1}{\varepsilon}} \int_0^\infty e^{-t}(t+\log(1+K_\varepsilon))\, \mathrm{d}t\,, \tag{5.30}$$

because $t + \log\frac{1}{\varepsilon} \geqslant \log\frac{1}{\varepsilon}$ and $1 + K_\varepsilon e^t \leqslant (1+K_\varepsilon)e^t$. Therefore

$$\sum_{\mathsf{i}'=\mathsf{i}+2}^{\mathsf{i}+K_\varepsilon} \int_0^\varepsilon \frac{1}{\varepsilon \mathsf{i}'-b-x}\, \frac{\log(1+\frac{\varepsilon K_\varepsilon}{x})}{\log \frac{1}{x}}\, \mathrm{d}x \leqslant C\, \frac{\log(1+K_\varepsilon)}{\log \frac{1}{\varepsilon}} \log\Big(1+\frac{\varepsilon K_\varepsilon}{\varepsilon \mathsf{i}-b}\Big).$$

Now for the case $\mathsf{i}' = \mathsf{i}+1$, we have

$$\begin{aligned}
&\int_0^{\varepsilon} \frac{1}{\varepsilon(\mathsf{i}+1)-b-x}\,\frac{\log(1+\frac{\varepsilon K_\varepsilon}{x})}{\log\frac{1}{x}}\,\mathrm{d}x \\
&\leqslant \frac{2}{\varepsilon}\int_0^{\varepsilon/2}\frac{\log(1+\frac{\varepsilon K_\varepsilon}{x})}{\log\frac{1}{x}}\mathrm{d}x + C\,\frac{\log(1+K_\varepsilon)}{\log\frac{1}{\varepsilon}}\int_{\varepsilon/2}^{\varepsilon}\frac{1}{\varepsilon(\mathsf{i}+1)-b-x}\,\mathrm{d}x \\
&\leqslant C\,\frac{\log(1+K_\varepsilon)}{\log\frac{1}{\varepsilon}}\left(1+\log\left(1+\frac{\varepsilon/2}{\varepsilon\mathsf{i}-b}\right)\right) \leqslant C\,\frac{\log(1+K_\varepsilon)}{\log\frac{1}{\varepsilon}}\log\left(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-b}\right).
\end{aligned}$$

Substituting the above bounds into (5.28) then gives

$$\begin{aligned}
\mathcal{W}_\varepsilon &\leqslant C'\,\frac{\log(1+K_\varepsilon)}{\log\frac{1}{\varepsilon}}\sum_{\mathsf{i}=1}^{\frac{1}{\varepsilon}}\int_{b\in\mathcal{T}_\varepsilon(\mathsf{i})}\mathrm{d}b\int_0^b \mathrm{d}s\;\Phi(s)\,G_\vartheta(b-s)\,\log\left(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-b}\right) \\
&= C'\,\frac{\log(1+K_\varepsilon)}{\log\frac{1}{\varepsilon}}\sum_{\mathsf{i}=1}^{\frac{1}{\varepsilon}}\int_0^{\varepsilon\mathsf{i}}\mathrm{d}s\,\Phi(s)\int_{\max\{s,\varepsilon(\mathsf{i}-1)\}}^{\varepsilon\mathsf{i}}\mathrm{d}b\,G_\vartheta(b-s)\,\log\left(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-b}\right).
\end{aligned} \tag{5.31}$$

Step 3. Given $s\in(0,1)$, we bound the integral over b in (5.31).

- For $s\in(\varepsilon(\mathsf{i}-2),\varepsilon\mathsf{i})$ we can bound, by (5.26) with $\delta=\varepsilon\mathsf{i}-s$,

$$\int_s^{\varepsilon\mathsf{i}} G_\vartheta(b-s)\,\log\left(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-b}\right)\mathrm{d}b \leqslant C\,\frac{\log(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-s})}{\log\frac{1}{\varepsilon\mathsf{i}-s}} \leqslant C'\,\log(1+K_\varepsilon)\,, \tag{5.32}$$

  where the last inequality (which is very rough) holds, say for $\varepsilon\in(0,\frac14)$, *uniformly for* $s\in(\varepsilon(\mathsf{i}-2),\varepsilon\mathsf{i})$ *and* $K_\varepsilon\geqslant 1$, because $x:=\frac{1}{\varepsilon\mathsf{i}-s}\geqslant\frac{1}{2\varepsilon}\geqslant 2$ and

$$\sup_{x\geqslant 2}\frac{\log(1+(\varepsilon K_\varepsilon)x)}{\log x}\leqslant\sup_{x\geqslant2}\frac{\log((1+K_\varepsilon)x)}{\log x}=\frac{\log((1+K_\varepsilon)2)}{\log 2}\leqslant\frac{2}{\log 2}\,\log(1+K_\varepsilon)\,.$$

- For $s\leqslant\varepsilon(\mathsf{i}-2)$ we can bound $G_\vartheta(b-s)\leqslant C\,G_\vartheta(\varepsilon(\mathsf{i}-1)-s)$, see (3.46), and

$$\int_{\varepsilon(\mathsf{i}-1)}^{\varepsilon\mathsf{i}}\log\left(1+\frac{\varepsilon K_\varepsilon}{\varepsilon\mathsf{i}-b}\right)\mathrm{d}b = \int_0^{\varepsilon}\log\left(1+\frac{\varepsilon K_\varepsilon}{x}\right)\mathrm{d}x \leqslant C'\,\varepsilon\,\log(1+K_\varepsilon)\,,$$

  by the change of variables $x:=\varepsilon\mathsf{i}-b$ and the estimate (5.27) with $\delta=\varepsilon$ and $z=\varepsilon K_\varepsilon$.

Substituting these bounds into (5.31) then gives

$$\begin{aligned}
\mathcal{W}_\varepsilon &\leqslant C''\,\frac{(\log(1+K_\varepsilon))^2}{\log\frac{1}{\varepsilon}}\sum_{\mathsf{i}=1}^{\frac{1}{\varepsilon}}\left\{\int_{\varepsilon(\mathsf{i}-2)^+}^{\varepsilon\mathsf{i}}\Phi(s)\,\mathrm{d}s + \varepsilon\int_0^{\varepsilon(\mathsf{i}-2)^+}\Phi(s)\,G_\vartheta(\varepsilon(\mathsf{i}-1)-s)\,\mathrm{d}s\right\} \\
&\leqslant C''\,\frac{(\log(1+K_\varepsilon))^2}{\log\frac{1}{\varepsilon}}\int_0^1\Phi(s)\left\{2+\sum_{\mathsf{i}=2+\lceil\frac{s}{\varepsilon}\rceil}^{\frac{1}{\varepsilon}}\varepsilon\,G_\vartheta(\varepsilon(\mathsf{i}-1)-s)\right\}\mathrm{d}s \\
&\leqslant C\,\frac{(\log(1+K_\varepsilon))^2}{\log\frac{1}{\varepsilon}}\int_0^1\Phi(s)\,\mathrm{d}s = C\,\frac{(\log(1+K_\varepsilon))^2}{\log\frac{1}{\varepsilon}}\|\varphi\|^2_{\mathcal{G}_1},
\end{aligned} \tag{5.33}$$

where the last sum can be seen as a Riemann sum and bounded by a multiple of $c_\vartheta = \int_0^1 G_\vartheta(x)\,\mathrm{d}x$. This bound gives the second term in (5.7) and completes its proof.

Bound for $\mathrm{I}_{N,\varepsilon}$: proof of (5.6). In view of (5.22) and (5.15), we need to show that

$$\mathcal{V}_\varepsilon := \sum_{\mathsf{I}\subset\{1,\dots,\frac{1}{\varepsilon}\},\,|\mathsf{I}|>(\log\frac{1}{\varepsilon})^2} \mathcal{I}_\varepsilon^\Phi(\mathsf{I}) \leqslant C\left(\int_0^1 \Phi(t)\,\mathrm{d}t\right)\frac{1}{\log\frac{1}{\varepsilon}}\,. \tag{5.34}$$

By Markov's inequality, we can bound

$$\mathcal{V}_\varepsilon \leqslant \frac{1}{(\log\frac{1}{\varepsilon})^2}\sum_{\mathsf{I}\subset\{1,\dots,\frac{1}{\varepsilon}\},\,|\mathsf{I}|>(\log\frac{1}{\varepsilon})^2} |\mathsf{I}|\,\mathcal{I}_\varepsilon^\Phi(\mathsf{I}) \leqslant \frac{1}{(\log\frac{1}{\varepsilon})^2}\sum_{\mathsf{j}=1}^{\frac{1}{\varepsilon}}\sum_{\mathsf{I}\subset\{1,\dots,\frac{1}{\varepsilon}\},\mathsf{I}\ni\mathsf{j}} \mathcal{I}_\varepsilon^\Phi(\mathsf{I})\,.$$

Recalling the renewal interpretation of $\mathcal{I}_\varepsilon^\Phi(\mathsf{I})$ after Lemma 5.3 and the renewal decomposition (5.20), we can integrate over the first renewal visit s, the last visited point $u\in\mathcal{T}_\varepsilon(\mathsf{j})$, the first visited point v after $\mathcal{T}_\varepsilon(\mathsf{j})$, and the last renewal visit $t\leqslant 1$, to obtain the bound

$$\begin{aligned}\mathcal{V}_\varepsilon &\leqslant \frac{1}{(\log\frac{1}{\varepsilon})^2}\sum_{1\leqslant\mathsf{j}\leqslant\frac{1}{\varepsilon}}\iiiint_{\substack{0<s<u<v<t\leqslant 1\\ u\in\mathcal{T}_\varepsilon(\mathsf{j}),v>\varepsilon\mathsf{j}}} \Phi(s)\,G_\vartheta(u-s)\,\frac{1}{v-u}\,G_\vartheta(t-v)\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\\ &\leqslant \frac{c_\vartheta}{(\log\frac{1}{\varepsilon})^2}\sum_{1\leqslant\mathsf{j}\leqslant\frac{1}{\varepsilon}}\iint_{0<s<u,\,u\in\mathcal{T}_\varepsilon(\mathsf{j})} \Phi(s)\,G_\vartheta(u-s)\,\log\frac{1}{\varepsilon\mathsf{j}-u}\,\mathrm{d}s\,\mathrm{d}t\,.\end{aligned}$$

Observe that the sum of the integrals is exactly the same as in the r.h.s. of (5.31) with K_ε replaced by $\frac{1}{\varepsilon}$. Therefore the bounds leading to (5.33) also applies, which gives

$$\mathcal{V}_\varepsilon \leqslant \frac{c_\vartheta}{(\log\frac{1}{\varepsilon})^2}C''\log\Big(1+\frac{1}{\varepsilon}\Big)\int_0^1\Phi(s)\mathrm{d}s \leqslant \frac{C}{\log\frac{1}{\varepsilon}}\int_0^1\Phi(s)\,\mathrm{d}s\,.$$

This matches our goal (5.34) and completes the proof of (5.6). □

5.2. Step 2: Diffusive truncation in space. In this step, we introduce our second approximation $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{diff})}(\varphi,\psi)$ and show that it is close to $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{no\,triple})}(\varphi,\psi)$.

For $\mathsf{a}=(\mathsf{a}^{(1)},\mathsf{a}^{(2)})\in\mathbb{Z}^2$, recall from (4.1) the mesoscopic spatial square

$$\begin{aligned}\mathcal{S}_{\varepsilon N}(\mathsf{a}) &:= \big((\mathsf{a}-(1,1))\sqrt{\varepsilon N},\,\mathsf{a}\sqrt{\varepsilon N}\,\big]\\ &= \big((\mathsf{a}^{(1)}-1)\sqrt{\varepsilon N},\,\mathsf{a}^{(1)}\sqrt{\varepsilon N}\,\big]\times\big((\mathsf{a}^{(2)}-1)\sqrt{\varepsilon N},\,\mathsf{a}^{(2)}\sqrt{\varepsilon N}\,\big]\,,\end{aligned} \tag{5.35}$$

to which we associate the mesoscopic space variable a. We now perform coarse-graining in space by grouping terms in the expansion of $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{no\,triple})}(\varphi,\psi)$ in (5.3) according to the mesoscopic spatial boxes visited by the space variables v,x_i,y_i,w in (5.3). Namely, recall the definition (4.1) of the mesoscopic time-space box

$$\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a}) = \big(\mathcal{T}_{\varepsilon N}(\mathsf{i})\times\mathcal{S}_{\varepsilon N}(\mathsf{a})\big)\ \cap\ \mathbb{Z}^3_{\mathrm{even}}\,. \tag{5.36}$$

We can rewrite (5.3) by introducing the *mesocopic space variables* $\mathsf{b}_0, \mathsf{a}_1, \mathsf{b}_1, \dots, \mathsf{a}_k, \mathsf{b}_k, \mathsf{a}_{k+1}$:

$$\begin{aligned}
\mathcal{Z}^{(\text{no triple})}_{N,\varepsilon}(\varphi,\psi) \;&=\; q^N_{0,N}(\varphi,\psi) \\
&+ \frac{1}{N} \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}^{(\text{no triple})}_\varepsilon} \sum_{\mathsf{b}_0\in\mathbb{Z}^2} \sum_{\mathsf{a}_1,\mathsf{b}_1,\dots,\mathsf{a}_k,\mathsf{b}_k\in\mathbb{Z}^2} \sum_{\substack{(d_1,x_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{a}_1)\\(f_1,y_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{b}_1)\\ \text{with } d_1\leqslant f_1}} \cdots \sum_{\substack{(d_k,x_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{a}_k)\\(f_k,y_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{b}_k)\\ \text{with } d_k\leqslant f_k}} \\
&\Bigg(\sum_{v\in\mathcal{S}_{\varepsilon N}(\mathsf{b})\cap\mathbb{Z}^2_{\text{even}}} \varphi_N(v)\, q_{0,d_1}(v,x_1)\Bigg) X_{d_1,f_1}(x_1,y_1) \Bigg\{\prod_{j=2}^{k} q_{f_{j-1},d_j}(y_{j-1},x_j)\, X_{d_j,f_j}(x_j,y_j)\Bigg\} \\
&\times \Bigg(\sum_{\mathsf{a}_{k+1}\in\mathbb{Z}^2} \sum_{w\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_{k+1})\cap\mathbb{Z}^2_{\text{even}}} q_{f_k,N}(y_k,w)\,\psi_N(w)\Bigg). \qquad (5.37)
\end{aligned}$$

We now perform a diffusive scale truncation by replacing each $X_{d_i,f_i}(x_i,y_i)$ in the above expansion by its truncated version $X^{(\text{diff})}_{d_i,f_i}(x_i,y_i)$ defined in (4.10). Let us stress that

$$X^{(\text{diff})}_{d,f}(x,y) = 0 \qquad \text{for } x\in\mathcal{S}_{\varepsilon N}(\mathsf{a}),\ y\in\mathcal{S}_{\varepsilon N}(\mathsf{b}) \text{ with } |b-a| > M_\varepsilon\,. \qquad (5.38)$$

Furthermore, we restrict the mesoscopic space variables $(\mathsf{a}_1,\, \mathsf{b}_1,\, \dots\, ,\mathsf{a}_k,\, \mathsf{b}_k)$ in (5.37) to a "diffusive set" that depends on the initial and final space variables b_0 and a_{k+1}, and time variables $(\mathsf{i}_0 := 0, \mathsf{i}_1, \dots, \mathsf{i}_k, \mathsf{i}_{k+1} := \frac{1}{\varepsilon})$:

$$\begin{aligned}
\mathcal{A}^{(\text{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}} := \Big\{&(\mathsf{a}_1,\, \mathsf{b}_1,\, \dots\, ,\mathsf{a}_k,\, \mathsf{b}_k)\in(\mathbb{Z}^2)^{2k} \ \text{ s.t. } \forall\, 1\leqslant j\leqslant k,\ |\mathsf{b}_j-\mathsf{a}_j|\leqslant M_\varepsilon, \\
&\text{and} \quad \forall\, 1\leqslant j\leqslant k+1,\ |\mathsf{a}_j-\mathsf{b}_{j-1}|\leqslant M_\varepsilon\sqrt{\mathsf{i}_j-\mathsf{i}_{j-1}}\ \Big\}. \qquad (5.39)
\end{aligned}$$

Remark 5.5. *Once* $(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}^{(\text{no triple})}_\varepsilon$ *are grouped into time blocks, see Definition 4.1, we can then group* $(\mathsf{a}_1,\, \mathsf{b}_1,\, \dots\, ,\mathsf{a}_k,\, \mathsf{b}_k)\in\mathcal{A}^{(\text{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ *into space-blocks, see Definition 4.2. The constraint* $\mathcal{A}^{(\text{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ *then maps to the constraint* $\vec{\mathcal{A}}^{(\text{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ *defined in* (4.6).

More explicitly, we can approximate $\mathcal{Z}^{(\text{no triple})}_{N,\varepsilon}(\varphi,\psi)$ from (5.37) by

$$\begin{aligned}
\mathcal{Z}^{(\text{diff})}_{N,\varepsilon}(\varphi,\psi) := q^N_{0,N}(\varphi,\psi) + \frac{1}{N} &\sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}^{(\text{no triple})}_\varepsilon} \sum_{\mathsf{b}_0,\mathsf{a}_{k+1}\in\mathbb{Z}^2} \sum_{(\mathsf{a}_1,\mathsf{b}_1,\dots,\mathsf{a}_k,\mathsf{b}_k)\in\mathcal{A}^{(\text{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}}} \\
&\sum_{\substack{(d_1,x_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{a}_1)\\(f_1,y_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{b}_1)\\ \text{with } d_1\leqslant f_1}} \cdots \sum_{\substack{(d_k,x_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{a}_k)\\(f_k,y_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{b}_k)\\ \text{with } d_k\leqslant f_k}} \sum_{y_0\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0)\cap\mathbb{Z}^2_{\text{even}}} \varphi_N(y_0) \qquad (5.40) \\
&\times \Bigg\{\prod_{j=1}^{k} q_{f_{j-1},d_j}(y_{j-1},x_j)\, X^{(\text{diff})}_{d_j,f_j}(x_j,y_j)\Bigg\} \sum_{x_{k+1}\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_{k+1})\cap\mathbb{Z}^2_{\text{even}}} q_{f_k,N}(y_k,x_{k+1})\,\psi_N(x_{k+1})\,,
\end{aligned}$$

where $f_0 := 0$. The main result of this subsection is the following, where the approximation error is much smaller than $\mathbb{V}\text{ar}(\mathcal{Z}^{\beta_N}_{N,1}(\varphi,\psi))$ in (3.55).

Lemma 5.6 (Diffusive bound). *Recall from* (4.2) *that* $M_\varepsilon = \log\log\frac{1}{\varepsilon}$ *and recall* $\|\cdot\|_{\mathcal{G}_t}$ *from* (3.52). *There exist* $\mathsf{c}, \mathsf{C} \in (0,\infty)$ *such that for* $\varepsilon > 0$ *small enough, we have: for all* φ *with* $\|\varphi\|_{\mathcal{G}_1}^2 < \infty$ *and* $\psi \in L^\infty(\mathbb{R}^2)$,

$$\limsup_{N\to\infty \text{ with } N\in 2\mathbb{N}} \big\| \big(\mathcal{Z}_{N,\varepsilon}^{(\mathrm{diff})} - \mathcal{Z}_{N,\varepsilon}^{(\mathrm{no\,triple})}\big)(\varphi,\psi)\big\|_{L^2}^2 \leqslant \mathsf{C}\, e^{-\mathsf{c}M_\varepsilon^2} \|\varphi\|_{\mathcal{G}_1}^2 \|\psi\|_\infty^2. \tag{5.41}$$

Proof. We argue as in the proof of Lemma 5.1. Note that the chaos expansion for $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{diff})}$ *is a restricted sum of terms in the expansion for* $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{no\,triple})}$ due to the following two effects:

(I) the replacement of $X_{d_j,f_j}(x_j,y_j)$ by $X_{d_j,f_j}^{(\mathrm{diff})}(x_j,y_j)$ (cf. (4.10)-(4.9) and (3.29));

(II) the restriction of $(\mathsf{a}_1, \mathsf{b}_1, \dots, \mathsf{a}_k, \mathsf{b}_k)$ to $\mathcal{A}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}}^{(\mathrm{diff})}$.

Since terms in the chaos expansion are mutually L^2 orthogonal, we can write

$$\big\| \big(\mathcal{Z}_{N,\varepsilon}^{(\mathrm{diff})} - \mathcal{Z}_{N,\varepsilon}^{(\mathrm{no\,triple})}\big)(\varphi,\psi)\big\|_{L^2}^2 \;=\; \mathrm{I}_{N,\varepsilon} \,+\, \mathrm{II}_{N,\varepsilon}\,, \tag{5.42}$$

where $\mathrm{I}_{N,\varepsilon}$ and $\mathrm{II}_{N,\varepsilon}$ are the squared L^2 error as we first make the replacement in (I) and then impose the restriction in (II).

To be more precise, we can define $X_{d,f}^{(\mathrm{superdiff})}(x,y)$ by the equality

$$X_{d,f}(x,y) = X_{d,f}^{(\mathrm{diff})}(x,y) + X_{d,f}^{(\mathrm{superdiff})}(x,y)\,. \tag{5.43}$$

In view of (3.30) and (3.40), we define

$$\overline{U}_{N,\,d,f}^{(\mathrm{diff})}(x,y) := \mathbb{E}\big[X_{d,f}^{(\mathrm{diff})}(x,y)^2\big]\,, \tag{5.44}$$

$$\overline{U}_{N,\,d,f}^{(\mathrm{superdiff})}(x,y) := \mathbb{E}\big[X_{d,f}^{(\mathrm{superdiff})}(x,y)^2\big]\,. \tag{5.45}$$

Note that $X_{d,f}^{(\mathrm{superdiff})}(x,y)$ and $X_{d,f}^{(\mathrm{diff})}(x,y)$ are orthogonal in L^2, see (4.10)-(4.9) and (3.29). As a consequence, if we plug (5.43) into (5.37) and expand the product, we obtain

$$\begin{aligned}
\mathrm{I}_{N,\varepsilon} \;=\; \frac{1}{N^2} \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}_\varepsilon^{(\mathrm{no\,triple})}} \sum_{\substack{\mathsf{b}_0,\mathsf{b}_0{}'\\ \mathsf{a}_1,\mathsf{b}_1,\dots,\mathsf{a}_k,\mathsf{b}_k}\in\mathbb{Z}^2} \sum_{\substack{(d_1,x_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{a}_1)\\ (f_1,y_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{b}_1)\\ \text{with } d_1\leqslant f_1}} \cdots \sum_{\substack{(d_k,x_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{a}_k)\\ (f_k,y_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{b}_k)\\ \text{with } d_k\leqslant f_k}} \\
\times \Bigg(\sum_{\substack{v\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0)\cap\mathbb{Z}^2_{\mathrm{even}}\\ v'\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0{}')\cap\mathbb{Z}^2_{\mathrm{even}}}} \varphi_N(v)\,\varphi_N(v')\; q_{0,d_1}(v,x_1)\, q_{0,d_1}(v',x_1) \Bigg) \\
\times \Bigg\{ \prod_{j=2}^{k} q_{f_{j-1},d_j}(y_{j-1},x_j)^2 \Bigg\} \Bigg(\sum_{w\in\mathbb{Z}^2_{\mathrm{even}}} q_{f_k,N}(y_k,w)\,\psi_N(w) \Bigg)^2 \\
\times \sum_{J\subseteq\{1,\dots,k\}:\,|J|\geqslant 1} \prod_{j\in J} \overline{U}_{N,\,d_j,f_j}^{(\mathrm{superdiff})}(x_j,y_j) \prod_{j\in\{1,\dots,k\}\setminus J} \overline{U}_{N,\,d_j,f_j}^{(\mathrm{diff})}(x_j,y_j)\,.
\end{aligned} \tag{5.46}$$

The term $\mathrm{II}_{N,\varepsilon}$ in (5.42) accounts for the further restriction to $(\mathsf{a}_1,\,\mathsf{b}_1,\,\dots\,,\mathsf{a}_k,\,\mathsf{b}_k)\in \mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}}\cap\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0',\mathsf{a}_{k+1}'}$ for some $\mathsf{b}_0,\mathsf{b}_0{}',\mathsf{a}_{k+1},\mathsf{a}_{k+1}'\in\mathbb{Z}^2$, and hence

$$\begin{aligned}\mathrm{II}_{N,\varepsilon} \;=\frac{1}{N^2}\sum_{k=1}^{(\log\frac{1}{\varepsilon})^2}\;\sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}_\varepsilon^{(\text{no triple})}}\;\sum_{\substack{\mathsf{b}_0,\mathsf{b}_0{}',\mathsf{a}_{k+1},\mathsf{a}_{k+1}'\in\mathbb{Z}^2\\ (\mathsf{a}_1,\,\mathsf{b}_1,\,\dots\,,\mathsf{a}_k,\,\mathsf{b}_k)\in(\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}})^c\cap(\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0{}',\mathsf{a}_{k+1}'})^c}}\;\sum_{\substack{(d_1,x_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{a}_1)\\ (f_1,y_1)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_1,\mathsf{b}_1)\\ \text{with } d_1\leqslant f_1}}\cdots\sum_{\substack{(d_k,x_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{a}_k)\\ (f_k,y_k)\in\mathcal{B}_{\varepsilon N}(\mathsf{i}_k,\mathsf{b}_k)\\ \text{with } d_k\leqslant f_k}}\\ \Bigg(\sum_{\substack{v\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0)\cap\mathbb{Z}^2_{\mathrm{even}}\\ v'\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0{}')\cap\mathbb{Z}^2_{\mathrm{even}}}}\varphi_N(v)\,\varphi_N(v')\;q_{0,d_1}(v,x_1)\,q_{0,d_1}(v',x_1)\Bigg)\\ \times\quad \overline{U}^{(\mathrm{diff})}_{N,\,d_1,f_1}(x_1,y_1)\Bigg\{\prod_{j=2}^{k}q_{f_{j-1},d_j}(y_{j-1},x_j)^2\,\overline{U}^{(\mathrm{diff})}_{N,\,d_j,f_j}(x_j,y_j)\Bigg\}\\ \times\Bigg(\sum_{\substack{x_{k+1}\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_{k+1})\cap\mathbb{Z}^2_{\mathrm{even}}\\ x'_{k+1}\in\mathcal{S}_{\varepsilon N}(\mathsf{a}'_{k+1})\cap\mathbb{Z}^2_{\mathrm{even}}}}q_{f_k,N}(y_k,x_{k+1})q_{f_k,N}(y_k,x'_{k+1})\psi_N(x_{k+1})\psi_N(x'_{k+1})\Bigg).\end{aligned}\tag{5.47}$$

To prove (5.41), it suffices to show that for some $\mathsf{c},\mathsf{C}\in(0,\infty)$,

$$\limsup_{N\to\infty\text{ with }N\in2\mathbb{N}}\mathrm{I}_{N,\varepsilon}\leqslant\mathsf{C}e^{-\mathsf{c}M_\varepsilon^2}\|\varphi\|^2_{\mathcal{G}_1}\|\psi\|^2_\infty\quad\text{and}\quad\limsup_{N\to\infty\text{ with }N\in2\mathbb{N}}\mathrm{II}_{N,\varepsilon}\leqslant\mathsf{C}e^{-\mathsf{c}M_\varepsilon^2}\|\varphi\|^2_{\mathcal{G}_1}\|\psi\|^2_\infty.\tag{5.48}$$

We need the following bound, which follows easily from Lemma 3.5. Recall $U_N(n)$ and $\overline{U}_N(n)$ from (3.35) and (3.40).

$$\begin{aligned}&\exists\,c,C\in(0,\infty)\;\;s.t.\;\;\forall\,N\in\mathbb{N},\;\varepsilon>0,\;d\leqslant f\in\mathbb{N}\text{ with }|f-d|\leqslant\varepsilon N:\\ &\qquad\sum_{y\in\mathbb{Z}^2}\overline{U}^{(\mathrm{superdiff})}_{N,\,d,f}(x,y)\leqslant C\,e^{-cM_\varepsilon^2}\,\overline{U}_N(f-d)\,.\end{aligned}\tag{5.49}$$

We are ready to bound $\mathrm{I}_{N,\varepsilon}$ in (5.46). As in the proof of Lemma 5.1, the last term $\big(\sum_{w\in\mathbb{Z}^2}q_{f_k,N}(y_k,w)\,\psi_N(w)\big)^2$ can be bounded as in (5.11). We then sum over all space variables in reverse order from y_k,x_k until y_1. We will use $\sum_{x\in\mathbb{Z}^2}q_{f,d}(y,x)^2=u(d-f)$ by (3.3), apply (5.49) together with the fact that

$$\sum_{y\in\mathbb{Z}^2}\overline{U}^{(\mathrm{diff})}_{N,\,d,f}(x,y)\leqslant\sum_{y\in\mathbb{Z}^2}\overline{U}_N(f-d,y-x)=\overline{U}_N(f-d)\,,\tag{5.50}$$

and finally sum over x_1, noting that, by (5.12),

$$\begin{aligned}&\sum_{\mathsf{a}_1\in\mathbb{Z}^2}\sum_{x_1\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_1)}\sum_{\mathsf{b}_0,\mathsf{b}_0{}'\in\mathbb{Z}^2}\sum_{\substack{v\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0)\cap\mathbb{Z}^2_{\mathrm{even}}\\ v'\in\mathcal{S}_{\varepsilon N}(\mathsf{b}_0{}')\cap\mathbb{Z}^2_{\mathrm{even}}}}\varphi_N(v)\,\varphi_N(v')\;q_{0,d_1}(v,x_1)\,q_{0,d_1}(v',x_1)\\ &=\sum_{v,v'\in\mathbb{Z}^2_{\mathrm{even}}}\varphi_N(v)\,\varphi_N(v')\,q_{2d_1}(v-v')\;=\;N\,\Phi_N\Big(\frac{d_1}{N}\Big).\end{aligned}\tag{5.51}$$

We then obtain an analogue of (5.13):

$$\begin{aligned}\mathrm{I}_{N,\varepsilon} \;\leqslant\; \frac{C\,\|\psi\|_\infty^2}{N} \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{(\mathsf{i}_1,\dots,\mathsf{i}_k)\in\mathcal{A}_\varepsilon^{(\text{no triple})}} \sum_{d_1\leqslant f_1\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_1),\,\dots,\,d_k\leqslant f_k\in\mathcal{T}_{\varepsilon N}(\mathsf{i}_k)} \\ \Bigg[\sum_{J\subseteq\{1,\dots,k\}:\,|J|\geqslant 1} C\,e^{-cM_\varepsilon^2|J|}\Bigg] \Phi_N\Big(\frac{d_1}{N}\Big)\,\overline{U}_N(f_1-d_1)\Bigg\{\prod_{j=2}^{k} u(d_j-f_{j-1})\,\overline{U}_N(f_j-d_j)\Bigg\}.\end{aligned} \tag{5.52}$$

For $k\leqslant(\log\frac{1}{\varepsilon})^2$, recalling that $M_\varepsilon=\log\log\frac{1}{\varepsilon}$, we can bound for $\varepsilon>0$ small enough

$$\Bigg[\sum_{J\subseteq\{1,\dots,k\}:\;|J|\geqslant 1} C\,e^{-cM_\varepsilon^2|J|}\Bigg] = C\left\{(1+e^{-cM_\varepsilon^2})^k-1\right\} \leqslant 2Ck\,e^{-cM_\varepsilon^2} \leqslant 2Ce^{-\frac{c}{2}M_\varepsilon^2}.$$

We now plug this bound into (5.52) and sum freely over all $0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}$. As $N\to\infty$, by Lemma 5.3 and similar to (5.21), we obtain (with $\mathsf{c}=\frac{c}{2}$)

$$\limsup_{N\to\infty}\;\mathrm{I}_{N,\varepsilon} \;\leqslant\; C\,\|\psi\|_\infty^2\,e^{-\mathsf{c}M_\varepsilon^2} \sum_{k=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}} \mathcal{I}_\varepsilon^\Phi(\mathsf{i}_1,\dots,\mathsf{i}_k)\,, \tag{5.53}$$

The renewal decomposition (5.20), together with $c_\vartheta=\int_0^1 G_\vartheta(x)\,\mathrm{d}x<\infty$, yields

$$\sum_{k=1}^{\infty} \sum_{0<\mathsf{i}_1<\ldots<\mathsf{i}_k\leqslant\frac{1}{\varepsilon}} \mathcal{I}_\varepsilon^\Phi(\mathsf{i}_1,\dots,\mathsf{i}_k) = \iint\limits_{0<s<t<1} \Phi(s)\,G_\vartheta(t-s)\,\mathrm{d}s\,\mathrm{d}t \leqslant c_\vartheta\int_0^1\Phi(s)\,\mathrm{d}s\,. \tag{5.54}$$

By (5.15), this proves the first bound in (5.48).

We now prove the bound on $\mathrm{II}_{N,\varepsilon}$ in (5.48). Note from (5.39) that

$$(\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}})^c = \bigcup_{l=1}^{k}\big\{|\mathsf{b}_l-\mathsf{a}_l|>M_\varepsilon\big\} \;\cup\; \bigcup_{j=1}^{k+1}\Big\{|\mathsf{a}_j-\mathsf{b}_{j-1}|>M_\varepsilon\sqrt{\mathsf{i}_j-\mathsf{i}_{j-1}}\Big\}. \tag{5.55}$$

The first union in (5.55) gives no contribution to (5.47) since $\overline{U}^{(\mathrm{diff})}_{N,d,f}(x,y)=0$ when $x\in\mathcal{S}_{\varepsilon N}(\mathsf{a})$ and $y\in\mathcal{S}_{\varepsilon N}(\mathsf{b})$ with $|\mathsf{b}-\mathsf{a}|>M_\varepsilon$, by (5.38) and (5.44). It remains to consider the contribution to (5.47) from the second union in (5.55), namely, $|\mathsf{a}_j-\mathsf{b}_{j-1}|>M_\varepsilon\sqrt{\mathsf{i}_j-\mathsf{i}_{j-1}}$ for some $j\in\{1,\dots,k,k+1\}$.

In contrast to the bound on $\mathrm{I}_{N,\varepsilon}$ where a small factor $Ce^{-cM_\varepsilon^2}$ comes from the bound on $\sum_{y_j\in\mathbb{Z}^2}\overline{U}^{(\mathrm{superdiff})}_{N,\,d_j,f_j}(x_j,y_j)$ in (5.49), in the bound for $\mathrm{II}_{N,\varepsilon}$, the same small factor now comes from the following estimates: there exists $c\in(0,\infty)$ such that for any $\mathsf{i}<\mathsf{i}'$, $f\in\mathcal{T}_{\varepsilon N}(\mathsf{i})$, $d'\in\mathcal{T}_{\varepsilon N}(\mathsf{i}')$, $\mathsf{b}\in\mathbb{Z}^2$, and $y\in\mathcal{S}_{\varepsilon N}(\mathsf{b})$, we have

$$\sum_{\mathsf{a}'\in\mathbb{Z}^2:\,|\mathsf{a}'-\mathsf{b}|>M_\varepsilon\sqrt{\mathsf{i}'-\mathsf{i}}} \sum_{x'\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')} q_{f,d'}(y,x') \;\leqslant\; C\,e^{-cM_\varepsilon^2}, \tag{5.56}$$

$$\sum_{\mathsf{a}'\in\mathbb{Z}^2:\,|\mathsf{a}'-\mathsf{b}|>M_\varepsilon\sqrt{\mathsf{i}'-\mathsf{i}}} \sum_{x'\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')} q_{f,d'}(y,x')^2 \;\leqslant\; C\,e^{-cM_\varepsilon^2}\,u(d'-f)\,, \tag{5.57}$$

where recall from (3.3) that $u(n) := \sum_{z\in\mathbb{Z}^2} q_n(z)^2 = q_{2n}(0)$. The first bound follows from the fact that $q_n(\cdot)$ has Gaussian tail decay. The second bound is a consequence of the first bound, because $\sup_z q_n(z) \leqslant \frac{C'}{n} \leqslant Cu(n)$ by the local limit theorem (3.21) and (3.3).

The bound on $\mathrm{II}_{N,\varepsilon}$ then follows the same steps as that for $\mathrm{I}_{N,\varepsilon}$, where we take a union bound over all $1 \leqslant j \leqslant k+1 \leqslant (\log\frac{1}{\varepsilon})^2 + 1$ with $|\mathsf{a}_j - \mathsf{b}_{j-1}| > M_\varepsilon\sqrt{\mathsf{i}_j - \mathsf{i}_{j-1}}$.

- For $j = k+1$, bounding ψ by $\|\psi\|_\infty$ and applying (5.56), the sum over w, w' and $\mathsf{a}_{k+1}, \mathsf{a}'_{k+1}$ in (5.47) under the super-diffusive constraint $|\mathsf{a}_{k+1} - \mathsf{b}_k| > M_\varepsilon\sqrt{\frac{1}{\varepsilon} - \mathsf{i}_k}$ leads to an extra factor of $Ce^{-cM_\varepsilon^2}$ compared with the bound when this constraint is not present (see (5.11)).
- For $2 \leqslant j \leqslant k$, by (5.57), the sum of $q_{f_{j-1},d_j}(y_{j-1}, x_j)^2$ over x_j in (5.47) under this super-diffusive constraint gives an extra factor of $Ce^{-cM_\varepsilon^2}$ compared to the case when this constraint is not present.
- For $j = 1$, given $\mathsf{b}_0, \mathsf{b}'_0 \in \mathbb{Z}^2$, we could have either $|\mathsf{a}_1 - \mathsf{b}_0| > M_\varepsilon\sqrt{\mathsf{i}_1}$ or $|\mathsf{a}_1 - \mathsf{b}'_0| > M_\varepsilon\sqrt{\mathsf{i}_1}$; either way, given $v \in \mathcal{S}_{\varepsilon N}(\mathsf{b}_0)$, $v' \in \mathcal{S}_{\varepsilon N}(\mathsf{b}'_0)$ and $d_1 \in \mathcal{T}_{\varepsilon N}(\mathsf{i}_1)$, the sum of $q_{0,d_1}(v, x_1)q_{0,d_1}(v', x_1)$ in (5.47) under this super-diffusive constraint gives a factor $Ce^{-cM_\varepsilon^2}$ compared to the case when this constraint is not present.

Since there are at most $(\log\frac{1}{\varepsilon})^2 + 1$ choices of such j, this leads to the same bound we had for $\mathrm{I}_{N,\varepsilon}$ in (5.53), which establishes the second inequality in (5.48).

This completes the proof of Lemma 5.6. □

5.3. Step 3: Kernel replacement. In this step, we introduce a last approximation $\mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ and show that it is close both to $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ and to the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$. This completes the proof of Theorem 4.7.

Let us first summarize the previous steps. So far, we have performed coarse-graining by grouping terms in the chaos expansion for $\mathcal{Z}_N(\varphi,\psi)$ in (3.19) according to the mesoscopic time-space boxes $\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})$ visited by the microscopic time-space renewal configuration $(n_1, z_1), \ldots, (n_r, z_r)$ in (3.19). Imposing suitable restrictions, we have defined the approximations $\mathcal{Z}^{(\text{no triple})}_{N,\varepsilon}(\varphi,\psi)$ in (5.3) and $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ in (5.40). The next step is to replace the relevant random walk transition kernels in the expansion (5.40) for $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ by heat kernels as in (2.6), i.e., replace the random walk transition kernels $q_{f_{j-1},d_j}(y_{j-1}, x_j)$ connecting the microscopic points $(f_{j-1}, y_{j-1}) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}_{j-1}, \mathsf{b}_{j-1})$ and $(d_j, x_j) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}_j, \mathsf{a}_j)$ by a heat kernel that depends on the mesoscopic time-space variables $(\mathsf{i}_{j-1}, \mathsf{b}_{j-1})$ and $(\mathsf{i}_j, \mathsf{a}_j)$. However, such an approximation is only good if $\mathsf{i}_j - \mathsf{i}_{j-1}$ is sufficiently large, say at least $K_\varepsilon = (\log\frac{1}{\varepsilon})^6$ as in (4.2).

This naturally leads to the decomposition of $(\mathsf{i}_1, \ldots, \mathsf{i}_k)$ into time blocks, where consecutive i_{j-1}, i_j with distance less than K_ε are grouped into a single block. The constraint $\mathcal{A}^{(\text{no triple})}_\varepsilon$ in (5.2) ensures that each time block consists of either a single i_j, or two consecutive i_{j-1}, i_j, leading to the definition of *time blocks* $\vec{\mathsf{i}}$ in Definition 4.1, while $\mathcal{A}^{(\text{no triple})}_\varepsilon$ is mapped to the constraint $\vec{\mathcal{A}}^{(\text{no triple})}_\varepsilon$ introduced in (4.4) for a sequence of time blocks. Given a sequence of time blocks $(\vec{\mathsf{i}}_1, \ldots, \vec{\mathsf{i}}_r)$ visited by the microscopic time-space renewal configuration, for each time block $\vec{\mathsf{i}}_l = (\mathsf{i}_l, \mathsf{i}'_l)$, we can identify the first mesoscopic box $\mathcal{S}_{\varepsilon N}(\mathsf{a}_l)$ visited by the

renewal configuration in the time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i}_l)$, as well as the last mesoscopic box $\mathcal{S}_{\varepsilon N}(\mathsf{a}'_l)$ visited by the renewal configuration in the time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i}'_l)$. This produces a *space block* $\vec{\mathsf{a}}_l := (\mathsf{a}_l, \mathsf{a}'_l)$ as in Definition 4.2.

Summarizing: we can *rewrite the expansion for* $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ *in* (5.40) *according to the sequence of visited time-space blocks* $(\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1), \dots, (\vec{\mathsf{i}}_r, \vec{\mathsf{a}}_r)$, where the diffusive constraint $\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ in (5.39) is mapped to $\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ defined in (4.6) for the sequence of space blocks. See Figure 2.

We are ready to define our last approximation $\mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$: having rewritten the expansion (5.40) for $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ in terms of the visited time-space blocks $(\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1), \dots, (\vec{\mathsf{i}}_r, \vec{\mathsf{a}}_r)$, we replace *each random walk transition kernel connecting two consecutive time-space blocks* by a heat kernel depending only on the mesoscopic variables $(\vec{\mathsf{i}}_\cdot, \vec{\mathsf{a}}_\cdot)$. More precisely, given $(f_{j-1}, y_{j-1}) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}'_{j-1}, \mathsf{a}'_{j-1})$ and $(d_j, x_j) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}_j, \mathsf{a}_j)$, we make within (5.40) the replacement

$$q_{f_{j-1},d_j}(y_{j-1}, x_j) \quad \rightsquigarrow \quad \frac{2}{\varepsilon N}\, g_{\frac{1}{2}(\mathsf{i}_j - \mathsf{i}'_{j-1})}(\mathsf{a}_j - \mathsf{a}'_{j-1}), \tag{5.58}$$

where the prefactor 2 is due to periodicity (note that $\mathsf{i}_j - \mathsf{i}'_{j-1} \geqslant K_\varepsilon$ by the constraint $\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}$ from (4.4)). We similarly replace the "boundary kernels" in (5.40), namely

$$q^N_{0,N}(\varphi,\psi) \;\rightsquigarrow\; \frac{1}{2}\, g_{\frac{1}{2}}(\varphi,\psi)\,, \tag{5.59}$$

$$q_{0,d_1}(y_0, x_1) \;\rightsquigarrow\; \frac{2}{\varepsilon N}\, g_{\frac{1}{2}\mathsf{i}_1}(\mathsf{a}_1 - \mathsf{b}_0) \qquad \text{for} \quad \begin{array}{l} y_0 \in \mathcal{S}_{\varepsilon N}(\mathsf{b}_0)\,, \\ (d_1, x_1) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}_1, \mathsf{a}_1)\,, \end{array} \tag{5.60}$$

$$q_{f_k,N}(y_k, x_{k+1}) \;\rightsquigarrow\; \frac{2}{\varepsilon N}\, g_{\frac{1}{2}(\frac{1}{\varepsilon} - \mathsf{i}_k)}(\mathsf{a}_{k+1} - \mathsf{b}_k) \qquad \text{for} \quad \begin{array}{l} (f_k, y_k) \in \mathcal{B}_{\varepsilon N}(\mathsf{i}_k, \mathsf{b}_k)\,, \\ x_{k+1} \in \mathcal{S}_{\varepsilon N}(\mathsf{a}_{k+1})\,, \end{array} \tag{5.61}$$

where the constraint $\mathcal{A}^{(\text{no triple})}_{\varepsilon}$ (which maps to $\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}$) ensures $\mathsf{i}_1 \geqslant K_\varepsilon$ and $\frac{1}{\varepsilon} - \mathsf{i}_k \geqslant K_\varepsilon$. We thus define $\mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ as the expression obtained from (5.40) via the replacements (5.58) and (5.59)-(5.61) (this description is useful to compare $\mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ with $\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$).

An alternative, equivalent description of $\mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ is through the following formula:

$$\begin{aligned} \mathcal{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi) = \frac{1}{2} g_{\frac{1}{2}}(\varphi,\psi) + \frac{1}{N} \sum_{r=1}^{(\log\frac{1}{\varepsilon})^2} \sum_{\substack{\mathsf{a}'_0, \mathsf{a}_{r+1} \in \mathbb{Z}^2, (\vec{\mathsf{a}}_1, \dots, \vec{\mathsf{a}}_r) \in \vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\mathsf{a}'_0,\mathsf{a}_{r+1}} \\ (\vec{\mathsf{i}}_1, \dots, \vec{\mathsf{i}}_r) \in \vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}}} \Bigg(\sum_{v \in \mathcal{S}_{\varepsilon N}(\mathsf{a}'_0) \cap \mathbb{Z}^2_{\mathrm{even}}} \varphi_N(v) \Bigg) \\ \times \prod_{l=1}^{r} g_{\frac{1}{2}(\mathsf{i}_l - \mathsf{i}'_{l-1})}(\mathsf{a}_l - \mathsf{a}'_{l-1}) \Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_l, \vec{\mathsf{a}}_l) \times \Bigg(g_{\frac{1}{2}(\frac{1}{\varepsilon} - \mathsf{i}'_r)}(\mathsf{a}_{r+1} - \mathsf{a}'_r) \frac{2}{\varepsilon N} \sum_{w \in \mathcal{S}_{\varepsilon N}(\mathsf{a}_{r+1}) \cap \mathbb{Z}^2_{\mathrm{even}}} \psi_N(w) \Bigg), \end{aligned} \tag{5.62}$$

where $\mathsf{i}'_0 := 0$ and $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_j, \vec{\mathsf{a}}_j)$ is the *coarse-grained disorder variable* defined in (4.11), which collects the contributions in (5.40) from a given visited time-space block $(\vec{\mathsf{i}}_j, \vec{\mathsf{a}}_j)$, and it arises thanks to the factorisations induced by the replacements in (5.58) and (5.59)-(5.61).

Remark 5.7. *Only the prefactor* $\frac{2}{\varepsilon N}$ *arising from the last replacement* (5.61) *appears explicitly in* (5.62)*: all other factors of* $\frac{2}{\varepsilon N}$ *arising from* (5.58) *and* (5.60) *have been absorbed in the coarse-grained disorder variable* $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_j, \vec{\mathsf{a}}_j)$ *following the replaced kernel* q*, see* (4.11).

Finally, to compare $\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ with the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$ defined in (4.8), we introduce the notation

$$\varphi^{(N)}_{\varepsilon}(\mathsf{a}'_0) := \frac{2}{\varepsilon N}\Bigg(\sum_{v\in\mathcal{S}_{\varepsilon N}(\mathsf{a}'_0)\cap\mathbb{Z}^2_{\mathrm{even}}}\varphi_N(v)\Bigg), \qquad \psi^{(N)}_{\varepsilon}(\mathsf{a}_{r+1}) := \frac{2}{\varepsilon N}\Bigg(\sum_{w\in\mathcal{S}_{\varepsilon N}(\mathsf{a}_{r+1})\cap\mathbb{Z}^2_{\mathrm{even}}}\psi_N(v)\Bigg), \tag{5.63}$$

which allows us to rewrite (5.62) more compactly as

$$\begin{aligned}\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi) \;=\; \frac{1}{2}\,g_{\frac12}(\varphi,\psi) + \frac{\varepsilon}{2}\sum_{r=1}^{(\log\frac{1}{\varepsilon})^2}\sum_{\mathsf{a}'_0,\mathsf{a}_{r+1}\in\mathbb{Z}^2}\sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}\\ (\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_{\varepsilon;\mathsf{a}'_0,\mathsf{a}_{r+1}}}} \\ \varphi^{(N)}_{\varepsilon}(\mathsf{a}'_0)\times\Bigg\{\prod_{l=1}^{r} g_{\frac12(\mathsf{i}_l-\mathsf{i}'_{l-1})}(\mathsf{a}_l-\mathsf{a}'_{l-1})\,\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_l,\vec{\mathsf{a}}_l)\Bigg\}\,g_{\frac12(\frac{1}{\varepsilon}-\mathsf{i}'_r)}(\mathsf{a}_{r+1}-\mathsf{a}'_r)\,\psi^{(N)}_{\varepsilon}(\mathsf{a}_{r+1})\,.\end{aligned} \tag{5.64}$$

Compare this with $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$ in (4.8), the only difference is that φ_ε and ψ_ε in (4.8) are now replaced by $\varphi^{(N)}_\varepsilon$ and $\psi^{(N)}_\varepsilon$.

The main result of this subsection is the following L^2 approximation, which completes the proof of Theorem 4.7.

Lemma 5.8 (Coarse graining). *Recall* $\mathscr{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ *from* (5.62)*,* $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})$ *from* (4.8) *and* $\|\cdot\|_{\mathcal{G}_t}$ *from* (3.52)*. There exists* $\mathsf{C}\in(0,\infty)$ *such that for* $\varepsilon>0$ *small enough, we have: for all* φ *with* $\|\varphi\|^2_{\mathcal{G}_1}<\infty$ *and* $\psi\in L^\infty(\mathbb{R}^2)$*,*

$$\limsup_{N\to\infty \text{ with } N\in2\mathbb{N}} \big\|\big(\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)-\mathscr{Z}^{(\mathrm{diff})}_{N,\varepsilon}\big)(\varphi,\psi)\big\|^2_{L^2} \;\leqslant\; \frac{\mathsf{C}}{\log\frac{1}{\varepsilon}}\,\|\varphi\|^2_{\mathcal{G}_1}\|\psi\|^2_\infty\,, \tag{5.65}$$

$$\lim_{N\to\infty \text{ with } N\in2\mathbb{N}} \big\|\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta^{(\mathrm{cg})}_{N,\varepsilon})-\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)\big\|^2_{L^2} \;=\; 0. \tag{5.66}$$

Proof. We first prove (5.65). To define $\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ from $\mathscr{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ in (5.40), we first replaced the summation constraint $1\leqslant k\leqslant(\log\frac{1}{\varepsilon})^2$ in (5.40) (on the number of visited mesoscopic time intervals) with the constraint $1\leqslant r\leqslant(\log\frac{1}{\varepsilon})^2$ in (5.62) (on the number of visited coarse-grained disorder variables), where each coarse-grained disorder variable can visit either one or two mesoscopic time intervals. This amounts to adding some terms with $(\log\frac{1}{\varepsilon})^2<k\leqslant 2(\log\frac{1}{\varepsilon})^2$ in (5.40). The error from such additions is bounded as in (5.34) and agrees with the bound in (5.65). We then replaced the random walk kernels by heat kernels as in (5.58) and (5.59)-(5.61). We will make these replacements sequentially and control the error in each step, showing that it is bounded by the r.h.s. of (5.65).

First note that the replacement (5.59) simply changes the first term in (5.40), which is a deterministic constant. Since the l.h.s. of (5.59) converges to the r.h.s. as $N\to\infty$, see (3.26), the L^2 cost of the replacement (5.59) vanishes as $N\to\infty$, which does not contribute to the bound in (5.65).

Next note that thanks to the diffusive constraint $\mathcal{A}^{(\mathrm{diff})}_{\varepsilon;\mathsf{b}_0,\mathsf{a}_{k+1}}$ in (5.40), which maps to $\vec{\mathcal{A}}'^{(\mathrm{diff})}_{\varepsilon;\mathsf{a}'_0,\mathsf{a}_{r+1}}$ in (5.62), the replacements in (5.58) and (5.60)-(5.61) are all of the form

$$q_{s,t}(x,y) \quad \rightsquigarrow \quad \frac{2}{\varepsilon N}\, g_{\frac{1}{2}(\mathsf{i}_j-\mathsf{i})}(\mathsf{b}-\mathsf{a})$$

for some $(s,x) \in \mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})$ and $(t,y) \in \mathcal{B}_{\varepsilon N}(\mathsf{j},\mathsf{b})$, with $(t-s, y-x) \in \mathbb{Z}^3_{\mathrm{even}}$, $\mathsf{j}-\mathsf{i} \geqslant K_\varepsilon$, and $|\mathsf{b}-\mathsf{a}| \leqslant M_\varepsilon\sqrt{\mathsf{j}-\mathsf{i}}$ (recall from (4.2) that $M_\varepsilon = \log\log\frac{1}{\varepsilon}$). We can then apply the local limit theorem (3.21) and refine the bounds in Lemma 3.2 as follows:

$$\begin{aligned} q_{s,t}(x,y) &= 2g_{\frac{t-s}{2}}(y-x)\exp\Big\{O\Big(\tfrac{1}{t-s}\Big) + O\Big(\tfrac{|y-x|^4}{(t-s)^3}\Big)\Big\} = 2g_{\frac{t-s}{2}}(y-x)\, e^{O\left(\frac{M_\varepsilon^4}{\varepsilon K_\varepsilon N}\right)} \\ &= \frac{2}{\varepsilon N} g_{\frac{\mathsf{j}-\mathsf{i}}{2}}(\mathsf{b}-\mathsf{a}) \cdot \frac{\varepsilon N g_{\frac{t-s}{2}}(y-x)}{g_{\frac{\mathsf{j}-\mathsf{i}}{2}}(\mathsf{b}-\mathsf{a})}\, e^{O\left(\frac{M_\varepsilon^4}{\varepsilon K_\varepsilon N}\right)} \\ &= \frac{2}{\varepsilon N} g_{\frac{\mathsf{j}-\mathsf{i}}{2}}(\mathsf{b}-\mathsf{a})\, e^{O\left(\frac{M_\varepsilon^4}{\varepsilon K_\varepsilon N}\right)} \cdot \frac{\varepsilon N(\mathsf{j}-\mathsf{i})}{t-s} \exp\Big\{ -\frac{|y-x|^2}{t-s} + \frac{|\mathsf{b}-\mathsf{a}|^2}{\mathsf{j}-\mathsf{i}}\Big\}, \end{aligned}$$

where since $|(t-s) - \varepsilon N(\mathsf{j}-\mathsf{i})| \leqslant \varepsilon N$ and $|(y-x) - \sqrt{\varepsilon N}(\mathsf{b}-\mathsf{a})| \leqslant \sqrt{2\varepsilon N}$ we can bound

$$\frac{\varepsilon N(\mathsf{j}-\mathsf{i})}{t-s} = 1 + O\big(\tfrac{1}{K_\varepsilon}\big), \qquad -\frac{|y-x|^2}{t-s} + \frac{|\mathsf{b}-\mathsf{a}|^2}{\mathsf{j}-\mathsf{i}} = O\Big(\tfrac{M_\varepsilon}{\sqrt{K_\varepsilon}}\Big),$$

so that for some $c > 0$, uniformly in $\varepsilon > 0$ small enough and N large, we have

$$e^{-cM_\varepsilon/\sqrt{K_\varepsilon}} \leqslant \frac{q_{s,t}(x,y)}{\frac{2}{\varepsilon N} g_{\frac{\mathsf{j}-\mathsf{i}}{2}}(\mathsf{b}-\mathsf{a})} \leqslant e^{cM_\varepsilon/\sqrt{K_\varepsilon}}. \tag{5.67}$$

Namely, every time we replace a random walk kernel by the corresponding heat kernel, we introduce an error factor of $e^{\pm cM_\varepsilon/\sqrt{K_\varepsilon}}$.

We first estimate the cost of the bulk replacements (5.61). Consider each term in the sum in (5.40), which we abbreviate by $\mathcal{Z}_i$ for simplicity, where i gathers the indices k and $\mathsf{i}_j, \mathsf{a}_j, \mathsf{b}_j, (d_j, x_j), (f_j, y_j)$ for $1 \leqslant j \leqslant k$ (excluding b_0, y_0 and $\mathsf{a}_{k+1}, x_{k+1}$). Note that within each term $\mathcal{Z}_i$, we replace at most $(\log\frac{1}{\varepsilon})^2$ random walk kernels, which amounts to replacing $\mathcal{Z}_i$ by $\gamma_i\mathcal{Z}_i$ with $e^{-c(\log\frac{1}{\varepsilon})^2 M_\varepsilon/\sqrt{K_\varepsilon}} \leqslant \gamma_i \leqslant e^{c(\log\frac{1}{\varepsilon})^2 M_\varepsilon/\sqrt{K_\varepsilon}}$. We then have

$$\mathbb{E}[(\gamma_i\mathcal{Z}_i - \mathcal{Z}_i)^2] = (\gamma_i-1)^2\,\mathbb{E}[\mathcal{Z}_i^2] \leqslant C\,\big(\log\tfrac{1}{\varepsilon}\big)^4 \frac{M_\varepsilon^2}{K_\varepsilon}\,\mathbb{E}[\mathcal{Z}_i^2] \leqslant \frac{C}{\log\frac{1}{\varepsilon}}\mathbb{E}[\mathcal{Z}_i^2], \tag{5.68}$$

since $M_\varepsilon = \log\log\frac{1}{\varepsilon}$ and $K_\varepsilon = \big(\log\frac{1}{\varepsilon}\big)^6$ by (4.2). Since the terms $\mathcal{Z}_i$ in the sum in (5.40) are *mutually orthogonal*, we can sum the bound above over i and we see that the contribution of the bulk replacements to (5.65) is at most

$$\frac{C}{\log\frac{1}{\varepsilon}} \limsup_{N\to\infty \text{ with } N\in 2\mathbb{N}} \mathbb{V}\mathrm{ar}(\mathcal{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)) \;\leqslant\; \frac{C}{\log\frac{1}{\varepsilon}}\|\psi\|_\infty^2\|\varphi\|_{\mathcal{G}_1}^2, \tag{5.69}$$

where the last bound follows from (5.23). This agrees with (5.65).

We next consider the boundary replacements (5.59) and (5.61). Replacing the leftmost random walk kernel $q_{0,d_1}(y_0,x_1)$ in (5.40) by the corresponding heat kernel introduces an error factor $e^{-cM_\varepsilon/\sqrt{K_\varepsilon}} \leqslant \gamma_{y_0,(d_1,x_1)} \leqslant e^{+cM_\varepsilon/\sqrt{K_\varepsilon}}$, see (5.67), which affects the L^2 norm

by $\{\sum_{y_0}\varphi_N(y_0)(\gamma_{y_0,(d_1,x_1)}-1)\}^2 \leqslant C\frac{M_\varepsilon^2}{K_\varepsilon}\{\sum_{y_0}|\varphi_N(y_0)|\}^2$ (no disorder variable is attached to y_0). Thus, as in (5.69), the left boundary replacement contributes to (5.65) at most by

$$\frac{M_\varepsilon^2}{K_\varepsilon}\limsup_{N\to\infty \text{ with } N\in 2\mathbb{N}} \mathbb{V}\mathrm{ar}(\mathcal{Z}_{N,\varepsilon}^{(\mathrm{diff})}(|\varphi|,\psi)) \;\leqslant\; \frac{C}{(\log\frac{1}{\varepsilon})^5}\|\psi\|_\infty^2\|\varphi\|_{\mathcal{G}_1}^2,$$

which is a stronger bound than (5.65). The right boundary replacement (5.61) is controlled in a similar fashion, which completes the proof of (5.65).

Lastly, we prove (5.66). As noted before, the only difference between $\mathcal{Z}_{N,\varepsilon}^{(\mathrm{cg})}(\varphi,\psi)$ and $\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon}^{(\mathrm{cg})})$ is that $\varphi_\varepsilon,\psi_\varepsilon$ in (4.8) are replaced by $\varphi_\varepsilon^{(N)},\psi_\varepsilon^{(N)}$ in (5.64). For $\mathsf{a}\in\mathbb{Z}^2$, $\varepsilon\varphi_\varepsilon(\mathsf{a})$ is the integral of φ over the square $\mathcal{S}_\varepsilon(\mathsf{a}) = (\sqrt{\varepsilon}\mathsf{a}-(\sqrt{\varepsilon},\sqrt{\varepsilon}),\sqrt{\varepsilon}\mathsf{a}]$, by the definition of φ_ε in (4.7). On the other hand, by the definition of $\varphi_\varepsilon^{(N)}$ in (5.63) and φ_N in (3.9), $\varepsilon\varphi_\varepsilon^{(N)}$ is the integral of φ over the region $\tilde{\mathcal{S}}_\varepsilon(\mathsf{a}) := \bigcup_{v\in\mathcal{S}_\varepsilon(\mathsf{a})\cap(\mathbb{Z}^2_{\mathrm{even}}/\sqrt{N})}\{x\in\mathbb{R}^2 : |x-v|_1\leqslant\frac{1}{\sqrt{N}}\}$. The difference between $\mathcal{S}_\varepsilon(\mathsf{a})$ and $\tilde{\mathcal{S}}_\varepsilon(\mathsf{a})$ is contained in a shell of thickness $1/\sqrt{N}$ around the boundary of $\mathcal{S}_\varepsilon(\mathsf{a})$. Therefore, if $\varphi:\mathbb{R}^2\to\mathbb{R}$ is locally integrable, then $\varphi_\varepsilon^{(N)}$ converges pointwise to φ_ε as functions on $\mathbb{Z}^2$, while if $\psi:\mathbb{R}^2\to\mathbb{R}$ is also bounded, then $\psi_\varepsilon^{(N)}$ converges uniformly to ψ_ε. If φ has bounded support, then (5.66) is easily seen to hold since we already have control over $\mathbb{V}\mathrm{ar}(\mathcal{Z}_{N,\varepsilon}^{(\mathrm{cg})}(\varphi,\psi))$ that is uniform in N. General φ can then be handled by truncating its support. This concludes the proof of Lemma 5.8. □

Combining Lemmas 5.1, 5.6 and 5.8 then gives Theorem 4.7.

5.4. A second moment bound for the coarse-grained model. In this subsection, we prove a second moment bound for the coarse-grained model, which is loosely speaking the analogue of Lemma 3.5. This is needed when we bound the fourth moment of the coarse-grained model in Section 8.

First we introduce some notation. Let us define the following variants of $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$ and $\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\,\mathsf{b},\mathsf{c}}^{(\mathrm{diff})}$ (see (4.4) and (4.6)), without dependence on the boundary conditions:

$$\begin{aligned}\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})} := \Big\{ &\text{time blocks } \vec{\mathsf{i}}_1 < \ldots < \vec{\mathsf{i}}_r \quad \text{such that } |\vec{\mathsf{i}}_j|\leqslant K_\varepsilon \;\; \forall j=1,\ldots,r \\ &\qquad\qquad \text{and } \operatorname{dist}(\vec{\mathsf{i}}_{j-1},\vec{\mathsf{i}}_j)\geqslant K_\varepsilon \;\; \forall j=2,\ldots,r\Big\},\end{aligned} \tag{5.70}$$

$$\begin{aligned}\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\mathrm{diff})} := \Big\{ &\text{space blocks } \vec{\mathsf{a}}_1,\ldots,\vec{\mathsf{a}}_r \quad \text{such that } |\vec{\mathsf{a}}_j|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}_j|} \quad \forall j=1,\ldots,r, \\ &\qquad \operatorname{dist}(\vec{\mathsf{a}}_{j-1},\vec{\mathsf{a}}_j)\leqslant M_\varepsilon\sqrt{\operatorname{dist}(\vec{\mathsf{i}}_{j-1},\vec{\mathsf{i}}_j)} \quad \forall j=2,\ldots,r\Big\}.\end{aligned} \tag{5.71}$$

Recall the definition (3.40)-(3.41) of $\overline{U}(n-m,x-y)$. We introduce an analogous quantity for the coarse-grained model defined in (4.8) (illustrated in Figure 2). Given $\mathsf{n}\in\mathbb{N}_0$ and

$\mathsf{x} \in \mathbb{Z}^2$, we define a coarse-grained analogue of $X_{d,f}(x,y)$ in (3.29):

$$\mathscr{X}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) := \begin{cases} \Theta^{(\mathrm{cg})}_{N,\varepsilon}\big((0,\mathsf{n}),(0,\mathsf{x})\big) & \text{if } \mathsf{n} < K_\varepsilon\,, \\ \displaystyle\sum_{r=2}^{(\log\frac{1}{\varepsilon})^2} \sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_\varepsilon \\ \mathsf{i}_1=0,\ \mathsf{i}'_r=\mathsf{n}}} \sum_{\substack{(\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\mathrm{diff})}_\varepsilon \\ \mathsf{a}_1=0,\ \mathsf{a}'_r=\mathsf{x}}} \Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1) & \\ \qquad \times \prod_{j=2}^{r} g_{\frac{1}{2}(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\,\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j) & \text{if } \mathsf{n} \geqslant K_\varepsilon\,, \end{cases} \tag{5.72}$$

and define

$$\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) := \mathbb{E}[(\mathscr{X}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}))^2]. \tag{5.73}$$

We prove the following analogue of Lemma 3.5 (with an extra sum in the time index).

Lemma 5.9. *For every $\mathsf{c} \in (0,\infty)$, there exist $\mathsf{C} = \mathsf{C}(\mathsf{c}) \in (0,\infty)$ and $\hat\lambda_0 = \hat\lambda_0(\mathsf{c}) \in (0,\infty)$ such that: there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ and $\hat\lambda \in (\hat\lambda_0,\infty)$, we have*

$$\limsup_{N\to\infty} \sum_{\mathsf{n}=0}^{2/\varepsilon} e^{-\hat\lambda\varepsilon\mathsf{n}} \sum_{\mathsf{x}\in\mathbb{Z}^2} e^{\mathsf{c}\sqrt{\varepsilon}|\mathsf{x}|}\,\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) \leqslant \frac{\mathsf{C}}{\log\hat\lambda}\,. \tag{5.74}$$

Proof. The basic strategy is to first undo the replacement of the random walk kernels by heat kernels in the definition of the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ in Section 5.3. We then undo the summation constraints imposed in Sections 5.1 and 5.2, which allows us to bound $\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x})$, the second moment of $\mathscr{X}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x})$, in terms of the second moment of the original partition function, so that Lemma 3.5 can be applied. The details are as follows.

Let us recall how the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{N,\varepsilon}(\varphi,\psi)$ in (5.62) was defined from $\mathscr{Z}^{(\mathrm{diff})}_{N,\varepsilon}(\varphi,\psi)$ in (5.40) by replacing the random walk kernels with heat kernels in the chaos expansion (see (5.58), (5.60)-(5.61)). It was shown in the proof of Lemma 5.8, in particular, in (5.67) and (5.68), that the aggregate effect of such replacements is to change the second moment by a factor that is bounded between $1 - C/\log\frac{1}{\varepsilon}$ and $1 + C/\log\frac{1}{\varepsilon}$. We can therefore undo these replacements, which only changes the second moment by a factor that is bounded between $1 - C'/\log\frac{1}{\varepsilon}$ and $1 + C'/\log\frac{1}{\varepsilon}$ ($\leqslant 2$ for ε small). More precisely, define

$$\mathcal{X}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) := \frac{2}{\varepsilon N} \sum_{r=1}^{\infty} \sum_{\substack{n_1<\dots<n_r \\ z_1,\dots,z_r\in\mathbb{Z}^2 \\ (n_1,z_1)\in\mathcal{B}_{\varepsilon N}(0,0),\,(n_r,z_r)\in\mathcal{B}_{\varepsilon N}(\mathsf{n},\mathsf{x})}} \xi_N(n_1,z_1) \prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1},z_j)\,\xi_N(n_j,z_j),$$

which is obtained by reversing the replacements of the random walk transition kernels by heat kernels in the definition of $\mathscr{X}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x})$ in (5.72), plugging in the chaos expansion for the coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ from (4.11), (4.10) and (4.9), and then relaxing the constraints on the time-space summation indices. The pre-factor $2/\varepsilon N$ comes from the first $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ in (5.72) and is a normalising factor in its definition in (4.11). Since relaxing the summation constraint only increases the second moment because the terms of the chaos

expansion are L^2 orthogonal, we have that for ε sufficiently small and all N large,

$$\begin{aligned}\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) = \mathbb{E}[(\mathscr{X}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}))^2] &\leqslant 2\,\mathbb{E}[(\mathcal{X}_{N,\varepsilon}(\mathsf{n},\mathsf{x}))^2] \\ &= \frac{8}{(\varepsilon N)^2} \sum_{(m,y)\in\mathcal{B}_{\varepsilon N}(0,0),\,(n,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{n},\mathsf{x})} \overline{U}_N(n-m,x-y)\,, \end{aligned} \tag{5.75}$$

where we have used the definition of $\overline{U}_N(n-m,x-y)$ from (3.40)-(3.41). Substituting this bound into the l.h.s. of (5.74) then gives

$$\begin{aligned}&\sum_{\mathsf{n}=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon\mathsf{n}} \sum_{\mathsf{x}\in\mathbb{Z}^2} e^{\mathsf{c}\sqrt{\varepsilon}|\mathsf{x}|}\,\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) \\ &\leqslant \frac{8}{(\varepsilon N)^2} \sum_{(m,y)\in\mathcal{B}_{\varepsilon N}(0,0)} \sum_{\mathsf{n}=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon\mathsf{n}} \sum_{\mathsf{x}\in\mathbb{Z}^2} e^{\mathsf{c}\sqrt{\varepsilon}|\mathsf{x}|} \sum_{(n,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{n},\mathsf{x})} \overline{U}_N(n-m,x-y)\,.\end{aligned}$$

We now observe that for $(m,y)\in\mathcal{B}_{\varepsilon N}(0,0)$ and $(n,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{n},\mathsf{x})$ we have $\frac{n-m}{\varepsilon N}\in[\mathsf{n}-1,\mathsf{n}+1]$ and $\frac{|x-y|}{\sqrt{\varepsilon N}}\in[|\mathsf{x}|-\sqrt{2},|\mathsf{x}|+\sqrt{2}]$, hence $\varepsilon\mathsf{n}=\frac{n-m}{N}+O(\varepsilon)$ and $\sqrt{\varepsilon}|\mathsf{x}|=\frac{x-y}{\sqrt{N}}+O(\sqrt{\varepsilon})$. Recalling that $|\mathcal{B}_{\varepsilon N}(0,0)|=O((\varepsilon N)^2)$, the change of variables $(n-m,x-y)=(l,z)$ then yields

$$\begin{aligned}&\limsup_{N\to\infty} \sum_{\mathsf{n}=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon\mathsf{n}} \sum_{\mathsf{x}\in\mathbb{Z}^2} e^{\mathsf{c}\sqrt{\varepsilon}|\mathsf{x}|}\,\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{n},\mathsf{x}) \;\leqslant\; C\limsup_{N\to\infty} \sum_{l=0}^{3N} e^{-\hat{\lambda}\frac{l}{N}} \sum_{z\in\mathbb{Z}^2} e^{\mathsf{c}\frac{|z|}{\sqrt{N}}}\,\overline{U}_N(l,z) \\ &\leqslant\; C\limsup_{N\to\infty} \sum_{l=0}^{3N} e^{-(\hat{\lambda}-c\mathsf{c}^2)\frac{l}{N}}\overline{U}_N(l) \;\leqslant\; C\limsup_{N\to\infty} \sum_{l=1}^{3N} e^{-(\hat{\lambda}-c\mathsf{c}^2)\frac{l}{N}}\frac{1}{N}G_\vartheta\Big(\frac{l}{N}\Big) \\ &=\; C\int_0^3 e^{-(\hat{\lambda}-c\mathsf{c}^2)s}G_\vartheta(s)\mathrm{d}s \leqslant C\int_0^{\frac{1}{\sqrt{\hat{\lambda}}}} e^{-\frac{\hat{\lambda}}{2}s}G_\vartheta(s)\mathrm{d}s + C\int_{\frac{1}{\sqrt{\hat{\lambda}}}}^3 e^{-\frac{\hat{\lambda}}{2}s}G_\vartheta(s)\mathrm{d}s \;\leqslant\; \frac{\mathsf{C}}{\log\hat{\lambda}},\end{aligned}$$

where we applied Lemma 3.5 and (3.43) in the second and third inequalities, then we chose $\hat{\lambda}\geqslant 2c\mathsf{c}^2=:\hat{\lambda}_0$ and applied (3.47) in the last line. This concludes the proof of Lemma 5.9. □

6. Higher moment bounds for averaged partition functions

In this section, we bound higher moments of the averaged partition function $\mathcal{Z}^\beta_N(\varphi,\psi)$ (see (3.8) and (3.19)) in the critical window as specified in Theorem 1.1 and (3.11)-(3.12). As noted in Section 1.5 and the discussion therein on Schrödinger operators with point interactions, these bounds are very delicate in the critical window. Unlike in the sub-critical regime considered in [CSZ17b, CSZ20], where the chaos expansion of $\mathcal{Z}^\beta_N(\varphi,\psi)$ is supported (up to a small L^2 error) on chaoses of finite order independent of N, for $\beta=\beta_N$ in the critical window, the expansion is supported on chaoses of order $\log N$, so hypercontractivity can no longer be used to bound higher moments in terms of the second moment. Instead, the expansion has to be controlled with much greater care. Bounds on the third moment were first obtained in [CSZ19b]. Bounds on higher moments of the averaged solution of the mollified stochastic heat equation (continuum analogues of $\mathcal{Z}^\beta_N(\varphi,\psi)$), for $\varphi,\psi\in L^2$, were then obtained in [GQT21] using techniques from the study of Schrödinger operator with point interactions (also called Delta-Bose gas) [DFT94, DR04]. The recent work [Che21] studied the semigroup associated with the Schrödinger operator and allowed φ to be delta functions.

Our goal is to develop similar moment bounds as in [GQT21] for the averaged polymer partition function $\mathcal{Z}_N^\beta(\varphi,\psi)$. The approach of [GQT21] used explicit Fourier calculations and the underlying space-time white noise, which cannot be easily adapted to lattice models with general disorder. We develop an alternative approach, where the key ingredient is a functional inequality for the Green's function of multiple random walks on $\mathbb{Z}^2$ (see Lemma 6.8). This leads to Theorem 6.1, which is the main result of this section, where instead of working with $\varphi,\psi \in L^2$ as in [GQT21], we will work with weighted L^p–L^q spaces with $\frac{1}{p}+\frac{1}{q}=1$, which allows $\psi(y)\equiv 1$ and $\varphi(x)=\varepsilon^{-1}\mathbb{1}_{|x|\leqslant\sqrt{\varepsilon}}$ to be an approximate delta function on the scale $\sqrt{\varepsilon}$, and it also gives spatial decay if the support of φ and ψ are far apart. Our approach is robust enough that it can be applied the coarse-grained disorder variables $\Theta_{N,\varepsilon}^{(\mathrm{cg})}$, which can be seen as an averaged partition functions (see Lemma 7.2), and it can also be adapted to the coarse-grained model $\mathscr{Z}_{\varepsilon,t}^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon}^{(\mathrm{cg})})$, as we will show in Theorem 8.1.

6.1. Statement and proof. Given a countable set $\mathbb{T}$ and a function $f:\mathbb{T}\to\mathbb{R}$, we use the standard notation

$$\|f\|_{\ell^p}=\|f\|_{\ell^p(\mathbb{T})}:=\Big(\sum_{z\in\mathbb{T}}|f(z)|^p\Big)^{1/p}\quad\text{for } p\in[1,\infty),\qquad \|f\|_{\ell^\infty}:=\sup_{z\in\mathbb{T}}|f(z)|, \tag{6.1}$$

while we let $\|g\|_p$ denote the usual L^p norm of $g:\mathbb{R}^2\to\mathbb{R}$. We will ignore parity issues, since this only affects the bounds by a constant multiple: for a locally integrable function $\varphi:\mathbb{R}^2\to\mathbb{R}$, we consider its discretization $\varphi_N:\mathbb{Z}^2\to\mathbb{R}$ in (3.9) to be defined on the whole $\mathbb{Z}^2$ (rather than just on $\mathbb{Z}^2_{\mathrm{even}}$). Here is the main result of this section.

Theorem 6.1 (Higher moments). *For $N\leqslant\tilde N\in\mathbb{N}$, let $\mathcal{Z}_N^{\beta_{\tilde N}}(\varphi,\psi)=\mathcal{Z}_{N,1}^{\beta_{\tilde N}}(\varphi,\psi)$ be the averaged partition function in* (3.8)*, where $\beta=\beta_{\tilde N}=\beta_{\tilde N}(\vartheta)$ satisfies* (3.11) *for some $\vartheta\in\mathbb{R}$. Fix $p,q\in(1,\infty)$ with $\frac{1}{p}+\frac{1}{q}=1$, an integer $h\geqslant 3$, and a weight function $w:\mathbb{R}^2\to(0,\infty)$ such that $\log w$ is Lipschitz continuous. Then there exist $\mathsf{C},\mathsf{C}'<\infty$ such that, uniformly in large $N\leqslant\tilde N\in\mathbb{N}$ and locally integrable $\varphi,\psi:\mathbb{R}^2\to\mathbb{R}$, we have*

$$\begin{aligned}\Big|\mathbb{E}\Big[\Big(\mathcal{Z}_N^{\beta_{\tilde N}}(\varphi,\psi)-\mathbb{E}[\mathcal{Z}_N^{\beta_{\tilde N}}(\varphi,\psi)]\Big)^h\Big]\Big| &\leqslant \frac{\mathsf{C}}{\log(1+\frac{\tilde N}{N})}\,\frac{1}{N^h}\,\Big\|\frac{\varphi_N}{w_N}\Big\|_{\ell^p}^h\,\|\psi_N\|_{\ell^\infty}^h\,\|w_N\mathbb{1}_{B_N}\|_{\ell^q}^h\\ &\leqslant \frac{\mathsf{C}'}{\log(1+\frac{\tilde N}{N})}\,\Big\|\frac{\varphi}{w}\Big\|_p^h\,\|\psi\|_\infty^h\,\|w\mathbb{1}_B\|_q^h\,,\end{aligned} \tag{6.2}$$

where $\varphi_N,\psi_N,w_N:\mathbb{Z}^2\to\mathbb{R}$ are defined from $\varphi,\psi,w:\mathbb{R}^2\to\mathbb{R}$ by (3.9)*, we denote by $B\subseteq\mathbb{R}^2$ a ball on which ψ is supported (possibly $B=\mathbb{R}^2$), and we set $B_N:=B\sqrt{N}$.*

Theorem 6.1 will be needed later in the proof of Lemma 7.2, where we consider $N=\varepsilon\tilde N$ with $\varepsilon\in(0,1)$; this is why we allow for $\beta=\beta_{\tilde N}$ with $\tilde N\geqslant N$.

Remark 6.2. *The second line of* (6.2) *follows from the first line by Riemann sum approximation (note that $w(\frac{x+y}{\sqrt N})=(1+O(\frac{|y|}{\sqrt N}))\,w(\frac{x}{\sqrt N})$ by the Lipschitz continuity of $\log w$).*

Remark 6.3. *Typically we will let $w(x)=e^{-|x|}$, which allows $\psi\equiv 1$ provided φ decays sufficiently fast at ∞. If ψ is bounded with compact support,* (6.2) *also gives exponential spatial decay as the support of ψ moves to infinity. This answers the conjecture in* [GQT21, Remark 1.2] *in our lattice setting, which can be extended to their continuum setting.*

Remark 6.4. *Similar to* [GQT21, Theorem 1.1]*, we can show that the moments in the l.h.s. of* (6.2) *converge as* $N \to \infty$*. However, the limits are expressed as series of iterated integrals and are not very informative, so we will not state them here.*

Remark 6.5. *In the bound* (6.2)*, we could first assign* $\varphi_N, \psi_N : \mathbb{Z}^2 \to \mathbb{R}$ *and then define the corresponding* $\varphi, \psi : \mathbb{R}^2 \to \mathbb{R}$*, e.g. by piecewise constant extension* $\varphi(x) := \varphi_N(\llbracket \sqrt{N} x \rrbracket)$ *and* $\psi(y) := \psi_N(\llbracket \sqrt{N} x \rrbracket)$*, because* $\mathcal{Z}_N^\beta(\varphi, \psi)$ *depends on the functions* $\varphi, \psi : \mathbb{R}^2 \to \mathbb{R}$ *only through their discretizations* φ_N, ψ_N *in* (3.9)*, see* (3.10)*.*

In particular, we can apply the bound (6.2) *to the point-to-plane partition function* $Z_N^{\beta_N}(0)$ *defined in* (1.3)*. More precisely, we can write* $Z_N^{\beta_N}(0) = \sum_{y \in \mathbb{Z}^2} Z_N^{\beta_N}(0, y) =: \mathcal{Z}_N^{\beta_N}(\varphi, \mathbb{1})$ *with* $\varphi_N(w) = N \mathbb{1}_{\{w=0\}}$ *and* $\psi_N(z) = \mathbb{1}(z) \equiv 1$*, cf.* (3.7) *and* (3.10)*, which correspond to* $\varphi(x) := \varphi_N(\llbracket \sqrt{N} x \rrbracket) = N \mathbb{1}_{\{|x_1|+|x_2| \leqslant 1/\sqrt{N}\}}$ *and* $\psi(y) \equiv 1$*. Note that* $\|\varphi\|_p = O(N^{1-\frac{1}{p}})$*. Then, applying* (6.2) *with* $w(x) = e^{-|x|}$ *implies that for any integer* $h \geqslant 3$ *and for any* $p > 1$*, there exists* $C_{p,h} < \infty$ *such that for all* $N \in \mathbb{N}$*,*

$$\Big| \mathbb{E}\Big[\Big(Z_N^{\beta_N}(0) - \mathbb{E}\big[Z_N^{\beta_N}(0)\big] \Big)^h \Big] \Big| \;=\; \Big| \mathbb{E}\Big[\big(Z_N^{\beta_N}(0) - 1 \big)^h \Big] \Big| \;\leqslant\; C_{p,h}\, N^{h(1-\frac{1}{p})}\,. \tag{6.3}$$

Since we can take any $p > 1$*, this shows that centered moments of any order* $h \geqslant 3$ *of the point-to-plane partition function* $Z_N^{\beta_N}(0)$ *diverge as* $N \to \infty$ *more slowly than any polynomial.*

Proof of Theorem 6.1. Our starting point is the polynomial chaos expansion of $\mathcal{Z}_N^{\beta_{\tilde{N}}}(\varphi, \psi)$ as in (3.19), which gives

$$\begin{aligned} M_{N,\tilde{N},h}^{\varphi,\psi} &:= \mathbb{E}\Big[\Big(\mathcal{Z}_N^{\beta_{\tilde{N}}}(\varphi, \psi) - \mathbb{E}[\mathcal{Z}_N^{\beta_{\tilde{N}}}(\varphi, \psi)] \Big)^h \Big] \qquad (6.4) \\ &= \frac{1}{N^h} \mathbb{E}\Big[\Big(\sum_{r=1}^{\infty} \sum_{\substack{z_1,\dots,z_r \in \mathbb{Z}^2 \\ 0 < n_1 < \dots < n_r \leqslant N}} q_{0,n_1}^N(\varphi, z_1)\, \xi(n_1, z_1) \Big\{ \prod_{j=2}^{r} q_{n_{j-1},n_j}(z_{j-1}, z_j)\, \xi(n_j, z_j) \Big\} q_{n_j,N}^N(z_j, \psi) \Big)^h \Big], \end{aligned}$$

where $\xi(n, z) = \xi_{\tilde{N}}(n, z)$ is as defined in (3.14) with β_N therein replaced by $\beta_{\tilde{N}}(\vartheta)$ so that $\mathbb{V}\mathrm{ar}(\xi) = \sigma_{\tilde{N}}^2$. We will expand the h-fold product above, which gives a sum over h *microscopic time-space renewals* $(n_1^i, z_1^i), \dots, (n_{r_i}^i, z_{r_i}^i)$, $1 \leqslant i \leqslant h$. Given these h renewals, each lattice point (m, x) will contribute a factor of $\mathbb{E}[\xi(m, x)^{\#}]$, where $\#$ is the number of times (m, x) appears among the h time-space renewals. Recall ξ and $\sigma_{\tilde{N}}$ from (3.14) and (3.11), we have

$$\mathbb{E}[\xi(m, x)] = 0, \quad \mathbb{E}[\xi(m, x)^2] = \sigma_{\tilde{N}}^2 \sim \frac{\pi}{\log \tilde{N}}, \quad \big| \mathbb{E}[\xi(m, x)^l] \big| \leqslant C \sigma_{\tilde{N}}^l \quad \text{for } l \geqslant 3. \tag{6.5}$$

Therefore a given configuration of h time-space renewals will give a non-zero contribution to the expansion in (6.4) if each (m, x) is visited by none or by at least two of the h renewals. We will rewrite the expansion by first summing over all possible choices of the set of time coordinates $\bigcup_{i=1}^h \{n_1^i, \dots, n_{r_i}^i\}$, then for each time n in this set, sum over the locations $x \in \mathbb{Z}^2$ such that (n, x) is visited by (at least two) of the h renewals $(n_1^i, z_1^i), \dots, (n_{r_i}^i, z_{r_i}^i)$, and lastly, determine which of the h renewals visits (n, x). Note that in the expansion (6.4), for each of the h renewal sequences that visits (n, x), there is a random walk transition kernel q entering (n, x) and another one exiting (n, x), while for each renewal that does not visit the

time plane $\{(n,y) : y \in \mathbb{Z}^2\}$, there is a transition kernel $q_{a,b}(x,z)$ with $a < n < b$, for which we will use Chapman-Kolmogorov to rewrite it as $q_{a,b}(x,z) = \sum_{y\in\mathbb{Z}^2} q_{a,n}(x,y)q_{n,b}(y,z)$.[†]

To expand the centred moment $M^{\varphi,\psi}_{N,\tilde{N},h}$ in (6.4) as described above, we first introduce some notation. Given $h \geqslant 2$, let $I \vdash \{1,...,h\}$ denote a partition $I = I(1) \sqcup \cdots \sqcup I(m)$ of $\{1,...,h\}$ into disjoint subsets $I(1),...,I(m)$ with cardinality $|I| = m$. Write $k \overset{I}{\sim} l$ if k and l belong to the same partition element of I. The interpretation is that, for a given time n, we have $k \overset{I}{\sim} l$ if the k-th and l-th time-space renewals visit the same time-space point (n,x) for some $x \in \mathbb{Z}^2$, which leads to a power of the disorder variable $\xi(n,x)$. Given $I \vdash \{1,\dots,h\}$, denote

$$\mathbb{E}[\xi^I] := \prod_{1\leqslant j\leqslant |I|, |I(j)|\geqslant 2} \mathbb{E}[\xi^{|I(j)|}]. \tag{6.6}$$

For $\boldsymbol{x} \in (\mathbb{Z}^2)^h$, we denote

$$\boldsymbol{x} \sim I \quad \text{if } x_k = x_l \ \forall\, k \overset{I}{\sim} l. \tag{6.7}$$

For $\boldsymbol{x}, \tilde{\boldsymbol{x}} \in (\mathbb{Z}^2)^h$, denote the h-component random walk transition probabilities by

$$Q_t(\boldsymbol{x},\tilde{\boldsymbol{x}}) := \prod_{i=1}^h q_t(\tilde{x}_i - x_i), \ \ Q^N_t(\varphi,\boldsymbol{x}) := \prod_{i=1}^h q^N_{0,t}(\varphi,x_i), \ \ Q^N_t(\boldsymbol{x},\psi) := \prod_{i=1}^h q^N_{0,t}(x_i,\psi), \tag{6.8}$$

where $q^N_{0,t}(\varphi,x_i)$ and $q^N_{0,t}(x_i,\psi)$ are defined in (3.16)–(3.17), and for $I,J \vdash \{1,\dots,h\}$, denote

$$Q^{I,J}_t(\boldsymbol{x},\tilde{\boldsymbol{x}}) := \mathbb{1}_{\{\boldsymbol{x}\sim I, \tilde{\boldsymbol{x}}\sim J\}} Q_t(\boldsymbol{x},\tilde{\boldsymbol{x}}). \tag{6.9}$$

We can then write

$$\begin{aligned} M^{\varphi,\psi}_{N,\tilde{N},h} = \frac{1}{N^h} \sum_{r=1}^{\infty} \sum_{\substack{1\leqslant n_1<\cdots<n_r\leqslant n_{r+1}:=N \\ I_1,\dots,I_r\vdash\{1,\dots,h\}, m_i:=|I_i|<h \\ \boldsymbol{y}_1,\dots,\boldsymbol{y}_r\in(\mathbb{Z}^2)^h}} & Q^N_{n_1}(\varphi,\boldsymbol{y}_1)\mathbb{1}_{\{\boldsymbol{y}_1\sim I_1\}}\mathbb{E}[\xi^{I_1}] \\ & \times \prod_{i=2}^r Q^{I_{i-1},I_i}_{n_i-n_{i-1}}(\boldsymbol{y}_{i-1},\boldsymbol{y}_i)\mathbb{E}[\xi^{I_i}] \times \mathbb{1}_{\{\boldsymbol{y}_r\sim I_r\}} Q^N_{n_{r+1}-n_r}(\boldsymbol{y}_r,\psi). \end{aligned} \tag{6.10}$$

First note that we can bound $|M^{\varphi,\psi}_{N,\tilde{N},h}|$ from above by replacing $\mathbb{E}[\xi^{I_i}]$, φ and ψ with their absolute values. To simplify notation, we assume from now on $\mathbb{E}[\xi^{I_i}]$, φ and ψ are all non-negative.

Next, we bound ψ by $\|\psi\|_\infty \mathbb{1}_B$, where B is the ball of radius $\varrho \in [1,\infty]$ containing the support of ψ. We then note that, uniformly in $1 \leqslant n_r \leqslant N \leqslant n_{r+1} \leqslant 2N$ and $\boldsymbol{y}_r$, we can find $C > 0$ such that

$$Q^N_{N-n_r}(\boldsymbol{y}_r, \|\psi\|_\infty \mathbb{1}_B) \leqslant C\, Q^N_{n_{r+1}-n_r}(\boldsymbol{y}_r, \|\psi\|_\infty \mathbb{1}_B). \tag{6.11}$$

Recalling the definition of $q^N_{0,t}(y, \|\psi\|_\infty \mathbb{1}_B)$ from (3.17), this bound follows readily from the observation that, given that a random walk starting from y reaches $\sqrt{N}B$ at time $N - n_r$, the probability of being inside $\sqrt{N}B$ at time $n_{r+1} - n_r \in [0, 2N]$ is uniformly bounded away

[†]This is the key difference between the expansions in [CSZ19b] and [GQT21]. This decomposition was used in [GQT21], which allows a functional analytic interpretation of the iterated sums and helps bypass the combinatorial complexity encountered in [CSZ19b], which the authors could control for the third moment but seemed intractable for higher moments.

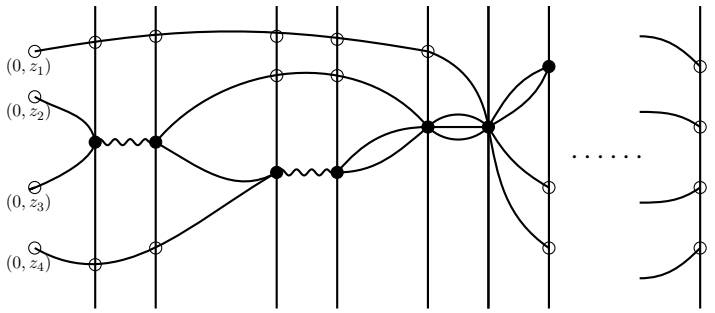


FIGURE 3. An illustration of the expansion for the fourth moment in (6.10). Solid dots are assigned weight $\mathbb{E}[\xi^{\#}]$ with # being the number of renewal sequences visiting the lattice site, while circles are assigned weight 1 and arise from the Chapman- Kolmogorov decomposition of $q_{a,b}(x,y)$ at an intermediate time. Curly lines between sites $(a,x),(b,y)$ (together with the solid dots at both ends) represent $U_{\tilde{N}}(b-a,y-y)$ as in (6.14) and (3.41), while solid lines between sites (either solid dots or circles) $(a,x),(b,y)$ are assigned weight $q_{a,b}(x,y)$. As an illustration of the expansion in (6.17), we see the sequence of operators $\mathsf{P}^{*,I_1} = Q^{*,I_1}U^{I_1}, \mathsf{P}^{I_1,I_2} = Q^{I_1,I_2}U^{I_2}, \mathsf{P}^{I_2,I_3} = Q^{I_2,I_3}, \mathsf{P}^{I_3,I_4} = Q^{I_3,I_4}, \mathsf{P}^{I_4,I_5} = Q^{I_4,I_5}$, with $|I_1|=3, |I_2|=3, |I_3|=2, |I_4|=1, |I_5|=3$.

from 0.[†] Therefore we can sum the r.h.s. of (6.10) over $N \leqslant n_{r+1} \leqslant 2N$ and then divide by N to get

$$|M^{\varphi,\psi}_{N,\tilde{N},h}| \leqslant \frac{C\|\psi\|_\infty^h}{N^{h+1}} \sum_{r=1}^{\infty} \sum_{\substack{1\leqslant n_1<\cdots<n_r\leqslant n_{r+1}\leqslant 2N \\ I_1,\ldots,I_r\vdash\{1,\ldots,h\}, m_i:=|I_i|<h \\ \boldsymbol{y}_1,\ldots,\boldsymbol{y}_r\in(\mathbb{Z}^2)^h}} Q^N_{n_1}(\varphi,\boldsymbol{y}_1)\mathbb{1}_{\{\boldsymbol{y}_1\sim I_1\}}\mathbb{E}[\xi^{I_1}] \times \prod_{i=2}^{r} Q^{I_{i-1},I_i}_{n_i-n_{i-1}}(\boldsymbol{y}_{i-1},\boldsymbol{y}_i)\mathbb{E}[\xi^{I_i}] \times \mathbb{1}_{\{\boldsymbol{y}_r\sim I_r\}} Q^N_{n_{r+1}-n_r}(\boldsymbol{y}_r,\mathbb{1}_B). \tag{6.12}$$

We first single out consecutive appearances of the same I among $I_1,\ldots,I_r$ with $|I| = h-1$ (that is, I consists of all singletons except for a pair $\{k,l\}$). Given $I \vdash \{1,\ldots,h\}$ with $|I| = h-1$, for $1 \leqslant s_1 < s_2 \leqslant N$ and $\boldsymbol{z}_1, \boldsymbol{z}_2 \in (\mathbb{Z}^2)^h$, define

$$\mathsf{U}^I_{s_2-s_1,\tilde{N}}(\boldsymbol{z}_1,\boldsymbol{z}_2) := \mathbb{1}_{\{\boldsymbol{z}_1,\boldsymbol{z}_2\sim I\}} \sum_{r=1}^{\infty} \mathbb{E}[\xi^2]^r \sum_{\substack{n_0:=s_1<n_1<\cdots<n_r:=s_2 \\ \boldsymbol{y}_i\in(\mathbb{Z}^2)^h, \boldsymbol{y}_0:=\boldsymbol{z}_1, \boldsymbol{y}_r:=\boldsymbol{z}_2}} \prod_{i=1}^{r} Q^{I,I}_{n_i-n_{i-1}}(\boldsymbol{y}_{i-1},\boldsymbol{y}_i) \tag{6.13}$$

(recall that the moments of ξ depend on $\tilde{N}$, see (6.5)) and define $\mathsf{U}^I_{0,\tilde{N}}(\boldsymbol{z}_1,\boldsymbol{z}_2) := \mathbb{1}_{\{\boldsymbol{z}_1=\boldsymbol{z}_2\sim I\}}$. If $\{k,l\}$ is the unique partition element of I with cardinality two, then we can write

$$\mathsf{U}^I_{s_2-s_1,\tilde{N}}(\boldsymbol{z}_1,\boldsymbol{z}_2) = U_{\tilde{N}}(s_2-s_1, \boldsymbol{z}_{2,k}-\boldsymbol{z}_{1,k}) \prod_{i\in\{1,\ldots,h\}\setminus\{k,l\}} q_{s_2-s_1}(\boldsymbol{z}_{2,i}-\boldsymbol{z}_{1,i}), \tag{6.14}$$

[†]For $3N/2 \leqslant n_{r+1} \leqslant 2N$, the inequality (6.11) holds for much more general B than balls of radius $\varrho \geqslant 1$, and Theorem 6.1 can be extended accordingly. But we use balls for simplicity.

where $U_{\tilde N}(n,x)$ was defined in (3.31) (with σ_N^2 therein replaced by $\sigma_{\tilde N}^2$). In (6.12), we can then contract the consecutive appearances of $I_i = I_{i+1} = \cdots I_j = I$ with $|I| = h-1$ into a single kernel $\mathsf{U}_\cdot^I(\cdot,\cdot)$, so that each I_i with $|I_i| = h-1$ does not appear twice in a row in the summation in (6.12).

With a slight overload of notation in order to avoid extra symbols, for $\lambda > 0$ we set

$$\begin{aligned} Q^{I,J}_{\lambda,N}(\boldsymbol{y},\boldsymbol{z}) &:= \sum_{n=1}^{2N} e^{-\lambda n}\, Q^{I,J}_n(\boldsymbol{y},\boldsymbol{z}), \qquad && \boldsymbol{y},\boldsymbol{z} \in (\mathbb{Z}^2)^h, \\ \mathsf{U}^J_{\lambda,N,\tilde N}(\boldsymbol{y},\boldsymbol{z}) &:= \sum_{n=0}^{2N} e^{-\lambda n}\, \mathsf{U}^J_{n,\tilde N}(\boldsymbol{y},\boldsymbol{z}), \qquad && \boldsymbol{y},\boldsymbol{z} \in (\mathbb{Z}^2)^h, \end{aligned} \tag{6.15}$$

for partitions $I, J \vdash \{1,\dots,h\}$, and we finally define, with operator notation,

$$\mathsf{P}^{I,J}_{\lambda,N,\tilde N} := \begin{cases} Q^{I,J}_{\lambda,N} & \text{if } |J| < h-1\,, \\ Q^{I,J}_{\lambda,N}\, \mathsf{U}^J_{\lambda,N,\tilde N} & \text{if } |J| = h-1\,. \end{cases} \tag{6.16}$$

To lighten notation, we will often omit the dependence of these operators on $N, \tilde N$.

The introduction of the parameter λ, especially the choice $\lambda = \hat\lambda/N$ we will take later, will be crucial in decoupling the sum over $n_1, n_2 - n_1, \dots, n_{r+1} - n_r = N - n_r$ in (6.10). Together with the replacement of $n_{r+1} = N$ by averaging n_{r+1} over $[N, 2N]$ in (6.12), this allows us to take Laplace transform and bound the r.h.s. of (6.10) in terms of the operators defined in (6.15)-(6.16). Furthermore, by taking $\hat\lambda$ large, we can extract a logarithmic decay in $\hat\lambda$ from $\mathsf{U}^J_{\lambda,N,\tilde N}$, see (6.24). These ideas were used in [GQT21] in a continuum setting.

To proceed with estimates along these lines, we first obtain an upper bound on (6.12) by inserting the factor $e^{2\lambda N} e^{-\lambda \sum_{i=1}^{r+1}(n_i - n_{i-1})} \geqslant 1$, for $\lambda > 0$ to be determined later, and enlarging the range of summation for each $n_i - n_{i-1}$ to $[1, 2N]$. Denoting by $I = *$ the partition consisting of singletons (namely, $I = \{1\} \sqcup \{2\} \cdots \sqcup \{h\}$), we can then rewrite the sum in (6.12) and obtain the bound

$$\big|M^{\varphi,\psi}_{N,\tilde N,h}\big| \leqslant \frac{C e^{2\lambda N} \|\psi\|_\infty^h}{N^{h+1}} \sum_{r=1}^{\infty} \sum_{I_1,\dots,I_r} \big\langle \varphi_N^{\otimes h},\, \mathsf{P}^{*,I_1}_\lambda\, \mathsf{P}^{I_1,I_2}_\lambda \cdots \mathsf{P}^{I_{r-1},I_r}_\lambda\, Q^{I_r,*}_\lambda\, \mathbb{1}_{B_N}^{\otimes h} \big\rangle \prod_{i=1}^{r} \mathbb{E}[\xi^{I_i}], \tag{6.17}$$

where the sum is over r partitions $I_1, \cdots, I_r \vdash \{1, \cdots, h\}$ such that $|I_i| \leqslant h-1$ for all $1 \leqslant i \leqslant r$ and there is no consecutive $I_{i-1} = I_i$ with $|I_i| = h-1$; we also applied the definition of $q^N_{0,n}(\varphi, z)$ and $q^N_{0,n}(z,\psi)$ from (3.16)–(3.17), we set $B_N := \sqrt{N} B$ and, given $f : \mathbb{Z}^2 \to \mathbb{R}$, for $\boldsymbol{y} \in (\mathbb{Z}^2)^h$ we define $f^{\otimes h}(\boldsymbol{y}) := \prod_{i=1}^h f(\boldsymbol{y}_i)$.

Our bounds will be in terms of the norms of operators acting on the function space $\ell^q((\mathbb{Z}^2)^h)$ for some $q > 1$. To allow for $\psi = \mathbb{1}_{B_N}^{\otimes h}$ in (6.17), it is necessary to introduce spatial weights, which incidentally will also give bounds on the spatial decay if ψ has compact support and we shift its support toward ∞. More precisely, for a function $w : \mathbb{R}^2 \to (0,\infty)$ such that $\log w$ is Lipschitz, we define its discretized version $w_N : \mathbb{Z}^2 \to \mathbb{R}$ by (3.9) and we introduce the weighted operators

$$\begin{aligned} \widehat Q^{I,J}_{\lambda,N}(\boldsymbol{y},\boldsymbol{z}) &:= \frac{w_N^{\otimes h}(\boldsymbol{y})}{w_N^{\otimes h}(\boldsymbol{z})}\, Q^{I,J}_{\lambda,N}(\boldsymbol{y},\boldsymbol{z}), \\ \widehat{\mathsf{U}}^J_{\lambda,N,\tilde N}(\boldsymbol{y},\boldsymbol{z}) &:= \frac{w_N^{\otimes h}(\boldsymbol{y})}{w_N^{\otimes h}(\boldsymbol{z})}\, \mathsf{U}^J_{\lambda,N,\tilde N}(\boldsymbol{y},\boldsymbol{z}), \end{aligned} \tag{6.18}$$

with $\widehat{\mathsf{P}}^{I,J}_{\lambda,N,\tilde{N}}$ defined from $\widehat{Q}$ and $\widehat{U}$ as in (6.16). Given a partition $I \vdash \{1,\dots,h\}$, denote

$$(\mathbb{Z}^2)^h_I := \{\boldsymbol{x} \in (\mathbb{Z}^2)^h : \boldsymbol{x} \sim I\}, \tag{6.19}$$

which is just a copy of $(\mathbb{Z}^2)^{|I|}$ embedded in $(\mathbb{Z}^2)^h$. Due to the delta function constraints in its definition ($\mathbb{1}_{\{\boldsymbol{x}\sim I, \tilde{\boldsymbol{x}}\sim J\}}$ in (6.9)), *we will regard* $\widehat{Q}^{I,J}_{\lambda,N}(\boldsymbol{x},\tilde{\boldsymbol{x}})$ *as an operator mapping from* $\ell^q((\mathbb{Z}^2)^h_J)$ *to* $\ell^q((\mathbb{Z}^2)^h_I)$ *for some* $q>1$, *and similarly for* $\widehat{\mathsf{U}}^J_{\lambda,N,\tilde{N}}$ *and* $\widehat{P}^{I,J}_{\lambda,N,\tilde{N}}$. For $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, by Hölder's inequality, we can then rewrite the bound (6.17) as

$$\begin{aligned}
\big|M^{\varphi,\psi}_{N,\tilde{N},h}\big| &\leqslant \frac{Ce^{2\lambda N}\|\psi\|^h_\infty}{N^{h+1}} \sum_{r=1}^{\infty} \sum_{I_1,\dots,I_r} \langle \frac{\varphi_N^{\otimes h}}{w_N^{\otimes h}}, \widehat{\mathsf{P}}^{*,I_1}_\lambda \, \widehat{\mathsf{P}}^{I_1,I_2}_\lambda \cdots \widehat{\mathsf{P}}^{I_{r-1},I_r}_\lambda \, \widehat{Q}^{I_r,*}_\lambda \, \mathbb{1}^{\otimes h}_{B_N} w_N^{\otimes h} \rangle \prod_{i=1}^r \mathbb{E}[\xi^{I_i}] \\
&\leqslant \frac{Ce^{2\lambda N}\|\psi\|^h_\infty}{N^{h+1}} \sum_{r=1}^{\infty} \sum_{I_1,\dots,I_r} \Big\| \frac{\varphi_N^{\otimes h}}{w_N^{\otimes h}} \Big\|_{\ell^p} \Big\| \widehat{\mathsf{P}}^{*,I_1}_\lambda \Big\|_{\ell^q\to\ell^q} \Big\| \widehat{\mathsf{P}}^{I_1,I_2}_\lambda \Big\|_{\ell^q\to\ell^q} \cdots \\
&\qquad \cdots \Big\| \widehat{\mathsf{P}}^{I_{r-1},I_r}_\lambda \Big\|_{\ell^q\to\ell^q} \Big\| \widehat{Q}^{I_r,*}_\lambda \Big\|_{\ell^q\to\ell^q} \Big\| \mathbb{1}^{\otimes h}_{B_N} w_N^{\otimes h} \Big\|_{\ell^q} \prod_{i=1}^r \mathbb{E}[\xi^{I_i}],
\end{aligned} \tag{6.20}$$

where $\|\cdot\|_{\ell^p(\mathbb{T})}$ is defined in (6.1), and given an operator $\mathsf{A} : \ell^q(\mathbb{T}) \to \ell^q(\mathbb{T}')$, we set

$$\|\mathsf{A}\|_{\ell^q\to\ell^q} := \sup_{g\neq 0} \frac{\|\mathsf{A}\, g\|_{\ell^q(\mathbb{T}')}}{\|g\|_{\ell^q(\mathbb{T})}} = \sup_{\|f\|_{\ell^p(\mathbb{T}')}\leqslant 1,\, \|g\|_{\ell^q(\mathbb{T})}\leqslant 1} \langle f, \mathsf{A}\, g\rangle. \tag{6.21}$$

In our case $\widehat{\mathsf{P}}^{I,J}_\lambda, \widehat{Q}^{I,J}_\lambda : \ell^q((\mathbb{Z}^2)^h_J) \to \ell^q((\mathbb{Z}^2)^h_I)$ (note that for $I=*$ we have $(\mathbb{Z}^2)^h_I = (\mathbb{Z}^2)^h$).

We will choose $\lambda := \hat{\lambda}/N$ with $\hat{\lambda}$ large but fixed so that $e^{\lambda N}$ remains bounded. We will show the following.

Proposition 6.6. *Fix* $p,q>1$ *with* $\frac{1}{p}+\frac{1}{q}=1$, *an integer* $h\geqslant 2$ *and* $\hat{\lambda}>0$. *Then there exists* $c = c_{p,q,h,\hat{\lambda}} < \infty$ *such that, uniformly for partitions* $I, J \vdash \{1,\dots,h\}$ *with* $1\leqslant |I|, |J| \leqslant h-1$ *and* $I\neq J$ *when* $|I|=|J|=h-1$, *for large* $N\leqslant \tilde{N}\in\mathbb{N}$ *and* $\lambda = \frac{\hat{\lambda}}{N}$ *we have*

$$\Big\| \widehat{Q}^{I,J}_{\lambda,N} \Big\|_{\ell^q\to\ell^q} \leqslant c\,; \tag{6.22}$$

$$\Big\| \widehat{Q}^{*,I}_{\lambda,N} \Big\|_{\ell^q\to\ell^q} \leqslant c\, N^{\frac{1}{q}}\,, \qquad \Big\| \widehat{Q}^{I,*}_{\lambda,N} \Big\|_{\ell^q\to\ell^q} \leqslant c\, N^{\frac{1}{p}}\,; \tag{6.23}$$

$$\textit{furthermore, for } |I|=h-1, \qquad \Big\| \widehat{\mathsf{U}}^I_{\lambda,N,\tilde{N}} \Big\|_{\ell^q\to\ell^q} \leqslant \frac{c}{(\log \hat{\lambda}\frac{\tilde{N}}{N})\,\sigma^2_{\tilde{N}}}. \tag{6.24}$$

Recall by (6.5) that $|\mathbb{E}[\xi^{I_i}]| = \sigma^2_{\tilde{N}}$ if $|I_i| = h-1$, while $|\mathbb{E}[\xi^{I_i}]| \leqslant c'\,\sigma^3_{\tilde{N}} = O((\log\tilde{N})^{-\frac{3}{2}})$ if $|I_i|<h-1$. Then, by the definition of $\widehat{\mathsf{P}}$ analogous to (6.16), Proposition 6.6 implies that in (6.20), for each $2\leqslant i\leqslant r$, we have for N sufficiently large

$$\mathbb{E}\big[\xi^{I_i}\big] \Big\| \widehat{\mathsf{P}}^{I_{i-1},I_i}_\lambda \Big\|_{\ell^q\to\ell^q} \leqslant \mathbb{1}_{\{|I_i|=h-1\}} \frac{c^2}{(\log\hat{\lambda}\frac{\tilde{N}}{N})} + \mathbb{1}_{\{|I_i|<h-1\}} c\,c'\sigma^3_{\tilde{N}} \leqslant \frac{c''}{\log\hat{\lambda}\frac{\tilde{N}}{N}}\,,$$

where for $|I_i| = h-1$ with $\widehat{\mathsf{P}}_\lambda^{I_{i-1},I_i} = \widehat{Q}_\lambda^{I_{i-1},I_i}\,\widehat{\mathsf{U}}_\lambda^{I_i}$, we used the fact that $\|\widehat{Q}_\lambda^{I_{i-1},I_i}\,\widehat{\mathsf{U}}_\lambda^{I_i}\|_{\ell^q\to\ell^q} \leqslant \|\widehat{Q}_\lambda^{I_{i-1},I_i}\|_{\ell^q\to\ell^q}\|\widehat{\mathsf{U}}_\lambda^{I_i}\|_{\ell^q\to\ell^q}$. Similarly,

$$\mathbb{E}\big[\xi^{I_1}\big]\Big\|\widehat{\mathsf{P}}_\lambda^{*,I_1}\Big\|_{\ell^q\to\ell^q} \leqslant cN^{\frac{1}{q}}\left(\frac{c\,\mathbb{1}_{\{|I_1|=h-1\}}}{\log\hat\lambda\frac{\tilde N}{N}} + \mathbb{1}_{\{|I_1|<h-1\}}c'\sigma_{\tilde N}^3\right) \leqslant \frac{c''}{\log\hat\lambda\frac{\tilde N}{N}}\,N^{\frac{1}{q}}\,.$$

Substituting these bounds into (6.20), bounding the number of choices for each $I_i \vdash \{1,\dots,h\}$ by a suitable constant $\mathfrak{c}_h$, choosing $\hat\lambda$ large such that $\mathfrak{c}_h\frac{c''}{\log\hat\lambda} \leqslant \frac12$, and using the fact that $|I_1|, |I_r| \leqslant h-1$, we then obtain that, for all N sufficiently large,

$$\begin{aligned}
\big|M_{N,\tilde N,h}^{\varphi,\psi}\big| &\leqslant \frac{Ce^{2\hat\lambda}\|\psi\|_\infty^h}{N^h}\,c\left\|\frac{\varphi_N}{w_N}\right\|_{\ell^p}^h\|w_N\mathbb{1}_{B_N}\|_{\ell^q}^h\sum_{r=1}^\infty\Big(\frac{\mathfrak{c}_h\,c''}{\log\hat\lambda\frac{\tilde N}{N}}\Big)^r \\
&\leqslant \frac{(2\,C\,c\,\mathfrak{c}_h\,c'')}{N^h\,\log\hat\lambda\frac{\tilde N}{N}}\|\psi\|_\infty^h\left\|\frac{\varphi_N}{w_N}\right\|_{\ell^p}^h\|w_N\mathbb{1}_{B_N}\|_{\ell^q}^h \qquad (6.25)\\
&\leqslant \frac{\mathsf{C}}{N^h\log(1+\frac{\tilde N}{N})}\,\|\psi\|_\infty^h\left\|\frac{\varphi_N}{w_N}\right\|_{\ell^p}^h\|w_N\mathbb{1}_{B_N}\|_{\ell^q}^h\,,
\end{aligned}$$

where the last inequality holds for $\hat\lambda \geqslant 2$. This concludes the proof of Theorem 6.1. □

6.2. Functional Inequalities. It only remains to prove the bounds in Proposition 6.6. The key ingredient is a Hardy-Littlewood-Sobolev type inequality. First we need the following bound on the Green's function of a random walk on $(\mathbb{Z}^2)^h$.

Lemma 6.7. *Given $N\in\mathbb{N}$, $\lambda\geqslant 0$, an integer $h\geqslant 2$ and $\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h$, denote $Q_{\lambda,N}(\boldsymbol{x},\boldsymbol{y}) := \sum_{n=1}^{2N}e^{-\lambda n}\prod_{i=1}^h q_n(y_i-x_i)$. Then for some $C\in(0,\infty)$ uniformly in λ, N and $\boldsymbol{x},\boldsymbol{y}$,*

$$Q_{\lambda,N}(\boldsymbol{x},\boldsymbol{y}) \leqslant \begin{cases} \dfrac{C}{(1+|\boldsymbol{y}-\boldsymbol{x}|^2)^{h-1}} & \text{for all } \boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h,\\[2ex] \dfrac{C}{N^{h-1}}e^{-\frac{|\boldsymbol{y}-\boldsymbol{x}|^2}{CN}} & \text{for } |\boldsymbol{x}-\boldsymbol{y}|>\sqrt{N}.\end{cases} \qquad (6.26)$$

Proof. We may assume $\lambda = 0$. Note that $\prod_{i=1}^h q_n(\cdot)$ is the transition kernel of a random walk on $\mathbb{Z}^{2h}$. By the local central limit theorem [LaLi10, Theorem 2.3.11] and a Gaussian concentration bound, we have

$$\prod_{i=1}^h q_n(y_i-x_i) \leqslant \frac{C_1}{n^h}e^{-C_2\frac{|\boldsymbol{y}-\boldsymbol{x}|^2}{n}}$$

for some $C_1, C_2\in(0,\infty)$ uniformly in $n\in\mathbb{N}$ and $\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h$. We then have

$$Q_\lambda(\boldsymbol{x},\boldsymbol{y}) \leqslant C_1\sum_{n=1}^{2N}n^{-h}e^{-C_2\frac{|\boldsymbol{y}-\boldsymbol{x}|^2}{n}} \leqslant C_1\min\left\{2,\ \frac{1}{N^{h-1}}\int_0^2 t^{-h}e^{-C_2\frac{|\boldsymbol{z}_N|^2}{t}}\mathrm{d}t\right\},$$

where we used a Riemann sum approximation and we set $\boldsymbol{z}_N := (\boldsymbol{y}-\boldsymbol{x})/\sqrt{N}$. When $\boldsymbol{z}_N = 0$, we just use the constant upper bound $Q_{\lambda,N}(\boldsymbol{x},\boldsymbol{y}) \leqslant 2C_1$. When $\boldsymbol{z}_N\neq 0$, we write

$$Q_{\lambda,N}(\boldsymbol{x},\boldsymbol{y}) \leqslant \frac{C_1}{N^{h-1}}\int_0^2 t^{-h}e^{-C_2\frac{|\boldsymbol{z}_N|^2}{t}}\mathrm{d}t = \frac{C_1}{(C_2N|\boldsymbol{z}_N|^2)^{h-1}}\int_{\frac12 C_2|\boldsymbol{z}_N|^2}^\infty\tau^{h-2}e^{-\tau}\mathrm{d}\tau,$$

where $N|\boldsymbol{z}_N|^2 = |\boldsymbol{y}-\boldsymbol{x}|^2$, while the integral is bounded uniformly in $\boldsymbol{z}_N$ and can be bounded by $C_3 e^{-C_4|\boldsymbol{z}_N|^2}$ when $|\boldsymbol{z}_N| \geqslant 1$. The bound (6.26) then follows. □

The following crucial lemma proves a Hardy-Littlewood-Sobolev type inequality. This generalizes an inequality of Dell'Antonio-Figari-Teta in [DFT94] (see Lemma 3.1 and inequalities (3.1) and (3.5) therein) which played an important role in [GQT21] for moment bounds with L^2 test functions and initial conditions.

Lemma 6.8. *Fix $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ and an integer $h \geqslant 2$. Consider partitions $I, J \vdash \{1, \dots, h\}$ with $1 \leqslant |I|, |J| \leqslant h-1$, and $I \neq J$ if $|I| = |J| = h-1$. Recall $(\mathbb{Z}^2)^h_I$ from (6.19) and the associated function space $\ell^p\big((\mathbb{Z}^2)^h_I\big)$. Let $f \in \ell^p\big((\mathbb{Z}^2)^h_I\big)$ and $g \in \ell^q\big((\mathbb{Z}^2)^h_J\big)$. Then there exists $C = C_{p,q,h} < \infty$, independent of f and g, such that*

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I\,,\,\boldsymbol{y}\in(\mathbb{Z}^2)^h_J} \frac{f(\boldsymbol{x})g(\boldsymbol{y})}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}} \leqslant C\,\|f\|_{\ell^p}\|g\|_{\ell^q}. \tag{6.27}$$

In [DFT94], an analogue of (6.27) was proved in the continuum for the special case $p = q = 2$ (i.e., L^2 test functions) and $|I| = |J| = h-1$ with $I \neq J$. They presented their inequality in Fourier space, but in the L^2 case, it is equivalent to (6.27) by the Plancherel theorem. Here we work on the lattice, which requires us to also consider partitions with $|I| < h-1$ or $|J| < h-1$, which cases are not present in the continuum. We also consider test functions in general ℓ^p-ℓ^q spaces (our proof steps can also be carried out in the continuum to extend the inequality of [DFT94] to L^p-L^q spaces). Instead of working in Fourier space as in [DFT94], we will work directly in real space, which allows the extensions mentioned above.

Remark 6.9. *The inequality (6.27) is not expected to hold for $|I| = |J| = h-1$ with $I = J$, because it is exactly the borderline case of the Hardy-Littlewood-Sobolev inequality when it fails, see* [LiLo01, Theorem 4.3].

Proof of Lemma 6.8. We first consider the case $|I| = |J| = h-1$, and $I \neq J$. Then I and J each contains a partition element with cardinality 2, say $\{k, l\}$ and $\{m, n\}$ respectively, and $\{k, l\} \neq \{m, n\}$. In particular, $x_k = x_l$ and $y_m = y_n$ for $\boldsymbol{x} \in (\mathbb{Z}^2)^h_I, \boldsymbol{y} \in (\mathbb{Z}^2)^h_J$.

Fix any $0 < a < \frac{1}{p\vee q}$. We then apply Hölder to bound the left hand side of (6.27) by

$$\begin{aligned}\Bigg(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I,\boldsymbol{y}\in(\mathbb{Z}^2)^h_J} \frac{f(\boldsymbol{x})^p}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}} \cdot \frac{(1+|x_m-x_n|^{2a})^p}{(1+|y_k-y_l|^{2a})^p}\Bigg)^{1/p} \\ \times \Bigg(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I,\boldsymbol{y}\in(\mathbb{Z}^2)^h_J} \frac{g(\boldsymbol{y})^q}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}} \cdot \frac{(1+|y_k-y_l|^{2a})^q}{(1+|x_m-x_n|^{2a})^q}\Bigg)^{1/q}.\end{aligned} \tag{6.28}$$

We now bound the first factor in (6.28). Note that since $y_m = y_n$, by the triangle inequality,

$$|x_m-y_m|^2 + |x_n-y_n|^2 \geqslant \frac{|x_m-x_n|^2 + |x_n-y_n|^2}{3}. \tag{6.29}$$

Substituting this inequality into $\sum_{i=1}^h |x_i - y_i|^2$ then bounds the first factor in (6.28) by

$$C\left(\sum_{\boldsymbol{x}} f(\boldsymbol{x})^p (1+|x_m-x_n|^{2a})^p \sum_{\boldsymbol{y}} \frac{1}{\left(1+|x_m-x_n|^2+\sum_{i\neq m}|x_i-y_i|^2\right)^{h-1}(1+|y_k-y_l|^{2a})^p}\right)^{1/p}. \tag{6.30}$$

Note that for $r>1$ we can bound $\sum_{y\in\mathbb{Z}^2}\frac{1}{(s+|y-x|^2)^r}\leqslant\frac{C}{s^{r-1}}$ uniformly for $s\geqslant 0$ and $x\in\mathbb{Z}^2$. We can then successively sum over the variables y_j with $j\neq k,l$: there are $|J|-2$ such variables y_j and they are all present in the sum $\sum_{i\neq m}|y_i-x_i|^2$ (recall that we sum over $\boldsymbol{y}\in(\mathbb{Z}^2)^h_J$, hence $y_m=y_n$ is a single variable), hence we get

$$\begin{aligned}&\sum_{\boldsymbol{y}\in(\mathbb{Z}^2)^h_J}\frac{1}{\left(1+|x_m-x_n|^2+\sum_{i\neq m}|y_i-x_i|^2\right)^{h-1}(1+|y_k-y_l|^{2a})^p}\\ &\leqslant C\sum_{y_k,y_l}\frac{1}{\left(1+|x_m-x_n|^2+|y_k-x_k|^2+|y_l-x_l|^2\right)^{h+1-|J|}(1+|y_k-y_l|^{2a})^p}.\end{aligned}$$

Note that $x_k=x_l$. Via the change of variables $\tilde{y}_1=y_k-y_l$ and $\tilde{y}_2=y_k+y_l-2x_k$, and the observation that $|y_k-x_k|^2+|y_l-x_l|^2=(\tilde{y}_1^2+\tilde{y}_2^2)/2$, we can bound the above sum by

$$\begin{aligned}&C\sum_{\tilde{y}_1,\tilde{y}_2}\frac{1}{\left(1+|x_m-x_n|^2+\tilde{y}_1^2+\tilde{y}_2^2\right)^{h+1-|J|}(1+|\tilde{y}_1|^{2a})^p}\\ &\leqslant C\sum_{\tilde{y}_1}\frac{1}{\left(1+|x_m-x_n|^2+\tilde{y}_1^2\right)^{h-|J|}(1+|\tilde{y}_1|^{2a})^p}\leqslant\frac{C}{1+|x_n-x_m|^{2(h-1-|J|)+2ap}},\end{aligned}$$

where the last inequality is obtained by summing separately over $|\tilde{y}_1|\leqslant|x_n-x_m|$ and $|\tilde{y}_1|>|x_n-x_m|$, plus the assumption that $ap<1$. Substituting this bound into (6.30), since we assume $|J|=h-1$, we obtain that the first factor in (6.28) is bounded by $C\|f\|_{\ell^p}$. The second factor in (6.28) can similarly be bounded by $C\|g\|_{\ell^q}$. This concludes the proof of (6.27), and hence also (6.22), for the case $|I|=|J|=h-1$ and $I\neq J$.

We can adapt the proof to the case $\min\{|I|,|J|\}<h-1$ as follows. If $|I|,|J|<h-1$, then there is no need to introduce the factor $\frac{1+|x_m-x_n|^{2a}}{1+|y_k-y_l|^{2a}}$ and its reciprocal in (6.28) because we already have $\sum_{\boldsymbol{y}\in(\mathbb{Z}^2)^h_J}\frac{1}{(1+\sum_{i=1}^h|x_i-y_i|^2)^{h-1}}<\infty$. If $|I|<h-1$ and $|J|=h-1$, then we can still find k,l in the same partition element of I, but not the same partition element of J. We should then replace the factor $\frac{1+|x_m-x_n|^{2a}}{1+|y_k-y_l|^{2a}}$ and its reciprocal in (6.28) by $\frac{1}{1+|y_k-y_l|^{2a}}$. The rest of the proof is essentially the same. □

6.3. Proof of Proposition 6.6. We now prove (6.22)–(6.24).

Proof of (6.22)**.** Note that (6.22) is equivalent to showing (recall (6.18))

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I,\boldsymbol{y}\in(\mathbb{Z}^2)^h_J} f(\boldsymbol{x})Q_\lambda^{I,J}(\boldsymbol{x},\boldsymbol{y})\frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})}\,g(\boldsymbol{y})\leqslant c\,\|f\|_{\ell^p}\,\|g\|_{\ell^q} \tag{6.31}$$

uniformly for all $f\in\ell^p((\mathbb{Z}^2)^h_I)$ and $g\in\ell^q((\mathbb{Z}^2)^h_J)$. To control the effect of the weight $w_N^{\otimes h}$, we split the summation into the regions

$$A_N:=\big\{(\boldsymbol{x},\boldsymbol{y})\colon|\boldsymbol{x}-\boldsymbol{y}|\leqslant C_0\sqrt{N}\big\}$$

and A_N^c for some C_0 to be chosen later. Note that $\log w(\frac{y}{\sqrt{N}}) - \log w(\frac{x}{\sqrt{N}}) = O(\frac{|x-y|}{\sqrt{N}})$, because $\log w$ is assumed to be Lipschitz. Since $w_N : \mathbb{Z}^2 \to \mathbb{R}$ is obtained from $w : \mathbb{R}^2 \to \mathbb{R}$ by (3.9), we have for all $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{Z}^2)^h$

$$\frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})} \leqslant e^{C|\boldsymbol{x}-\boldsymbol{y}|/\sqrt{N}}, \tag{6.32}$$

which is bounded by e^{CC_0} in A_N. Therefore, the contribution of this region to the l.h.s. of (6.31) is controlled by the following uniform bound, that we prove below:

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)_I^h, \boldsymbol{y}\in(\mathbb{Z}^2)_J^h} f(\boldsymbol{x})\, Q_{\lambda,N}^{I,J}(\boldsymbol{x},\boldsymbol{y})\, g(\boldsymbol{y}) \leqslant c_1 \|f\|_{\ell^p} \|g\|_{\ell^q}. \tag{6.33}$$

In the region A_N^c, since $Q_{\lambda,N}^{I,J} \leqslant Q_{\lambda,N}$ (recall (6.8)-(6.9)), we can apply Lemma 6.7 to bound

$$\frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})} Q_{\lambda,N}(\boldsymbol{x},\boldsymbol{y}) \leqslant \frac{C}{N^{h-1}} \exp\Big\{ -\tfrac{|\boldsymbol{x}-\boldsymbol{y}|^2}{CN} + \tfrac{C|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt{N}} \Big\} \leqslant \frac{C}{N^{h-1}} \exp\Big\{ -\tfrac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt{N}} \Big\}, \tag{6.34}$$

where the last inequality holds for $|\boldsymbol{x}-\boldsymbol{y}| > C_0\sqrt{N}$ with $C_0 := C(C+1)$. Thus

$$\begin{aligned}
&\sum_{(\boldsymbol{x},\boldsymbol{y})\in A_N^c} f(\boldsymbol{x}) \cdot \frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})}\, Q_{\lambda,N}^{I,J}(\boldsymbol{x},\boldsymbol{y}) \cdot g(\boldsymbol{y}) \\
\leqslant\ & \frac{C}{N^{h-1}} \sum_{\boldsymbol{x},\boldsymbol{y}} f(\boldsymbol{x})\, e^{-\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt{N}}}\, g(\boldsymbol{y}) \\
\leqslant\ & \frac{C}{N^{h-1}} \Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)_I^h, \boldsymbol{y}\in(\mathbb{Z}^2)_J^h} |f(\boldsymbol{x})|^p\, e^{-\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt{N}}} \Big)^{1/p} \Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)_I^h, \boldsymbol{y}\in(\mathbb{Z}^2)_J^h} |g(\boldsymbol{y})|^q\, e^{-\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt{N}}} \Big)^{1/q} \\
\leqslant\ & CN^{\frac{|J|}{p}+\frac{|I|}{q}-(h-1)}\, \|f\|_{\ell^p}\, \|g\|_{\ell^q},
\end{aligned} \tag{6.35}$$

where the prefactor is bounded if $|I|, |J| \leqslant h-1$. Combined with (6.33), this implies (6.31). It only remains to prove (6.33), which follows from Lemmas 6.7 and 6.8 above. □

Proof of (6.23)**.** It suffices to show that for $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ and for $|I| \leqslant h-1$

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)_I^h, \boldsymbol{y}\in(\mathbb{Z}^2)^h} f(\boldsymbol{x})\, Q_{\lambda,N}^{I,*}(\boldsymbol{x},\boldsymbol{y}) \frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})}\, g(\boldsymbol{y}) \leqslant cN^{\frac{1}{p}} \|f\|_{\ell^p} \|g\|_{\ell^q} \tag{6.36}$$

uniformly in $f \in \ell^p((\mathbb{Z}^2)_I^h)$ and $g \in \ell^q((\mathbb{Z}^2)^h)$, which proves the second relation in (6.23); the first relation follows by interchanging f and g. (We recall that $J = *$ denotes the partition of $\{1, \dots, h\}$ consisting of h singletons, i.e. $J = \{1\}, \{2\}, \dots, \{h\}$.)

The proof is similar to that of (6.31). When the sum in (6.36) is restricted to A_N^c with $A_N := \big\{(\boldsymbol{x},\boldsymbol{y}) \colon |\boldsymbol{x}-\boldsymbol{y}| \leqslant C_0\sqrt{N}\big\}$, the same bound in (6.34)-(6.35) holds, which gives an upper bound of

$$CN^{\frac{h}{p}+\frac{|I|}{q}-(h-1)} \|f\|_{\ell^p} \|g\|_{\ell^q} = CN^{1-\frac{h-|I|}{q}} \|f\|_{\ell^p} \|g\|_{\ell^q} \leqslant CN^{\frac{1}{p}} \|f\|_{\ell^p} \|g\|_{\ell^q}\,. \tag{6.37}$$

It only remains to bound the sum in (6.36) restricted to A_N and show the following analogue of (6.27):

$$\sum_{\substack{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I\,,\,\boldsymbol{y}\in(\mathbb{Z}^2)^h\\ |\boldsymbol{x}-\boldsymbol{y}|\leqslant C_0\sqrt{N}}} \frac{f(\boldsymbol{x})g(\boldsymbol{y})}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}} \leqslant CN^{\frac{1}{p}}\|f\|_{\ell^p}\|g\|_{\ell^q}\,. \tag{6.38}$$

W.l.o.g., we may assume that 1 and 2 belong to the same partition element of I, so that $x_1 = x_2$. By Hölder's inequality, we can bound the l.h.s. by

$$\begin{aligned}&\Bigg(\sum_{\substack{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I\,,\,\boldsymbol{y}\in(\mathbb{Z}^2)^h\\ |\boldsymbol{x}-\boldsymbol{y}|\leqslant C_0\sqrt{N}}} \frac{f^p(\boldsymbol{x})\big(\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)\big)^{\frac{p}{q}}}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}}\Bigg)^{\frac{1}{p}}\\ &\times\Bigg(\sum_{\substack{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I\,,\,\boldsymbol{y}\in(\mathbb{Z}^2)^h\\ |\boldsymbol{x}-\boldsymbol{y}|\leqslant C_0\sqrt{N}}} \frac{g(\boldsymbol{y})^q}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)}\Bigg)^{\frac{1}{q}}.\end{aligned} \tag{6.39}$$

In the second factor, since $x_1 = x_2$, we can bound $|x_2-y_2|^2 + |x_1-y_1|^2 \geqslant \frac{|y_1-y_2|^2+|x_1-y_1|^2}{3}$ as in (6.29) to replace $|x_2-y_2|^2$ by $|y_1-y_2|^2$ inside $\sum_{i=1}^h |x_i-y_i|^2$. By the same argument as that following (6.30), we can sum out the variables x_i for $i \geqslant 3$. Since there are $|I|-1$ such variables in $(\mathbb{Z}^2)^h_I$, for $|I| \leqslant h-1$ we get

$$\begin{aligned}&\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^h_I} \frac{1}{\big(1+|y_1-y_2|^2+\sum_{i\neq 2}|x_i-y_i|^2\big)^{h-1}\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)}\\ &\leqslant \frac{1}{\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)} \sum_{\substack{x_1\in\mathbb{Z}^2\\ |x_1-y_1|\leqslant C_0\sqrt{N}}} \frac{1}{\big(1+|y_1-y_2|^2+|x_1-y_1|^2\big)^{h-|I|}} \leqslant C\,,\end{aligned} \tag{6.40}$$

where the last bound holds because $\sum_{x\in\mathbb{Z}^2:\,|x|\leqslant k} \frac{1}{s+|x|^2} \leqslant C\log(1+\frac{k}{\sqrt{s}})$ uniformly in $k, s \geqslant 1$, and furthermore $|y_1-y_2| \leqslant |y_1-x_1| + |x_2-y_2| \leqslant 2C_0\sqrt{N}$ by $|\boldsymbol{x}-\boldsymbol{y}| \leqslant C_0\sqrt{N}$. This implies that the second factor in (6.39) can be bounded by $C\|g\|_{\ell^q}$.

For the first factor in (6.39), we can first sum over $\boldsymbol{y} \in (\mathbb{Z}^2)^h$ to bound

$$\sum_{\boldsymbol{y}\in(\mathbb{Z}^2)^h} \frac{\big(\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)\big)^{\frac{p}{q}}}{\big(1+\sum_{i=1}^h |x_i-y_i|^2\big)^{h-1}} \leqslant C \sum_{\substack{y_1,y_2\in\mathbb{Z}^2\\ |y_1-x_1|,|y_2-x_2|\leqslant C_0\sqrt{N}}} \frac{\big(\log\big(1+\frac{C_0^2N}{1+|y_1-y_2|^2}\big)\big)^{\frac{p}{q}}}{1+|y_1-x_1|^2+|y_2-x_2|^2}\,.$$

Recall $x_1 = x_2$. Let $z_1 := y_1 - y_2$ and $z_2 := y_1 + y_2 - 2x_1$, so that $|z_1|, |z_2| \leqslant 2C_0\sqrt{N}$ and $|z_1|^2 + |z_2|^2 = 2(|y_1-x_1|^2 + |y_2-x_2|^2)$. Summing over z_2 then leads to the bound

$$\sum_{|z_1|\leqslant 2C_0\sqrt{N}} \Big(\log\big(1+\frac{C_0^2N}{1+|z_1|^2}\big)\Big)^{1+\frac{p}{q}} \leqslant CN$$

by a Riemann sum approximation. Therefore the first factor in (6.39) can be bounded by $CN^{\frac{1}{p}}\|f\|_{\ell^p}$. Together with (6.40), this implies (6.38) and concludes the proof of (6.23). □

Proof of (6.24)**.** Note that (6.24) is equivalent to showing

$$\sum_{\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h_I} f(\boldsymbol{x})\,\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N,\tilde N}(\boldsymbol{x},\boldsymbol{y})\,g(\boldsymbol{y})\,\frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})} \leqslant \frac{c\log\tilde N}{\log\hat\lambda\frac{\tilde N}{N}}\|f\|_{\ell^p}\|g\|_{\ell^q} \tag{6.41}$$

uniformly in $f\in\ell^p((\mathbb{Z}^2)^h_I)$ and $g\in\ell^q((\mathbb{Z}^2)^h_I)$. Without loss of generality, we may assume $I\vdash\{1,\dots,h\}$ consists of partition elements $\{1,2\},\{3\},\dots,\{h\}$, so that $x_1=x_2$ and $y_1=y_2$.

Recall from (6.15) and (6.14) that

$$\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N,\tilde N}(\boldsymbol{x},\boldsymbol{y}) = \mathbb{1}_{\{\boldsymbol{x}=\boldsymbol{y}\}} + \sum_{n=1}^{2N} e^{-\hat\lambda\frac{n}{N}}U_{\tilde N}(n,y_1-x_1)\prod_{i=3}^h q_n(y_i-x_i), \tag{6.42}$$

where $U_{\tilde N}(n,x)$ is defined in (3.31), with σ_N^2 replaced by $\sigma_{\tilde N}^2$. Let us set $T:=\frac{\tilde N}{N}\geqslant 1$ for short. By (3.35), (3.43) where $\overline{U}_{\tilde N}=\sigma_{\tilde N}^2U_{\tilde N}$, and (3.46), we have

$$\begin{aligned}\sum_{\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N,\tilde N}(\boldsymbol{x},\boldsymbol{y}) &\leqslant 1+\sum_{n=1}^{2N}e^{-\hat\lambda\frac{n}{N}}U_{\tilde N}(n)\\ &\leqslant 1+C\,\frac{\log\tilde N}{\tilde N}\sum_{n=1}^{2\tilde N/T}e^{-\hat\lambda\frac{nT}{\tilde N}}G_\vartheta\big(\tfrac{n}{\tilde N}\big)\leqslant 1+C\log\tilde N\int_0^{\frac{2}{T}}e^{-\hat\lambda Tt}G_\vartheta(t)\mathrm{d}t\\ &\leqslant 1+Ce^{-\frac{\hat\lambda T}{2}}\log\tilde N+C\log\tilde N\int_0^{\frac12\wedge\frac{2}{T}}\frac{e^{-\hat\lambda Tt}}{t(\log\frac1t)^2}\mathrm{d}t\leqslant C\frac{\log\tilde N}{\log\hat\lambda T},\end{aligned} \tag{6.43}$$

where the last inequality follows by bounding the integral separately over $\big(0,\frac{1}{(\hat\lambda T)^{1/2}}\big)$ and $\big(\frac{1}{(\hat\lambda T)^{1/2}},\frac12\wedge\frac{2}{T}\big)$, with the dominant contribution coming from the first interval.

On the other hand, for any $C>0$, by (6.42) we have

$$\begin{aligned}&\sum_{\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N,\tilde N}(\boldsymbol{x},\boldsymbol{y})e^{C\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt N}}\\ &\leqslant 1+\sum_{n=1}^{2N}e^{-\hat\lambda\frac{n}{N}}\sum_{y_1\in\mathbb{Z}^2}U_{\tilde N}(n,y_1-x_1)e^{C\frac{|y_1-x_1|}{\sqrt N}}\prod_{i=3}^h\Big(\sum_{y_i\in\mathbb{Z}^2}q_n(y_i-x_i)e^{C\frac{|x_i-y_i|}{\sqrt N}}\Big)\\ &\leqslant 1+C\sum_{n=1}^{2N}e^{-\hat\lambda\frac{n}{N}}U_{\tilde N}(n)\leqslant C\frac{\log\tilde N}{\log\hat\lambda T},\end{aligned} \tag{6.44}$$

where we applied (3.48) and (6.43).

We can now bound the l.h.s. of (6.41) as follows, recalling (6.32):

$$\begin{aligned}&\sum_{\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}f(\boldsymbol{x})\,\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N}(\boldsymbol{x},\boldsymbol{y})\,g(\boldsymbol{y})\frac{w_N^{\otimes h}(\boldsymbol{x})}{w_N^{\otimes h}(\boldsymbol{y})}\leqslant\sum_{\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}f(\boldsymbol{x})\,\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N}(\boldsymbol{x},\boldsymbol{y})e^{C\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt N}}\,g(\boldsymbol{y})\\ &\leqslant C\Big(\sum_{\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}|f(\boldsymbol{x})|^p\,\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N}(\boldsymbol{x},\boldsymbol{y})e^{C\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt N}}\Big)^{1/p}\Big(\sum_{\boldsymbol{x},\boldsymbol{y}\in(\mathbb{Z}^2)^h_I}|g(\boldsymbol{y})|^q\,\mathsf{U}^I_{\hat\lambda_{\frac{\tilde N}{N}},N}(\boldsymbol{x},\boldsymbol{y})e^{C\frac{|\boldsymbol{x}-\boldsymbol{y}|}{\sqrt N}}\Big)^{1/q}.\\ &\leqslant C\frac{\log\tilde N}{\log\hat\lambda T}\|f\|_{\ell^p}\|g\|_{\ell^q},\end{aligned}$$

where we applied (6.44). Recalling that $T = \frac{\tilde{N}}{N}$, this concludes the proof of (6.41). □

7. Moment estimates for coarse-grained disorder

In this section, we derive second and fourth moment estimates for the coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ defined in (4.11). These will be used later to bound moments of the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon}(\varphi,\psi|\Theta)$ introduced in Definition 4.4, with $\Theta = \Theta^{(\mathrm{cg})}_{N,\varepsilon}$.

7.1. Second moment estimates.

We first study the second moment of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$.

Lemma 7.1. *For each time-space block $(\vec{\mathsf{i}},\vec{\mathsf{a}}) = ((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}')) \in \mathbb{T}_\varepsilon$ defined in (4.5), the coarse-grained disorder variable $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ as defined in (4.11) has mean 0 and its second moment converges to a finite limit*

$$\sigma^2_\varepsilon(\vec{\mathsf{i}},\vec{\mathsf{a}}) := \lim_{N\to\infty} \mathrm{E}\Big[\big(\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})\big)^2\Big], \tag{7.1}$$

see (7.4)-(7.5) below. Furthermore, there exist $c, C > 0$ independent of ε, $\vec{\mathsf{i}}$, $\vec{\mathsf{a}}$ such that

$$\sigma^2_\varepsilon(\vec{\mathsf{i}},\vec{\mathsf{a}}) \leqslant \frac{C e^{-c|\vec{\mathsf{a}}|^2/|\vec{\mathsf{i}}|}\mathbb{1}_{\{|\vec{\mathsf{a}}|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}\}}}{\big(\log\frac{1}{\varepsilon}\big)^{1+\mathbb{1}_{\{|\vec{\mathsf{a}}|>0\}\cup\{|\vec{\mathsf{i}}|\geqslant 2\}}}\big|\vec{\mathsf{i}}\big|^2}. \tag{7.2}$$

Proof. (I) Random walk representation. We first express $\mathrm{E}\big[\big(\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})\big)^2\big]$ in terms of the time-space renewal $(\tau^{(N)}_\cdot, S^{(N)}_\cdot)$ defined in (3.32). First consider the case $|\vec{\mathsf{i}}| = 1$, i.e., $\mathsf{i} = \mathsf{i}'$. Recall from (4.11), (4.9) and (4.10) that

$$\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}}) = \frac{2}{\varepsilon N}\Bigg\{ \sum_{\substack{d\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x\in\mathcal{S}_{\varepsilon N}(\mathsf{a})\cap\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}}\xi_N(d,x) + \sum_{\substack{d<f\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x\in\mathcal{S}_{\varepsilon N}(\mathsf{a}),\, y\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}} q_{d,f}(x,y)\xi_N(d,x)\xi_N(f,y)$$
$$+ \sum_{r=1}^{\infty} \sum_{\substack{d:=n_0<f:=n_{r+1}\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x:=z_0\in\mathcal{S}_{\varepsilon N}(\mathsf{a}),\, y:=z_{r+1}\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}} \sum_{\substack{d<n_1<\ldots<n_r<f\\ z_1,\ldots,z_r\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}})}} \xi_N(d,x)\prod_{j=1}^{r+1} q_{n_{j-1},n_j}(z_{j-1},z_j)\,\xi_N(n_j,z_j)\Bigg\},$$

where we note that the terms are uncorrelated because $\xi_N(\cdot,\cdot)$ are independent centred random variables with mean 0 and variance σ^2_N (recall (3.14) and (3.11)). Therefore

$$\mathrm{E}\Big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})^2\Big] = \frac{4}{(\varepsilon N)^2}\Bigg\{ \sum_{\substack{d\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x\in\mathcal{S}_{\varepsilon N}(\mathsf{a})\cap\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}}\sigma^2_N + \sum_{\substack{d<f\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x\in\mathcal{S}_{\varepsilon N}(\mathsf{a}),\, y\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}}\sigma^4_N q^2_{d,f}(x,y)$$
$$+ \sum_{r=1}^{\infty}\sigma^{2(r+1)}_N \sum_{\substack{d:=n_0<f:=n_{r+1}\in\mathcal{T}_{\varepsilon N}(\mathsf{i})\\ x:=z_0\in\mathcal{S}_{\varepsilon N}(\mathsf{a}),\, y:=z_{r+1}\in\mathcal{S}_{\varepsilon N}(\mathsf{a}')}} \mathbb{1}_{(d,x)\in\mathbb{Z}^3_{\mathrm{even}}} \sum_{\substack{d<n_1<\ldots<n_r<f\\ z_1,\ldots,z_r\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_{\varepsilon N}(\tilde{\mathsf{a}})}} \prod_{j=1}^{r+1} q^2_{n_{j-1},n_j}(z_{j-1},z_j)\Bigg\}.$$

Note that this sum admits a representation in terms of the space-time random walk $(\tau_k^{(N)}, S_k^{(N)})_{k\geqslant 0}$ defined in (3.32), namely,

$$\begin{aligned}\mathrm{E}\Big[\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})^2\Big] &= 2\sigma_N^2\sum_{k=0}^{\infty}(\sigma_N^2 R_n)^k \\ &\times \mathrm{P}_{\mathsf{i},\mathsf{a}}^{N,\varepsilon}\left(\frac{S_i^{(N)}}{\sqrt{N}}\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_\varepsilon(\tilde{\mathsf{a}})\ \forall\, 1\leqslant i<k;\ \left(\frac{\tau_k^{(N)}}{N},\frac{S_k^{(N)}}{\sqrt{N}}\right)\in\mathcal{T}_\varepsilon(\mathsf{i})\times\mathcal{S}_\varepsilon(\mathsf{a}')\right),\end{aligned} \tag{7.3}$$

where $\mathrm{P}_{\mathsf{i},\mathsf{a}}^{N,\varepsilon}$ denotes probability for $(\tau_k^{(N)}, S_k^{(N)})_{k\geqslant 0}$ with $(\tau_0^{(N)}, S_0^{(N)})$ sampled uniformly from $\mathcal{T}_{\varepsilon N}(\mathsf{i})\times\mathcal{S}_{\varepsilon N}(\mathsf{a})\cap\mathbb{Z}^3_{\mathrm{even}}$. Changing variable $k = s\log N$, using $\sigma_N^2 R_N = 1+(\vartheta+o(1))/\log N$, and applying Lemma 3.3 on the convergence of $(\tau_k^{(N)}, S_k^{(N)})_{k\geqslant 0}$ to a Lévy process $\boldsymbol{Y}_\cdot = (Y_\cdot, V_\cdot)$, we find that the sum above converges to the Riemann integral

$$\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) = 2\pi\int_0^\infty e^{\vartheta s}\mathrm{P}_{\mathsf{i},\mathsf{a}}^{\varepsilon}\Big(V_u\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_\varepsilon(\tilde{\mathsf{a}})\ \forall\, u\in(0,s);\ (Y_s,V_s)\in\mathcal{T}_\varepsilon(\mathsf{i})\times\mathcal{S}_\varepsilon(\mathsf{a}')\Big)\,\mathrm{d}s, \tag{7.4}$$

where $\mathrm{P}_{\mathsf{i},\mathsf{a}}^{\varepsilon}$ denotes the law of the Lévy process $\boldsymbol{Y}_u = (Y_u, V_u)$ with $\boldsymbol{Y}_0$ sampled uniformly from $\mathcal{T}_\varepsilon(\mathsf{i})\times\mathcal{S}_\varepsilon(\mathsf{a})$.

For $|\vec{\mathsf{i}}|\geqslant 2$, $\Theta_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ is defined in (4.11). The same argument as for the case $|\vec{\mathsf{i}}| = 1$ gives

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) = &\sum_{\substack{\mathsf{b}:\,|\mathsf{b}-\mathsf{a}|\leqslant M_\varepsilon\\ \mathsf{b}':\,|\mathsf{b}'-\mathsf{a}'|\leqslant M_\varepsilon,\,|\mathsf{b}'-\mathsf{b}|\leqslant M_\varepsilon\sqrt{|\mathsf{i}'-\mathsf{i}|}}}2\pi\int_0^\infty \mathrm{d}s\, e^{\vartheta s}\\ &\times\mathrm{P}_{\mathsf{i},\mathsf{a}}^{\varepsilon}\Big(\exists\, t\in(0,s)\ s.t.\ (Y_{t^-},V_{t^-})\in\mathcal{T}_\varepsilon(\mathsf{i})\times\mathcal{S}_\varepsilon(\mathsf{b}),\ (Y_t,V_t)\in\mathcal{T}_\varepsilon(\mathsf{i}')\times\mathcal{S}_\varepsilon(\mathsf{b}');\\ &\forall\, u\in(0,t):\, V_u\in\bigcup_{|\tilde{\mathsf{a}}-\mathsf{a}|\leqslant M_\varepsilon}\mathcal{S}_\varepsilon(\tilde{\mathsf{a}});\ \forall\, v\in(t,s):\, V_v\in\bigcup_{|\tilde{\mathsf{b}}-\mathsf{b}'|\leqslant M_\varepsilon}\mathcal{S}_\varepsilon(\tilde{\mathsf{b}});\ (Y_s,V_s)\in\mathcal{T}_\varepsilon(\mathsf{i}')\times\mathcal{S}_\varepsilon(\mathsf{a}')\Big).\end{aligned} \tag{7.5}$$

Here t is the time $(Y_\cdot, V_\cdot)$ jumps from $\mathcal{T}_\varepsilon(\mathsf{i})\times\mathcal{S}_\varepsilon(\mathsf{b})$ to $\mathcal{T}_\varepsilon(\mathsf{i}')\times\mathcal{S}_\varepsilon(\mathsf{b}')$.

(II) Proof of (7.2)**.** First consider the case $|\vec{\mathsf{i}}| = 1$. By translation invariance, we may assume $\mathsf{i} = \mathsf{i}' = 1$ and $\mathsf{a} = 0$. First note that

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) &\leqslant 2\pi\int_0^\infty e^{\vartheta s}\mathrm{P}_{0,0}^{\varepsilon}\Big((Y_s,V_s)\in\mathcal{T}_\varepsilon(0)\times\mathcal{S}_\varepsilon(\mathsf{a}')\Big)\,\mathrm{d}s\\ &= \varepsilon^{-2}2\pi\iint\limits_{\substack{0<s<t<\varepsilon\\ x\in\mathcal{S}_\varepsilon(0),y\in\mathcal{S}_\varepsilon(\mathsf{a}')}}G_\vartheta(t-s)\,g_{\frac{t-s}{4}}(y-x)\mathrm{d}s\mathrm{d}t\mathrm{d}x\mathrm{d}y,\end{aligned} \tag{7.6}$$

where $G_\vartheta(t-s, y-x) := G_\vartheta(t-s)\,g_{\frac{t-s}{4}}(y-x)$ is the weighted Green's function defined for the Lévy process $\boldsymbol{Y}_s$, see (3.38) and (3.37).

When $\mathsf{a}' = 0$, we can relax the domain of integration in (7.6), use standard bounds on the Gaussian kernel g, and set $u := t-s$ to obtain

$$\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) \leqslant 2\pi\iint\limits_{u\in(0,\varepsilon)}G_\vartheta(u)g_{\frac{u}{4}}(y)\mathrm{d}u\mathrm{d}y \leqslant C\int_0^\varepsilon G_\vartheta(u)\mathrm{d}u \leqslant \frac{C}{\log\frac{1}{\varepsilon}}, \tag{7.7}$$

where we applied the asymptotics for $\int_0^\varepsilon G_\vartheta$ in (3.47).

When $\mathsf{a}' \neq 0$, the bound for $\sigma_\varepsilon^2(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ can be improved with an extra factor of $\frac{e^{-c|\mathsf{a}'|^2}}{\log \frac{1}{\varepsilon}}$. Indeed, using polar coordinates (with respect to the $|\cdot|_\infty$ norm) for $x \in \mathcal{S}_\varepsilon(0)$, we have

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}}, \vec{\mathsf{a}}) &\leqslant \varepsilon^{-2} 2\pi \iint\limits_{\substack{0<s<t<\varepsilon \\ x\in\mathcal{S}_\varepsilon(0), y\in\mathcal{S}_\varepsilon(\mathsf{a}')}} G_\vartheta(t-s)\, g_{\frac{t-s}{4}}(y-x)\mathrm{d}s\mathrm{d}t\mathrm{d}x\mathrm{d}y \\ &\leqslant \varepsilon^{-2} 2\pi \iint\limits_{0<s<t<\varepsilon} \mathrm{d}s\mathrm{d}t\, G_\vartheta(t-s) \int\limits_{0<r<1} 2\varepsilon r \mathrm{d}r \int\limits_{(|\mathsf{a}'|-r)\sqrt{\varepsilon}\leqslant|z|\leqslant(|\mathsf{a}'|+2)\sqrt{\varepsilon}} g_{\frac{t-s}{4}}(z)\mathrm{d}z. \quad (7.8)\end{aligned}$$

If $|\mathsf{a}'| \geqslant 2$, then we can use (3.46) to bound the right hand side of (7.8) by

$$\begin{aligned}&\varepsilon^{-1} 2\pi \iint\limits_{0<s<t<\varepsilon} \mathrm{d}s\mathrm{d}t\, G_\vartheta(t-s) \int\limits_{|\mathsf{a}'|-1\leqslant|w|} g_{\frac{t-s}{4\varepsilon}}(w)\mathrm{d}w \\ &\leqslant \varepsilon^{-1} 2\pi \iint\limits_{0<s<t<\varepsilon} G_\vartheta(t-s) e^{-\frac{|\mathsf{a}'|^2}{2(t-s)}\varepsilon}\, \mathrm{d}s\mathrm{d}t \leqslant 4\pi \int_0^1 \frac{1}{u(\log\frac{1}{u}+\log\frac{1}{\varepsilon})^2} e^{-\frac{|\mathsf{a}'|^2}{2u}}\, \mathrm{d}u \\ &\leqslant \frac{4\pi}{(\log\frac{1}{\varepsilon})^2} \int_0^1 \frac{1}{u} e^{-\frac{|\mathsf{a}'|^2}{2u}}\, \mathrm{d}u = \frac{4\pi}{(\log\frac{1}{\varepsilon})^2} \int_{|\mathsf{a}'|^2}^\infty \frac{1}{v} e^{-\frac{v}{2}} \mathrm{d}v \leqslant \frac{4\pi e^{-c|\mathsf{a}'|^2}}{(\log\frac{1}{\varepsilon})^2}.\end{aligned}$$

If $1 \leqslant |\mathsf{a}'| < 2$, then we can bound the right hand side of (7.8) by

$$\begin{aligned}&\varepsilon^{-1} 2\pi \iint\limits_{0<s<t<\varepsilon} \mathrm{d}s\mathrm{d}t\, G_\vartheta(t-s) \int_0^1 2r\mathrm{d}r \int\limits_{1-r\leqslant|w|} g_{\frac{t-s}{4\varepsilon}}(w)\mathrm{d}w \\ &\leqslant 2\pi \int\limits_{0<u<\varepsilon} G_\vartheta(u) \int_0^1 2r e^{-\frac{(1-r)^2}{u}\varepsilon}\, \mathrm{d}r\, \mathrm{d}u \leqslant\ 4\pi \int\limits_{0<u<\varepsilon} G_\vartheta(u) \int_0^1 e^{-\frac{r^2}{u}\varepsilon}\, \mathrm{d}r\, \mathrm{d}u \\ &\leqslant C \int_0^1 \frac{1}{v(\log\frac{1}{v}+\log\frac{1}{\varepsilon})^2} \int_0^1 e^{-\frac{r^2}{v}}\, \mathrm{d}r\, \mathrm{d}v \\ &\leqslant C \int_0^1 \frac{1}{\sqrt{v}(\log\frac{1}{v}+\log\frac{1}{\varepsilon})^2} \int_0^\infty e^{-s^2}\, \mathrm{d}s\, \mathrm{d}v \leqslant \frac{C}{(\log\frac{1}{\varepsilon})^2} \leqslant \frac{Ce^{-c|\mathsf{a}'|^2}}{(\log\frac{1}{\varepsilon})^2}.\end{aligned}$$

This concludes the proof of the upper bound in (7.2).

We now bound $\sigma_\varepsilon^2(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ for the case $|\vec{\mathsf{i}}| = 2$. By relaxing all the constraints in (7.5) except $(Y_s, V_s) \in \mathcal{T}_\varepsilon(\mathsf{i}') \times \mathcal{S}_\varepsilon(\mathsf{a}')$, we note that except for a change of constants, the bound in (7.2) for $|\vec{\mathsf{i}}| = 1$ also applies in this case. In particular, the bound in (7.2) holds for $|\vec{\mathsf{i}}| = 2$ and $|\vec{\mathsf{a}}| \neq 0$. For $|\vec{\mathsf{i}}| = 2$ and $|\vec{\mathsf{a}}| = 0$, let us assume for simplicity that $\mathsf{i} = 1$, $\mathsf{i}' = 2$, and $\mathsf{a} = 0$.

Again, relaxing all constraints in (7.5) except the constraint on (Y_s, V_s), we have

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) &\leqslant C\varepsilon^{-2}\iint\limits_{\substack{0<s<\varepsilon<t<2\varepsilon\\ x,y\in\mathcal{S}_\varepsilon(0)}} G_\vartheta(t-s)g_{\frac{t-s}{4}}(y-x)\mathrm{d}x\mathrm{d}y\mathrm{d}s\mathrm{d}t\\ &\leqslant C\varepsilon^{-1}\iint\limits_{0<s<\varepsilon<t<2\varepsilon} G_\vartheta(t-s)\mathrm{d}s\mathrm{d}t \leqslant C\int_0^\varepsilon uG_\vartheta(u)\mathrm{d}u\\ &\leqslant C\int_0^\varepsilon \frac{1}{(\log\frac{1}{u})^2}\mathrm{d}u \leqslant \frac{C}{(\log\frac{1}{\varepsilon})^2}.\end{aligned} \tag{7.9}$$

The upper bound in (7.2) also holds.

We now consider the case $|\vec{\mathsf{i}}| \geqslant 3$. We first ignore the constraints on V_r for $r \in (0,t)\cup(t,s)$ in (7.5). Using the weighted Green's function G_ϑ and the Lévy measure $\frac{\mathbb{1}_{(0,1)}}{t}g_{t/4}\mathrm{d}t\mathrm{d}x$ for the Lévy process $\boldsymbol{Y}_s = (Y_s, V_s)$ (see [CSZ19a, Section 2]), we obtain the bound

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) \leqslant \sum_{\substack{\mathsf{b}:\,|\mathsf{b}|\leqslant M_\varepsilon\\ \mathsf{b}':\,|\mathsf{b}'-\mathsf{a}'|\leqslant M_\varepsilon,\,|\mathsf{b}'-\mathsf{b}|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}}} C\varepsilon^{-2}\iint\limits_{\substack{0<s<t<\varepsilon\\ x\in\mathcal{S}_\varepsilon(0),y\in\mathcal{S}_\varepsilon(\mathsf{b})}}\iint\limits_{\substack{(\mathsf{i}'-1)\varepsilon<s'<t'<\mathsf{i}'\varepsilon\\ x'\in\mathcal{S}_\varepsilon(b'),y'\in\mathcal{S}_\varepsilon(\mathsf{a}')}} \mathrm{d}x\mathrm{d}y\mathrm{d}x'\mathrm{d}y'\mathrm{d}s\mathrm{d}t\mathrm{d}s'\mathrm{d}t'\\ G_\vartheta(t-s)g_{\frac{t-s}{4}}(y-x)\cdot\frac{g_{\frac{s'-t}{4}}(x'-y)}{s'-t}\cdot G_\vartheta(t'-s')\cdot g_{\frac{t'-s'}{4}}(y'-x').\end{aligned}$$

Since $|\vec{\mathsf{i}}| = \mathsf{i}'-\mathsf{i}+1 \geqslant 3$, we can bound $\frac{1}{s'-t} \leqslant \frac{1}{(|\vec{\mathsf{i}}|-2)\varepsilon} \leqslant \frac{3}{|\vec{\mathsf{i}}|\varepsilon}$ to obtain

$$\begin{aligned}\sigma_\varepsilon^2(\vec{\mathsf{i}},\vec{\mathsf{a}}) &\leqslant \frac{C}{|\vec{\mathsf{i}}|\varepsilon^3}\iint\limits_{\substack{0<s<t<\varepsilon\\ (\mathsf{i}'-1)\varepsilon<s'<t'<\mathsf{i}'\varepsilon}}\iint\limits_{\substack{x\in\mathcal{S}_\varepsilon(0)\\ y'\in\mathcal{S}_\varepsilon(\mathsf{a}')}} G_\vartheta(t-s)G_\vartheta(t'-s')g_{\frac{t'-s}{4}}(y'-x)\,\mathrm{d}s\mathrm{d}t\mathrm{d}s'\mathrm{d}t'\mathrm{d}x\mathrm{d}y'\\ &\leqslant \frac{C}{|\vec{\mathsf{i}}|\varepsilon^3}\Big(\int_0^\varepsilon G_\vartheta(u)\mathrm{d}u\Big)^2\iint\limits_{\substack{0<s<\varepsilon\\ (\mathsf{i}'-1)\varepsilon<t'<\mathsf{i}'\varepsilon}}\iint\limits_{\substack{x\in\mathcal{S}_\varepsilon(0)\\ y'\in\mathcal{S}_\varepsilon(\mathsf{a}')}} g_{\frac{t'-s}{4}}(y'-x)\,\mathrm{d}s\mathrm{d}t'\mathrm{d}x\mathrm{d}y'\\ &\leqslant \frac{C}{(\log\frac{1}{\varepsilon})^2}\frac{e^{-c|\vec{\mathsf{a}}|^2/|\vec{\mathsf{i}}|}}{|\vec{\mathsf{i}}|^2},\end{aligned} \tag{7.10}$$

where we first relaxed the constraints on b and b', then successively integrated out y, x', s', and t and applied (3.47), while in the last inequality, we applied a uniform bound on the heat kernel $g_{\frac{t'-s}{4}}(y'-x)$. This concludes the proof of the upper bound in (7.2). □

7.2. Fourth moment estimates. We next study the fourth moment.

Lemma 7.2. *Let $\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ be defined as in (4.11), with $(\vec{\mathsf{i}},\vec{\mathsf{a}}) = ((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}')) \in \mathbb{T}_\varepsilon$ defined in (4.5). There exist $c, C \in (0,\infty)$ uniform in $(\vec{\mathsf{i}},\vec{\mathsf{a}})$, such that for all $\varepsilon > 0$ sufficiently small,*

$$\limsup_{N\to\infty}\mathbb{E}\big[\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})^4\big] < \frac{Ce^{-c|\vec{\mathsf{a}}|/\sqrt{|\vec{\mathsf{i}}|}}\mathbb{1}_{\{|\vec{\mathsf{a}}|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}\}}}{\log\frac{1}{\varepsilon}}. \tag{7.11}$$

Proof. We first prove (7.11) for $|\vec{\mathsf{i}}| = \mathsf{i}' - \mathsf{i} + 1 \leqslant 2$. Consider a time-space block $(\vec{\mathsf{i}}, \vec{\mathsf{a}}) = ((\mathsf{i}, \mathsf{i}'), (\mathsf{a}, \mathsf{a}'))$ with $|\vec{\mathsf{a}}| \leqslant M_\varepsilon \sqrt{|\vec{\mathsf{i}}|}$ and assume without loss of generality that $\mathsf{i} = \mathsf{i}' = 1$ and $\mathsf{a} = 0$. The case $|\vec{\mathsf{i}}| = 2$ is similar (just replace ε by 2ε). We will compare $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{1}, \vec{\mathsf{a}})$ with an averaged partition function so that Theorem 6.1 can be applied.

Let us recall the polynomial chaos expansion of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ from (4.11)

$$\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{1}, \vec{\mathsf{a}}) := \frac{2}{\varepsilon N} \sum_{\substack{(d,x) \in \mathcal{B}_{\varepsilon N}(1,0) \\ (f,y) \in \mathcal{B}_{\varepsilon N}(1,\mathsf{a}') \\ \text{with } d \leqslant f}} X^{(\mathrm{diff})}_{d,f}(x,y),$$

which is essentially an average of point-to-point partition functions with average over (d,x) and (f,y) in the bulk instead of through boundary conditions at time 0 and εN respectively. To compare with an averaged partition function as in Theorem 6.1, we replace $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{1}, \vec{\mathsf{a}})$ by

$$\Theta := \frac{2}{\varepsilon N} \sum_{z_1, z_2 \in \mathbb{Z}^2} \sum_{\substack{(d,x) \in \mathcal{B}_{\varepsilon N}(1,0) \\ (f,y) \in \mathcal{B}_{\varepsilon N}(1,\mathsf{a}') \\ \text{with } d \leqslant f}} \mathbb{1}_{\mathcal{S}_{\varepsilon N}(0)}(z_1) q_{0,d}(z_1, x) X^{(\mathrm{diff})}_{d,f}(x,y) q_{f,\varepsilon N}(y, z_2) \mathbb{1}_{\mathcal{S}_{\varepsilon N}(\mathsf{a}')}(z_2), \quad (7.12)$$

where $\mathcal{S}_{\varepsilon N}(\mathsf{a}) = ((\mathsf{a} - (1,1))\sqrt{\varepsilon N}, \mathsf{a}\sqrt{\varepsilon N}]$ and we note that uniformly in $(d,x) \in \mathcal{B}_{\varepsilon N}(1,0)$ and $(f,y) \in \mathcal{B}_{\varepsilon N}(1,\mathsf{a}')$, we have

$$\sum_{z_1, z_2 \in \mathbb{Z}^2} \mathbb{1}_{\mathcal{S}_{\varepsilon N}(0)}(z_1) q_{0,d}(z_1, x) q_{f,\varepsilon N}(y, z_2) \mathbb{1}_{\mathcal{S}_{\varepsilon N}(\mathsf{a}')}(z_2) \geqslant C > 0.$$

Therefore $\mathrm{E}[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, \vec{\mathsf{a}})^4] \leqslant C\mathrm{E}[\Theta^4]$ because in the expansion for the fourth moment, all terms are non-negative if we assume $\mathrm{E}[\xi_N^3] \geqslant 0$, which we may assume from now on since our bounds are in terms of $|\mathrm{E}[\xi_N^k]|$ for $1 \leqslant k \leqslant 4$ (see the proof of Theorem 6.1). In the definition of Θ, we can further remove the constraint on y and the summation constraints in the definition of $X^{(\mathrm{diff})}_{d,f}(x,y)$ in (4.9)–(4.10), which gives the centred partition function

$$\mathcal{Z}^{\beta_N}_{\varepsilon N}(\varphi, \psi) - \mathrm{E}[\mathcal{Z}^{\beta_N}_{\varepsilon N}(\varphi, \psi)]$$

as defined in Theorem 6.1, with $\varphi(x) = \mathbb{1}_{\mathcal{S}_1(0)}(x)$ and $\psi(x) = \mathbb{1}_{\mathcal{S}_1(\mathsf{a}')}(x)$. Therefore for N large, we have

$$\begin{aligned} \mathrm{E}\big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, \vec{\mathsf{a}})^4\big] &\leqslant C\big[(\mathcal{Z}^{\beta_N}_{\varepsilon N}(\varphi, \psi) - \mathrm{E}[\mathcal{Z}^{\beta_N}_{\varepsilon N}(\varphi, \psi)])^4\big] \\ &\leqslant \frac{C}{\log \frac{1}{\varepsilon}} \Big\| \frac{\varphi}{w} \Big\|_2^4 \|\psi\|_\infty^4 \|w \mathbb{1}_{\mathcal{S}_1(\mathsf{a}')}\|_2^4 \leqslant \frac{C e^{-|\mathsf{a}' - \mathsf{a}|}}{\log \frac{1}{\varepsilon}}, \end{aligned} \quad (7.13)$$

where we applied Theorem 6.1 with N set to εN, $T = 1/\varepsilon$, $p = q = 2$, $h = 4$, and $w(x) = e^{-|x|}$. This proves (7.11) for $|\vec{\mathsf{i}}| \leqslant 2$.

We now consider the case $|\vec{\mathsf{i}}| \geqslant 3$. Recall the definition of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ from (4.11), we can rewrite it as

$$\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, \vec{\mathsf{a}}) = \sum_{\mathsf{b}, \mathsf{b}'} \Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}, (\mathsf{a}, \mathsf{b}), (\mathsf{b}', \mathsf{a}')), \quad (7.14)$$

where

$$\begin{aligned}&\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))\\&=\frac{2}{\varepsilon N}\sum_{\substack{(d,x)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})\\(f',y')\in\mathcal{B}_{\varepsilon N}(\mathsf{i}',\mathsf{a}')}}\sum_{\substack{(f,y)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{b})\\(d',x')\in\mathcal{B}_{\varepsilon N}(\mathsf{i}',\mathsf{b}')\\d\leqslant f,\,d'\leqslant f'}}X^{(\mathrm{diff})}_{d,f}(x,y)\,q_{f,d'}(y,x')\,X^{(\mathrm{diff})}_{d',f'}(x',y').\end{aligned}\tag{7.15}$$

For each (b,b'), because $\mathsf{i}'-\mathsf{i}\geqslant 2$, we can apply Lemma 3.2 (with $m=\varepsilon N$) to bound

$$q_{f,d'}(y,x')\leqslant\frac{C}{\varepsilon N|\vec{\mathsf{i}}|}\,e^{-c|\mathsf{b}-\mathsf{b}'|^2/|\vec{\mathsf{i}}|}$$

uniformly in $(f,y)\in\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{b})$ and $(d',x')\in\mathcal{B}_{\varepsilon N}(\mathsf{i}',\mathsf{b}')$. We can make this replacement in the bound for the fourth moment to obtain

$$\mathbb{E}\big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})^4\big]\leqslant C\,\mathbb{E}\Big[\Big(\sum_{\mathsf{b},\mathsf{b}'}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}))\frac{e^{-c|\mathsf{b}-\mathsf{b}'|^2/|\vec{\mathsf{i}}|}}{|\vec{\mathsf{i}}|}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i}',\mathsf{i}'),(\mathsf{b}',\mathsf{a}'))\Big)^4\Big].\tag{7.16}$$

By triangle inequality, we can split the sum over b,b' into three parts (with overlaps): (1) $|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3$; (2) $|\mathsf{b}'-\mathsf{a}'|\geqslant|\vec{\mathsf{a}}|/3$; $|\mathsf{b}'-\mathsf{b}|\geqslant|\vec{\mathsf{a}}|/3$. It suffices to bound the fourth moment of each part.

For part (1), we can bound

$$\begin{aligned}&\mathbb{E}\Big[\Big(\sum_{\substack{\mathsf{b},\mathsf{b}'\\|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3}}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}))\frac{e^{-c|\mathsf{b}-\mathsf{b}'|^2/|\vec{\mathsf{i}}|}}{|\vec{\mathsf{i}}|}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i}',\mathsf{i}'),(\mathsf{b}',\mathsf{a}'))\Big)^4\Big]\\&\leqslant\mathbb{E}\Big[\Big(\sum_{\mathsf{b}:|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}))\Big)^4\Big]\mathbb{E}\Big[\Big(\sum_{\mathsf{b}'}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i}',\mathsf{i}'),(\mathsf{b}',\mathsf{a}'))\Big)^4\Big],\end{aligned}\tag{7.17}$$

where the inequality can be justified if we first expand the power and take expectation and then bound $\frac{e^{-c|\mathsf{b}-\mathsf{b}'|^2/|\vec{\mathsf{i}}|}}{|\vec{\mathsf{i}}|}<1$; we also used the independence of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),\cdot)$ and $\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i}',\mathsf{i}'),\cdot)$. For the first factor in (7.17), we can expand the power and bound

$$\mathbb{E}\Big[\Big(\sum_{|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3}\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}))\Big)^4\Big]=\sum_{\substack{|\mathsf{b}_i-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3\\\text{for }1\leqslant i\leqslant 4}}\mathbb{E}\Big[\prod_{i=1}^4\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}_i))\Big]$$

$$\leqslant\sum_{\substack{|\mathsf{b}_i-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3\\\text{for }1\leqslant i\leqslant 4}}\prod_{i=1}^4\mathbb{E}\Big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}_i))^4\Big]^{\frac14}=\Big(\sum_{|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3}\mathbb{E}\Big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\mathsf{i},\mathsf{i}),(\mathsf{a},\mathsf{b}))^4\Big]^{\frac14}\Big)^4\leqslant\frac{Ce^{-c|\vec{\mathsf{a}}|}}{\log\frac{1}{\varepsilon}},$$

where in the last inequality, we applied the fourth moment bound (7.13) for $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\cdot)$ with $|\vec{\mathsf{i}}|=1$. The second factor in (7.17) can be bounded the same way without the factor $e^{-c|\vec{\mathsf{a}}|}$. This implies that when the sum in (7.16) is restricted to $|\mathsf{b}-\mathsf{a}|\geqslant|\vec{\mathsf{a}}|/3$, we get a fourth moment bound of $Ce^{-c|\vec{\mathsf{a}}|}/(\log\frac{1}{\varepsilon})^2$. The same bound holds if the sum is restricted to $|\mathsf{b}'-\mathsf{a}'|\geqslant|\vec{\mathsf{a}}|/3$.

When the sum in (7.16) is restricted to $|\mathsf{b}-\mathsf{b}'|\geqslant|\vec{\mathsf{a}}|/3$, we can bound $\frac{e^{-c|\mathsf{b}-\mathsf{b}'|^2/|\vec{\mathsf{i}}|}}{|\vec{\mathsf{i}}|}\leqslant e^{-c|\vec{\mathsf{a}}|^2/|\vec{\mathsf{i}}|}$. The rest of the calculations is the same as before, which leads to a fourth moment bound

of $Ce^{-c|\vec{\mathsf{a}}|^2/|\vec{\mathsf{i}}|}/(\log\frac{1}{\varepsilon})^2$. Combined with the previous estimates, it is clear that (7.11) holds. This concludes the proof of Lemma 7.2. □

8. Moment estimates for the coarse-grained model

In this section, we will prove an analogue of Theorem 6.1 for the coarse-grained model (defined in (4.8)), that we rewrite for convenience:

$$\begin{aligned}\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta) := \frac{1}{2}g_{\frac12}(\varphi,\psi) \,+\, \frac{\varepsilon}{2}\sum_{r=1}^{(\log\frac{1}{\varepsilon})^2}\sum_{\mathsf{b},\mathsf{c}\in\mathbb{Z}^2}\sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}\\(\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\,b,c}^{(\mathrm{diff})}}}\varphi_\varepsilon(\mathsf{b})g_{\frac12\mathsf{i}_1}(\mathsf{a}_1-\mathsf{b})\Theta(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1)\\ \times\Bigg\{\prod_{j=2}^{r}g_{\frac12(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\,\Theta(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j)\Bigg\}g_{\frac12(\lfloor\frac{1}{\varepsilon}\rfloor-\mathsf{i}'_r)}(\mathsf{c}-\mathsf{a}'_r)\psi_\varepsilon(\mathsf{c})\,,\end{aligned}\tag{8.1}$$

with coarse-grained disorder variables $\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}}) := \Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ (see (4.11)) indexed by time-space blocks $(\vec{\mathsf{i}},\vec{\mathsf{a}}) = ((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}'))$ in the set $\mathbb{T}_\varepsilon$ (see (4.5)), while $\varphi_\varepsilon,\psi_\varepsilon:\mathbb{Z}^2\to\mathbb{R}$ are defined by (4.7) from $\varphi\in C_c(\mathbb{R}^2)$ and $\psi\in C_b(\mathbb{R}^2)$.

We will prove the following analogue of Theorem 6.1 for the 4-th moment of the coarse-grained model.

Theorem 8.1. *Let $\mathscr{Z}_{N,\varepsilon}(\varphi,\psi) := \mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta_{N,\varepsilon}^{(\mathrm{cg})})$ be the coarse-grained model defined above. Further assume that $\|\psi\|_\infty<\infty$ and ψ is supported on a ball B (possibly $B=\mathbb{R}^d$). Then for any $p,q\in(1,\infty)$ with $\frac1p+\frac1q=1$ and any $w:\mathbb{R}^2\to(0,\infty)$ such that $\log w$ is Lipschitz continuous, there exists $C\in(0,\infty)$ such that uniformly in $\varepsilon\in(0,1)$,*

$$\limsup_{N\to\infty}\mathbb{E}\Big[\big(\mathscr{Z}_{N,\varepsilon}(\varphi,\psi)-\mathbb{E}[\mathscr{Z}_{N,\varepsilon}(\varphi,\psi)]\big)^4\Big]\leqslant C\varepsilon^{\frac{4}{p}}\Big\|\frac{\varphi_\varepsilon}{w_\varepsilon}\Big\|_{\ell^p(\mathbb{Z}^2)}^4\|\psi\|_\infty^4\|w\mathbb{1}_B\|_q^4,\tag{8.2}$$

where $w_\varepsilon:\mathbb{Z}^2\to\mathbb{R}$ is defined from w by (4.7).

Proof. We will adapt the proof of Theorem 6.1 to the current setting. The complication is that the coarse-grained disorder variables $\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ are assigned to time-space blocks $(\vec{\mathsf{i}},\vec{\mathsf{a}})$ instead of individual lattice sites. We will therefore divide the proof into two parts: first, expand the fourth moment and perform a resummation to bring it into a similar framework as in the proof of Theorem 6.1; second, give the necessary bounds analogous to those in the proof of Theorem 6.1.

Part I. Expansion. First, as in (6.4) in the proof of Theorem 6.1, denote

$$\begin{aligned}\mathcal{M}_{N,\varepsilon}^{\varphi,\psi} :=&\mathbb{E}\Big[\big(\mathscr{Z}_{N,\varepsilon}(\varphi,\psi)-\mathbb{E}[\mathscr{Z}_{N,\varepsilon}(\varphi,\psi)]\big)^4\Big]\\ =&\frac{\varepsilon^4}{16}\,\mathbb{E}\Big[\Big(\sum_{r=1}^{(\log\frac{1}{\varepsilon})^2}\sum_{\mathsf{b},\mathsf{c}\in\mathbb{Z}^2}\sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}\\(\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\,b,c}^{(\mathrm{diff})}}}\varphi_\varepsilon(\mathsf{b})g_{\frac12\mathsf{i}_1}(\mathsf{a}_1-\mathsf{b})\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1)\\ &\times\Big\{\prod_{j=2}^{r}g_{\frac12(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\Theta_{N,\varepsilon}^{(\mathrm{cg})}(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j)\Big\}g_{\frac12(\frac{1}{\varepsilon}-\mathsf{i}'_r)}(\mathsf{c}-\mathsf{a}'_r)\psi_\varepsilon(\mathsf{c})\Big)^4\Big].\end{aligned}\tag{8.3}$$

By assumption, we have $|\psi_\varepsilon| \leqslant \|\psi\|_\infty \mathbb{1}_{B_\varepsilon}$, where $B_\varepsilon = B/\sqrt{\varepsilon}$. By the same reasoning as in the proof of Theorem 6.1 (see the discussion leading to (6.12)), we can replace ψ by $\|\psi\|_\infty \mathbb{1}_{B_\varepsilon}$ and replace $g_{\frac{1}{2}(\frac{1}{\varepsilon}-\mathsf{i}'_r)}(\cdot)$ by $g_{\frac{1}{2}(\tilde{n}-\mathsf{i}'_r)}(\cdot)$ (with $\tilde{n}$ first summed over $[\varepsilon^{-1}, 2\varepsilon^{-1}]$, then extended to $[1, 2\varepsilon^{-1}]$) to obtain the following bound

$$\begin{aligned}\mathcal{M}^{\varphi,\psi}_{N,\varepsilon} \leqslant C\|\psi\|_\infty^4 \,\varepsilon^5 \sum_{\tilde{n}=1}^{2/\varepsilon} \mathbb{E}\Big[\Big(\sum_{r=1}^{\infty} \sum_{\mathsf{b},\mathsf{c}\in\mathbb{Z}^2} \sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_\varepsilon \\ (\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in\vec{\boldsymbol{\mathcal{A}}}^{(\text{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}}} \varphi_\varepsilon(\mathsf{b}) g_{\frac{1}{2}\mathsf{i}_1}(\mathsf{a}_1-\mathsf{b})\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_1,\vec{\mathsf{a}}_1) \\ \times\Big\{\prod_{j=2}^{r} g_{\frac{1}{2}(\mathsf{i}_j-\mathsf{i}'_{j-1})}(\mathsf{a}_j-\mathsf{a}'_{j-1})\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_j,\vec{\mathsf{a}}_j)\Big\} g_{\frac{1}{2}(\tilde{n}-\mathsf{i}'_r)}(\mathsf{c}-\mathsf{a}'_r)\mathbb{1}_{B_\varepsilon}(\mathsf{c})\Big)^4\Big].\end{aligned} \tag{8.4}$$

We then expand the product in (8.4) to obtain the sum over 4 sequences of time-space blocks, each time-space block contributing a $\Theta^{(\text{cg})}_{N,\varepsilon}$ variable. Because we will bound the sum by taking the absolute value of each summand, we can relax the summation constraint on r to obtain an upper bound. Also note that thanks to the assumption $\varphi \in C_c(\mathbb{R}^2)$ and the diffusive constraint $\vec{\boldsymbol{\mathcal{A}}}^{(\text{diff})}_{\varepsilon;\mathsf{b},\mathsf{c}}$ (see (4.6)), we have a sum with finitely many terms, which allows us to pass the limit $N\to\infty$ inside the sum later. For each $\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$ with $|\vec{\mathsf{i}}| = \mathsf{i}'-\mathsf{i}+1 \geqslant 2$, we further expand it as

$$\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}}) = \sum_{\substack{\mathsf{b}:\,|\mathsf{b}-\mathsf{a}|\leqslant M_\varepsilon \\ \mathsf{b}':\,|\mathsf{b}'-\mathsf{a}'|\leqslant M_\varepsilon \\ |\mathsf{b}'-\mathsf{b}|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}}} \Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}')), \tag{8.5}$$

where $\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))$ is defined as $\Theta^{(\text{cg})}_{N,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}})$, except the sum in its definition in (4.11) is restricted to a fixed choice of $\vec{\mathsf{b}} := (\mathsf{b},\mathsf{b}')$. The expansion of the product in (8.4) then gives 4 sequences of coarse-grained disorder variables $\Theta^{(\text{cg})}_{N,\varepsilon}$, some of which may visit two distinct mesoscopic time intervals with indices i,i' due to the expansion in (8.5). If we record the indices of the visited mesoscopic time intervals and the mesoscopic spatial boxes of entry and exit in each time interval, then we obtain 4 sequences of time-space indices $(\mathsf{i}^j_1,\mathsf{a}^j_1,\mathsf{b}^j_1)$, …, $(\mathsf{i}^j_{r_j},\mathsf{a}^j_{r_j},\mathsf{b}^j_{r_j})$, $1\leqslant j\leqslant 4$. We will call each such sequence a *mesoscopic time-space renewal sequence*, or just *renewal sequence* (see Figure 4).

We will rearrange the expansion of (8.4) as follows:

(1) Sum over the set $\bigcup_{j=1}^4\{\mathsf{i}^j_1,\dots,\mathsf{i}^j_{r_j}\} =: \{n_1,\dots,n_r\}$.

(2) For each time index n_i, $1\leqslant i\leqslant r$, sum over the set of indices $J_i \subset \{1,\dots,4\}$, which determine the renewal sequences that visit time interval n_i.

(3) For each $j\in J_i$, i.e., a renewal sequence that visits time interval n_i, sum over the indices $(\mathsf{a}^j_i,\mathsf{b}^j_i)$ that determine the spatial boxes of entry and exit in that time interval.

Given a choice of these summation indices, the summand contains a product of coarse grained disorder variables of either the form $\Theta^{(\text{cg})}_{N,\varepsilon}((n_i,n_i),(\mathsf{a}_i,\mathsf{b}_i)) =: \Theta^{(\text{cg})}_{N,\varepsilon}(n_i;\mathsf{a}_i,\mathsf{b}_i)$ or the form $\Theta^{(\text{cg})}_{N,\varepsilon}((n_i,n_j),(\mathsf{a}_i,\mathsf{b}_i),(\mathsf{a}_j,\mathsf{b}_j))$, connected by heat kernels $g_{\frac{1}{2}(n_k-n_l)}(\mathsf{a}_k-\mathsf{b}_l)$. For such a product to have non-zero expectation, we have the following constraints (see Figure 4):

(a) $|J_i|\geqslant 2$ for each $1\leqslant i\leqslant r$;

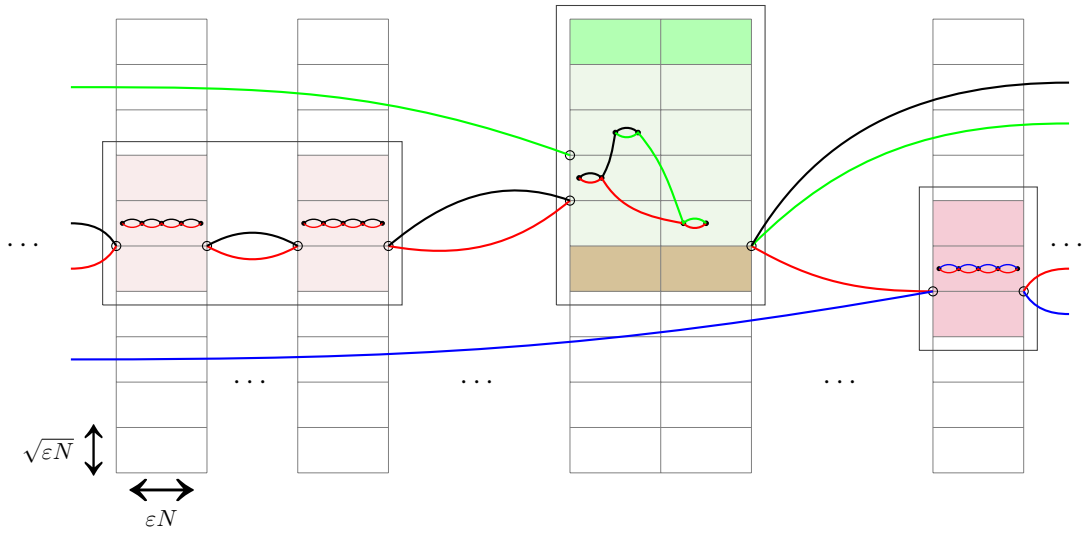


FIGURE 4. A depiction of the expansion of the fourth moment of the coarse-grained model that satisfy conditions (a)-(g). There are four time-space renewal sequences, each depicted in a different colour. Different sequences visit different mesoscopic boxes $\mathcal{B}_{\varepsilon N}(\mathsf{i},\mathsf{a})$, but each visited box must be visited by at least two sequences to give a non-zero contribution. The first two time strips are visited by the two renewal sequences coloured black and red, which match in the disorder they sample. These two strips are grouped together as a *block of type* U. The third and fourth time strips are visited by three renewal sequences, coloured black, red, and green, which form a *block of type* V and its width cannot exceed 4. Within this V block, the spatial boxes of entry by the three renewals are all within distance $2M_\varepsilon$ of each other, 2 of which match exactly. The last time strip is only visited by two renewal sequences, coloured blue and red, which also forms a U block.

(b) If $|J_i| = 2$, say $J_i = \{k,l\} \subset \{1,\dots,4\}$, then we must have $\mathsf{a}_i^k = \mathsf{a}_i^l$ and $\mathsf{b}_i^k = \mathsf{b}_i^l$;

(c) If $|J_i| \geqslant 3$, then for each sequence $k \in J_i$, there must be another sequence $l \in J_i$ such that $|\mathsf{a}_i^k - \mathsf{a}_i^l| \leqslant 2M_\varepsilon$, where $M_\varepsilon = \log\log\frac{1}{\varepsilon}$ as in (4.2).

If (c) is violated, then by the spatial constraint in the definition of $\Theta_{N,\varepsilon}^{(\mathrm{cg})}$ in (4.11), there will be a coarse-grained disorder variable visiting time interval n_i, which is independent of all other coarse-grained disorder variables in the product, and hence leads to zero expectation.

The summation constraints $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$ and $\vec{\boldsymbol{\mathcal{A}}}_{\varepsilon;\mathsf{b},\mathsf{c}}^{(\text{diff})}$ (see (4.4) and (4.6)) in the definition (8.1) of the coarse-grained model implies the following additional constraints on the summation indices r, $(n_i)_{1\leqslant i\leqslant r}$, $(J_i)_{1\leqslant i\leqslant r}$, and $(\mathsf{a}_i^j,\mathsf{b}_i^j)_{1\leqslant i\leqslant r, j\in J_i}$:

(d) For all $1 \leqslant i \leqslant r$ and each renewal sequence with index $j \in J_i$, $|\mathsf{b}_i^j - \mathsf{a}_i^j| \leqslant M_\varepsilon$;

(e) For $1 \leqslant j \leqslant 4$ and $1 \leqslant i_1 < i_2 \leqslant r$, if $j \in J_{i_1} \cap J_{i_2}$ and $j \notin J_i$ for all $i_1 < i < i_2$ (namely renewal sequence j visits the mesoscopic time intervals with indices n_{i_1} and n_{i_2}, but nothing in between), then $|\mathsf{a}_{i_2}^j - \mathsf{b}_{i_1}^j| \leqslant M_\varepsilon\sqrt{n_{i_2} - n_{i_1}}$;

(f) $K_\varepsilon \leqslant n_1 < n_2 < \cdots < n_r \leqslant \frac{1}{\varepsilon} - K_\varepsilon$, where $K_\varepsilon = (\log\frac{1}{\varepsilon})^6$ as in (4.2);

(g) $(n_1, \dots, n_r)$ can be partitioned into consecutive stretches $\mathcal{D}_1, \dots, \mathcal{D}_m$ such that each $\mathcal{D}_i$ consists of consecutive integers, with a gap between $\mathcal{D}_i$ and $\mathcal{D}_{i+1}$. Then each $\mathcal{D}_\cdot = (n_i, n_{i+1} = n_i + 1, \dots, n_j = n_i + (j-i))$ has width $n_j - n_i + 1 \leqslant 4$, since $|J_{n_l}| \geqslant 2$ for $i \leqslant l \leqslant j$ (namely the mesoscopic time interval with index n_l is visited by at least two renewal sequences), and each sequence can visit at most two mesoscopic time intervals with indices among $n_i, n_{i+1}, \dots, n_j$ by the constraint $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$.

Conditions (d)-(e) follow from the definitions of $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{diff})}$ in (4.6) and $\Theta_{N,\varepsilon}^{(\text{cg})}$ in (4.11), while conditions (f)-(g) follow from the definition of $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$ in (4.4).

To handle the dependency among the coarse-grained disorder variables in the expansion of (8.4), we perform a further resummation. First partition $(n_1, \dots, n_r)$ into consecutive stretches $\mathcal{D}_1, \dots, \mathcal{D}_m$ as in (g), so that $\{n_1, \dots, n_r\} = \bigcup_{i=1}^m \mathcal{D}_i$. For each $\mathcal{D}_i$, let $\tilde{J}_i := \bigcup_{j \in \mathcal{D}_i} J_j$, which records which of the 4 renewal sequences visits the stretch $\mathcal{D}_i$. Next we group together consecutive $\mathcal{D}_{i_1}, \mathcal{D}_{i_1+1}, \dots, \mathcal{D}_{i_2}$ with the same $\tilde{J}_i = \{k, l\}$ for some $k \neq l \in \{1, \dots, 4\}$, and only keep track of $s := \min \mathcal{D}_{i_1}$ and $t := \max \mathcal{D}_{i_2}$, thus effectively replacing $\bigcup_{i_1 \leqslant i \leqslant i_2} \mathcal{D}_i$ by $[s, t]$. This allows us to identify from $(n_1, \dots, n_r)$ a sequence of disjoint time intervals (which we call blocks) $\mathcal{I}_i = [s_i, t_i] \cap \mathbb{N}$, $1 \leqslant i \leqslant k$, each associated with a label set $\mathcal{J}_i \subset \{1, \dots, 4\}$. Some of these intervals arise from $\bigcup_{i_1 \leqslant i \leqslant i_2} \mathcal{D}_i$ as above, which are visited by exactly 2 renewal sequences, the rest coincide with the original $\mathcal{D}_i$'s. We can then rewrite the expansion of (8.4) as follows:

(1') Sum over integers $K_\varepsilon < s_1 \leqslant t_1 < s_2 \leqslant t_2 < \cdots < s_k \leqslant t_k < \tilde{n} \leqslant \frac{2}{\varepsilon}$, with $s_{i+1} - t_i \geqslant 2$ for each i (recall the summation index $\tilde{n}$ from (8.4)). Denote $\mathcal{I}_i := [s_i, t_i] \cap \mathbb{N}$.

(2') For each block $\mathcal{I}_i$, sum over the set of indices $\mathcal{J}_i \subset \{1, \dots, 4\}$ with $|\mathcal{J}_i| \geqslant 2$. If $|\mathcal{J}_i| = 2$, we call $\mathcal{I}_i$ a *block of type* U because it leads to contributions similar to $\mathsf{U}_s(\boldsymbol{z}^1, \boldsymbol{z}^2)$ in (6.14) (see also (3.41)); otherwise we call it a *block of type* V. There are no consecutive blocks $\mathcal{I}_i$, $\mathcal{I}_{i+1}$ of both type U with the same label set $\mathcal{J}_i = \mathcal{J}_{i+1}$, and each block $\mathcal{I}_i$ of type V must have length $|\mathcal{I}_i| \leqslant 4$. See Figure 4 for an illustration.

(3') For each block $\mathcal{I}_i$ and each renewal sequence $j \in \mathcal{J}_i$ that visits block $\mathcal{I}_i$, sum over time-space indices $(\sigma_i^j, \mathsf{a}_i^j)$ and $(\tau_i^j, \mathsf{b}_i^j)$ with $s_i \leqslant \sigma_i^j \leqslant \tau_i^j \leqslant t_i$ and $\mathsf{a}_i^j, \mathsf{b}_i^j \in \mathbb{Z}^2$, which identifies the mesoscopic time-space blocks of entry and exit by the j-th renewal sequence in the time interval $\mathcal{I}_i = [s_i, t_i] \cap \mathbb{N}$.

The constraints imposed in (d)-(g) carry over, so we do not repeat them here.

To rewrite the expansion of (8.4) in a form that fits the framework developed in the proof of Theorem 6.1, we will carry out the following steps, that we describe below.

(A) To decouple different blocks, replace each coarse-grained disorder variable $\Theta_{N,\varepsilon}^{(\text{cg})}$ (it will arise as a summand in (8.5)) that visits two consecutive blocks $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ by the product of two coarse-grained disorder variables of the form $\Theta_{N,\varepsilon}^{(\text{cg})}(\vec{\mathsf{i}}, \vec{\mathsf{a}})$ with $|\vec{\mathsf{i}}| = 1$, joined by a heat kernel.

(B) Bound the moments of the effective disorder variable associated with each U block and V block.

(C) Modification of the heat kernels $g_{\frac{1}{2}(s_{i+1} - t_i)}(\cdot)$ connecting different blocks. In particular, carry out a Chapman-Kolmogorov type decomposition for the heat kernels as in the

proof of Theorem 6.1, so that we sum over the spatial locations for all 4 renewal sequences at the beginning and end of each block $\mathcal{I}_i$.

This rewriting will introduce a constant factor for each block, but it will not affect our final result. We now give the details for **(A)-(C)**.

(A) Note the technical complication that given a realisation of the summation indices in (1')–(3'), there could be coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\tau^j_{i_1},\sigma^j_{i_2}),(\mathsf{a}^j_{i_1},\mathsf{b}^j_{i_1}),(\mathsf{a}^j_{i_2},\mathsf{b}^j_{i_2}))$ (see (8.5)) that visit two distinct blocks $\mathcal{I}_{i_1}=[s_{i_1},t_{i_1}]\cap\mathbb{N}$ and $\mathcal{I}_{i_2}=[s_{i_2},t_{i_2}]\cap\mathbb{N}$ for some $i_1<i_2$, due to the contribution from the j-th renewal sequence for some $j\in\mathcal{J}_{i_1}\cap\mathcal{J}_{i_2}$. In particular, $\tau^j_{i_1}\in\mathcal{I}_{i_1}$ and $\sigma^j_{i_2}\in\mathcal{I}_{i_2}$. Recall from (8.5) and (4.11) that

$$\begin{aligned}&\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\tau^j_{i_1},\sigma^j_{i_2}),(\mathsf{a}^j_{i_1},\mathsf{b}^j_{i_1}),(\mathsf{a}^j_{i_2},\mathsf{b}^j_{i_2}))\\&=\frac{2}{\varepsilon N}\sum_{\substack{(d,x)\in\mathcal{B}_{\varepsilon N}(\tau^j_{i_1},\mathsf{a}^j_{i_1})\\(f',y')\in\mathcal{B}_{\varepsilon N}(\sigma^j_{i_2},\mathsf{b}^j_{i_2})}}\sum_{\substack{(f,y)\in\mathcal{B}_{\varepsilon N}(\tau^j_{i_1},\mathsf{b}^j_{i_1})\\(d',x')\in\mathcal{B}_{\varepsilon N}(\sigma^j_{i_2},\mathsf{a}^j_{i_2})\\d\leqslant f,\,d'\leqslant f'}}X^{(\mathrm{diff})}_{d,f}(x,y)\,q_{f,d'}(y,x')\,X^{(\mathrm{diff})}_{d',f'}(x',y').\end{aligned}\tag{8.6}$$

Note that by the definition of a U block, if $\mathcal{I}_{i_1}$ is a block of type U, then we must have $\tau^j_{i_1}=t_{i_1}$, the last time index in the block $\mathcal{I}_{i_1}$; while if $\mathcal{I}_{i_1}$ is a block of type V, then we must have $\tau^j_{i_1}\geqslant s_{i_1}\geqslant t_{i_1}-3$ because V blocks of length at most 4. Similarly, if $\mathcal{I}_{i_2}$ is of type of U, then we must have $\sigma^j_{i_2}=s_{i_2}$, the first time in the block $\mathcal{I}_{i_2}$; while if $\mathcal{I}_{i_2}$ is type V, then we must have $\sigma^j_{i_2}\leqslant t_{i_2}\leqslant s_{i_2}+3$. Therefore $d'-f\leqslant(s_{i_2}-t_{i_1}+7)\varepsilon N$. On the other hand, $\mathcal{I}_{i_1}$ and $\mathcal{I}_{i_2}$ are distinct blocks and hence $s_{i_2}-t_{i_1}\geqslant 2$ and $d'-f\geqslant\varepsilon N$. We can therefore apply Lemma 3.2 with $m=\varepsilon N$, $n_1=d'-f$, $n_2=s_{i_2}-t_{i_1}$ and $\varrho_1=10$ to bound

$$\sup_{\substack{(f,y)\in\mathcal{B}_{\varepsilon N}(\tau^j_{i_1},\mathsf{b}^j_{i_1})\\(d',x')\in\mathcal{B}_{\varepsilon N}(\sigma^j_{i_2},\mathsf{a}^j_{i_2})}}q_{f,d'}(y,x')\leqslant\frac{C}{\varepsilon N}\,g_{10(s_{i_2}-t_{i_1})}\,(\mathsf{a}^j_{i_2}-\mathsf{b}^j_{i_1}).\tag{8.7}$$

Applying this bound in (8.6) then allows us to make the replacement (recall the definition of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ from (4.11))

$$\begin{aligned}&\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\tau^j_{i_1},\sigma^j_{i_2}),(\mathsf{a}^j_{i_1},\mathsf{b}^j_{i_1}),(\mathsf{a}^j_{i_2},\mathsf{b}^j_{i_2}))\\&\rightsquigarrow C\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\tau^j_{i_1};\mathsf{a}^j_{i_1},\mathsf{b}^j_{i_1})\;g_{10(s_{i_2}-t_{i_1})}\,(\mathsf{a}^j_{i_2}-\mathsf{b}^j_{i_1})\;\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\sigma^j_{i_2};\mathsf{a}^j_{i_2},\mathsf{b}^j_{i_2}).\end{aligned}\tag{8.8}$$

Of course this is not an upper bound since $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ could be negative. However, when we compute the moment in (8.4), we end up with products of the moments of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$'s and its constituent ξ_N's, which are then be bounded by their absolute values. Applying (8.7) at this point gives a true upper bound, which has the same effect as making the replacement (8.8) in the expansion before taking expectation, and then compute the moment as in (8.4). To keep the notation simple, we will assume this replacement from now on, so that the expansion of (8.4) now contains only $\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\mathsf{i};\mathsf{a},\mathsf{b})$ that visits a single mesoscopic time interval $\mathcal{T}_{\varepsilon N}(\mathsf{i})$, which simplifies the expansion from (8.4). The cost is replacing some heat kernel $g_{\frac{\mathsf{i}}{2}}(\cdot)$ (more accurately, the associated random walk transition kernel) by $Cg_{10\mathsf{i}}(\cdot)$ as in (8.7).

(B) We now consider a U block $\mathcal{I} = [s,t] \cap \mathbb{N}$. Assuming w.l.o.g. $\mathcal{J} = \{1,2\}$ so that only renewal sequences 1 and 2 in the expansion of (8.4) visit block $\mathcal{I}$. Let $\mathsf{a}, \mathsf{b} \in \mathbb{Z}^2$ be spatial indices for the mesoscopic boxes of entry and exit in the time interval $\mathcal{I}$. Then the coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ visited by renewal sequences 1 and 2 in the time interval $\mathcal{I}$ must match perfectly in order to have non-zero expectation. Taking expectation in (8.4), each U block in the expansion therein leads to the following quantity analogous to $\overline{U}_N(n,x)$ defined in (3.40) and (3.41):

$$\overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(t-s, \mathsf{b}-\mathsf{a}) := \sum_{\substack{(\vec{\mathsf{i}}_1,\dots,\vec{\mathsf{i}}_r)\in \vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon} \\ (\vec{\mathsf{a}}_1,\dots,\vec{\mathsf{a}}_r)\in \vec{\boldsymbol{\mathcal{A}}}^{(\text{diff})}_{\varepsilon} \\ \mathsf{i}_1 = s, \mathsf{i}'_r = t, \mathsf{a}_1 = \mathsf{a}, \mathsf{a}'_r = \mathsf{b}}} \mathbb{E}[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1)^2] \prod_{j=2}^{r} g^2_{\frac{1}{2}(\mathsf{i}_j - \mathsf{i}'_{j-1})}(\mathsf{a}_j - \mathsf{a}'_{j-1}) \mathbb{E}[\Theta^{(\mathrm{cg})}_{N,\varepsilon}(\vec{\mathsf{i}}_j, \vec{\mathsf{a}}_j)^2], \quad (8.9)$$

where $\vec{\boldsymbol{\mathcal{A}}}^{(\text{no triple})}_{\varepsilon}$ and $\vec{\boldsymbol{\mathcal{A}}}^{(\text{diff})}_{\varepsilon}$ are defined in (5.70) and (5.71). Because the sum above is a sum over finitely many terms, by Lemma 7.1, the following limit exists

$$\overline{U}^{(\mathrm{cg})}_{\infty,\varepsilon}(t-s, \mathsf{b}-\mathsf{a}) := \lim_{N\to\infty} \overline{U}^{(\mathrm{cg})}_{N,\varepsilon}(t-s, \mathsf{b}-\mathsf{a}). \quad (8.10)$$

We next consider a V block $\mathcal{I} = [s,t] \cap \mathbb{N}$ with size $t-s+1 \leqslant 4$. Let $\mathcal{J} \subset \{1,2,3,4\}$ denote the set of renewal sequences that visit $\mathcal{I}$. To have non-zero expectation in the expansion of (8.4), we must have $|\mathcal{J}| \geqslant 3$ ($|\mathcal{J}| = 2$ would make it a U block instead). Each renewal sequence can visit at most two mesoscopic time intervals with indices in $\mathcal{I}$. For each renewal sequence $j \in \mathcal{J}$ that visits block $\mathcal{I}$, let (σ^j, a^j) and (τ^j, b^j) be the indices of the mesoscopic time-space boxes of entry and exit in $\mathcal{I}$. In the expectation in (8.4), such a V block then leads to the following factor

$$\begin{aligned} V^{(\mathrm{cg})}_{N,\varepsilon}((\sigma^j, \tau^j, \mathsf{a}^j, \mathsf{b}^j)_{j\in\mathcal{J}}) &:= \mathbb{E}\Big[\prod_{j\in\mathcal{J}} \Theta^{(\mathrm{cg})}_{N,\varepsilon}((\sigma^j, \tau^j), (\mathsf{a}^j, \mathsf{b}^j)) \Big] \\ &\leqslant \prod_{j\in\mathcal{J}} \mathbb{1}_{\{\exists\, j\neq k\in\mathcal{J}, |\mathsf{a}^j - \mathsf{a}^k| \leqslant 6M_\varepsilon\}} \mathbb{E}\Big[\Theta^{(\mathrm{cg})}_{N,\varepsilon}((\sigma^j, \tau^j), (\mathsf{a}^j, \mathsf{b}^j))^4 \Big]^{\frac{1}{4}}, \end{aligned}$$

where the indicators follow from the local dependence of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$ from its definition in (4.11). Applying (7.11) with $|\vec{\mathsf{i}}| \leqslant 4$ (since V blocks have length at most 4) then gives

$$\begin{aligned} \limsup_{N\to\infty} V^{(\mathrm{cg})}_{N,\varepsilon}((\sigma^j, \tau^j, \mathsf{a}^j, \mathsf{b}^j)_{j\in\mathcal{J}}) &\leqslant \frac{C}{(\log\frac{1}{\varepsilon})^{\frac{|\mathcal{J}|}{4}}} \prod_{j\in\mathcal{J}} \mathbb{1}_{\substack{\{\exists\, k\in\mathcal{J}, k\neq j, |\mathsf{a}^j - \mathsf{a}^k| \leqslant 6M_\varepsilon; \\ |\mathsf{b}^j - \mathsf{a}^j| \leqslant M_\varepsilon\}}} e^{-|\mathsf{b}^j - \mathsf{a}^j|} \\ &=: V^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}(\mathsf{a}, \mathsf{b}). \end{aligned} \quad (8.11)$$

(C) We next modify the heat kernels connecting different blocks $\mathcal{I}_i$. First, we will contract each V block $\mathcal{I}_i = [s_i, t_i] \cap \mathbb{N}$ into a block of size 1. Note that every heat kernel in the expansion connects two different blocks $\mathcal{I}_{i_1} = [s_{i_1}, t_{i_1}] \cap \mathbb{N}$ and $\mathcal{I}_{i_2} = [s_{i_2}, t_{i_2}] \cap \mathbb{N}$, $i_1 < i_2$, and are of the form $g_{\frac{1}{2}(\sigma_{i_2} - \tau_{i_1})}(\mathsf{a}_{i_2} - \mathsf{b}_{i_1})$ for some $\tau_{i_1} \in \mathcal{I}_{i_1}$ and $\sigma_{i_2} \in \mathcal{I}_{i_2}$ with $|\mathsf{a}_{i_2} - \mathsf{b}_{i_1}| \leqslant M_\varepsilon \sqrt{\sigma_{i_2} - \tau_{i_1}}$. The heat kernel $g_{\frac{1}{2}(\sigma_{i_2} - \tau_{i_1})}(\cdot)$ from time τ_{i_1} to σ_{i_2} may jump over multiple blocks of type V. If we contract the time span $[\tau_{i_1}, \sigma_{i_2}]$ of the heat kernel by shrinking each block of type V that intersects $[\tau_{i_1}, \sigma_{i_2}]$ into a block of size 1, and let u denote the length of the reduced time span for the heat kernel, then $u \geqslant \frac{1}{4}(\sigma_{i_2} - \tau_{i_1})$ since

blocks of type V have length at most 4. Therefore

$$g_{\frac{1}{2}(\sigma_{i_2}-\tau_{i_1})}(\mathsf{a}_{i_2}-\mathsf{b}_{i_1}) \leqslant 8 g_{2u}(\mathsf{a}_{i_2}-\mathsf{b}_{i_1}).$$

The heat kernels introduced in (8.7) are of the form $g_{10(\sigma_{i_2}-\tau_{i_1})}(\mathsf{a}_{i_2}-\mathsf{b}_{i_1})$ and can similarly be bounded by $8g_{80u}(\mathsf{a}_{i_2}-\mathsf{b}_{i_1})$. For consistency, we will further bound $g_{2u}(\cdot) \leqslant 20 g_{80u}(\cdot)$. This shows that at the cost of introducing a constant factor C for each block $\mathcal{I}_i$, we can assume from now on that all blocks $\mathcal{I}_i$ of type V have length 1, namely $s_i = t_i$, and all heat kernels appearing in the expansion (as an upper bound for the expansion of (8.4)) are of the form $g_{80(s_{i_2}-t_{s_{i_1}})}(\cdot)$.

Lastly, we perform a Chapman-Kolmogorov type decomposition for each heat kernel $g_{80(s_{i_2}-t_{i_1})}(\mathsf{b}_{i_1},\mathsf{a}_{i_2}) := g_{80(s_{i_2}-t_{i_1})}(\mathsf{a}_{i_2}-\mathsf{b}_{i_1})$ at each $s_j, t_j \in (t_{i_1}, s_{i_2}) \cap \mathbb{N}$, similar to what was done in the proof of Theorem 6.1. To simplify notation, let $u_0 = t_{i_1}, u_1, \dots, u_{k-1}, u_k = s_{i_2}$, with $u_1, \dots, u_{k-1}$ being the times at which we want to perform the decomposition. Let $\mathsf{x}_0 := \mathsf{b}_{i_1}$, $\mathsf{x}_k := \mathsf{a}_{i_2}$. Then we can bound

$$\begin{aligned} g_{80(u_k-u_0)}(\mathsf{x}_0,\mathsf{x}_k) &= \int\cdots\int_{x_1,\dots,x_{k-1}\in\mathbb{R}^2} g_{80(u_1-u_0)}(\mathsf{x}_0,x_1)\cdots g_{80(u_k-u_{k-1})}(x_{k-1},\mathsf{x}_k)\mathrm{d}x_1\cdots\mathrm{d}x_{k-1} \\ &\leqslant \sum_{\mathsf{x}_1,\dots,\mathsf{x}_{k-1}\in\mathbb{Z}^2}\prod_{i=1}^{k}\big(2g_{80(u_i-u_{i-1})}(\mathsf{x}_{i-1},\mathsf{x}_i)\big), \end{aligned}$$

where we have discretized the spatial integral into a sum over the lattice and introduced a factor 2 for each intermediate time u_i, $1 \leqslant i < k$, as a crude upper bound.

The steps **(A)-(C)** we have performed so far basically allow us to bound the expansion (8.4) in a form that is similar to (6.12), and ready to lead to the analogue of (6.17). The U blocks we have introduced correspond to U^I introduced in (6.14), while V blocks correspond to the disorder variable ξ, which even after contracting each V block into a block length 1, still has non-trivial spatial dependence. Due to the heavy notation, we will not write down the analogue of (6.12) here. Instead, we explain below how the analogues of (6.17) and (6.20) can be derived.

Part II. Bounds. Based on the considerations above, we can write down an upper bound for (8.4) that allows us to adapt the proof of Theorem 6.1. First, we introduce some notation that parallels those in the proof of Theorem 6.1. Similar to (6.8), for $\mathsf{a} = (\mathsf{a}^j)_{1\leqslant j\leqslant 4}, \mathsf{b} = (\mathsf{b}^j)_{1\leqslant j\leqslant 4} \in (\mathbb{Z}^2)^4$ and $\varphi : \mathbb{Z}^2 \to \mathbb{R}$, define

$$\mathcal{Q}_t(\mathsf{a},\mathsf{b}) := \prod_{j=1}^{4} g_{80t}(\mathsf{a}^j,\mathsf{b}^j),$$

$$\mathcal{Q}_t(\varphi,\mathsf{b}) := \prod_{j=1}^{4}\Big(\sum_{\mathsf{c}^j\in\mathbb{Z}^2}\varphi(\mathsf{c}^j)g_{80t}(\mathsf{b}^j-\mathsf{c}^j)\Big), \quad \mathcal{Q}_t(\mathsf{a},\psi) := \prod_{j=1}^{4}\Big(\sum_{\mathsf{c}^j\in\mathbb{Z}^2} g_{80t}(\mathsf{c}^j-\mathsf{a}^j)\psi(\mathsf{c}^j)\Big).$$

Similar to (6.14), for $\mathcal{J} = \{k,l\} \subset \{1,2,3,4\}$, define (recall $\overline{U}^{(\mathrm{cg})}_{\infty,\varepsilon}$ from (8.10))

$$\mathsf{U}^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}(t;\mathsf{a},\mathsf{b}) := \mathbb{1}_{\{\mathsf{a}^k=\mathsf{a}^l,\mathsf{b}^k=\mathsf{b}^l\}}\overline{U}^{(\mathrm{cg})}_{\infty,\varepsilon}(t,\mathsf{b}^k-\mathsf{a}^k)\prod_{j\notin\mathcal{J}} g_{80t}(\mathsf{a}^j,\mathsf{b}^j), \tag{8.12}$$

with $g_0(\mathsf{a}^j, \mathsf{b}^j) := \mathbb{1}_{\{\mathsf{a}^j = \mathsf{b}^j\}}$. Similarly, define (recall $\mathsf{V}^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}(\mathsf{a},\mathsf{b})$ from (8.11))

$$\mathsf{V}^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}(\mathsf{a},\mathsf{b}) := V^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}(\mathsf{a},\mathsf{b}) \prod_{j\notin\mathcal{J}} \mathbb{1}_{\{\mathsf{a}^j=\mathsf{b}^j\}}. \tag{8.13}$$

To be consistent with the notation in the proof of Theorem 6.1, we will replace $\mathcal{J} \subset \{1,2,3,4\}$, which determines which mesoscopic renewal sequences collect coarse-grained disorder variables at time t, by a partition $I \vdash \{1,2,3,4\}$, which specifies which sequences interact with each other through the coarse-grained disorder variables at time t. In particular, in $\mathsf{U}^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}$ corresponding to a U block, the associated partition I consists of $\mathcal{J}$ and $\{j\}$ for $j \notin \mathcal{J}$, so that $|I| = 3$. In $V^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon}$ corresponding to a V block, if $|\mathcal{J}| = 3$, then the associated partition I consists of $\mathcal{J}$ and $\{j\}$ for $j \notin \mathcal{J}$, so $|I| = 2$; if $|\mathcal{J}| = 4$, then the associated partition I is given by the connected components of $\{1,2,3,4\}$ with an edge between i and j whenever $|\mathsf{a}^i - \mathsf{a}^j| \leqslant 6M_\varepsilon$, and there can be no singletons in the partition to ensure that $\mathsf{V}^{(\mathrm{cg}),\mathcal{J}}_{\infty,\varepsilon} \neq 0$ (in particular, $|I| = 1$ or 2).

From now on, we write $\mathsf{U}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(t;\mathsf{a},\mathsf{b})$ and $\mathsf{V}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(t;\mathsf{a},\mathsf{b})$, replacing $\mathcal{J}$ by the associated partition $I \vdash \{1,2,3,4\}$. Define

$$\mathsf{W}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(t,\mathsf{a},\mathsf{b}) := \mathbb{1}_{\{|I|=3\}}\mathsf{U}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(t,\mathsf{a},\mathsf{b}) + \mathbb{1}_{\{|\mathrm{I}|<3\}}\mathsf{V}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(\mathsf{a},\mathsf{b}). \tag{8.14}$$

We can now write down the following upper bound for (8.4) in the limit $N \to \infty$:

$$\begin{aligned}\mathcal{M}^{\varphi,\psi}_{\infty,\varepsilon} := \limsup_{N\to\infty} \mathcal{M}^{\varphi,\psi}_{N,\varepsilon} \leqslant \|\psi\|^4_\infty\, \varepsilon^5 \sum_{r=1}^{\infty} C^r \sum_{\substack{I_1,\dots,I_r\vdash\{1,2,3,4\},|I_i|\leqslant 3\\ K_\varepsilon<s_1\leqslant t_1<s_2\leqslant\cdots<s_r\leqslant t_r<s_{r+1}\leqslant\frac{2}{\varepsilon}\\ \mathsf{a}_1,\mathsf{b}_1,\dots,\mathsf{a}_r,\mathsf{b}_r\in(\mathbb{Z}^2)^4}} \\ \mathcal{Q}_{80s_1}(\varphi_\varepsilon,\mathsf{a}_1)\prod_{i=1}^{r}\mathsf{W}^{(\mathrm{cg}),I_i}_{\infty,\varepsilon}(t_i-s_i,\mathsf{a}_i,\mathsf{b}_i)\mathcal{Q}_{80(s_{i+1}-t_i)}(\mathsf{b}_i,\mathsf{a}_{i+1}),\end{aligned} \tag{8.15}$$

where C does not depend on ε, $\mathsf{a}_{r+1} := \mathbb{1}_{B_\varepsilon}$, the sum in (8.15) contains no consecutive $I_i = I_{i+1}$ with $|I_i| = 3$, and when $|I_i| \leqslant 2$, we must have $s_i = t_i$ thanks to the contraction of the V blocks.

For $\lambda > 0$ to be chosen later, we can insert the factor $e^{\frac{2\lambda}{\varepsilon}} e^{-\lambda\sum_{i=1}^{r+1}(s_i-t_{i-1})-\lambda\sum_{i=1}^{r}(t_i-s_i)} \geqslant 1$ into (8.15) to obtain a bound similar to (6.17):

$$\left|\mathcal{M}^{\varphi,\psi}_{\infty,\varepsilon}\right| \leqslant e^{\frac{2\lambda}{\varepsilon}}\|\psi\|^4_\infty\varepsilon^5\sum_{r=1}^{\infty}C^r\sum_{I_1,\dots,I_r}\langle\varphi_\varepsilon^{\otimes 4},\mathsf{P}^{*;I_1}_{\lambda,\varepsilon}\mathsf{P}^{I_1;I_2}_{\lambda,\varepsilon}\cdots\mathsf{P}^{I_{r-1};I_r}_{\lambda,\varepsilon}\mathcal{Q}^{I_r;*}_{\lambda,\varepsilon}\mathbb{1}^{\otimes 4}_{B_\varepsilon}\rangle, \tag{8.16}$$

where given two partitions $I, J \vdash \{1,\dots,4\}$, with $I = *$ denoting the partition consisting of singletons, $\mathsf{P}^{I,J}_{\lambda,\varepsilon}$ are integral operators with kernels given by

$$\mathsf{P}^{I;J}_{\lambda,\varepsilon} := \begin{cases}\mathcal{Q}^{I;J}_{\lambda,\varepsilon}\mathsf{V}^{J}_{\lambda,\varepsilon} & \text{if } |J|<3,\\ \mathcal{Q}^{I;J}_{\lambda,\varepsilon}\mathsf{U}^{J}_{\lambda,\varepsilon} & \text{if } |J|=3,\end{cases} \tag{8.17}$$

where for $\mathsf{a}, \mathsf{b} \in (\mathbb{Z}^2)^4$,

$$\begin{aligned}
\mathcal{Q}^{I;J}_{\lambda,\varepsilon}(\mathsf{b},\mathsf{a}) &:= \mathbb{1}_{\{\mathsf{b}\sim I, \mathsf{a}\sim J\}} \sum_{n=1}^{2/\varepsilon} e^{-\lambda n} \mathcal{Q}_{80n}(\mathsf{b},\mathsf{a}), \\
\mathsf{U}^{I}_{\lambda,\varepsilon}(\mathsf{a},\mathsf{b}) &:= \mathbb{1}_{\{\mathsf{a},\mathsf{b}\sim I\}} \sum_{n=0}^{2/\varepsilon} e^{-\lambda n} \mathsf{U}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(n,\mathsf{a},\mathsf{b}), && |I| = 3, \\
\mathsf{V}^{I}_{\lambda,\varepsilon}(\mathsf{a},\mathsf{b}) &:= \mathbb{1}_{\{\mathsf{a},\mathsf{b}\sim I\}} \mathsf{V}^{(\mathrm{cg}),I}_{\infty,\varepsilon}(\mathsf{a},\mathsf{b}), && |I| < 3.
\end{aligned} \tag{8.18}$$

Here, given $\mathsf{a} = (\mathsf{a}^j)_{1\leqslant j\leqslant 4} \in (\mathbb{Z}^2)^4$ and a partition $I \vdash \{1,2,3,4\}$, with $k \overset{I}{\sim} l$ denoting k, l belonging to the same partition element in I,

$$\mathsf{a} \sim I \quad \text{denotes the constraint} \quad \begin{cases} \forall\, k \overset{I}{\sim} l,\ \mathsf{a}^k = \mathsf{a}^l & \text{if } |I| = 3, \\ \forall\, k \overset{I}{\sim} l,\ |\mathsf{a}^k - \mathsf{a}^l| \leqslant 20 M_\varepsilon & \text{if } |I| \leqslant 2. \end{cases} \tag{8.19}$$

We will denote $(\mathbb{Z}^2)^4_I := \{\mathsf{b} \in (\mathbb{Z}^2)^4 : \mathsf{b} \sim I\}$. The main difference from (6.17) and (6.15) is that the spatial constraint there are delta functions (see (6.7)), that is, $\mathsf{a} \sim I$ with M_ε set to 0. Here we also have the additional operator $\mathsf{V}^J_{\lambda,\varepsilon}(\mathsf{a},\mathsf{b})$ because we allow $\mathsf{b} \neq \mathsf{a}$. The analogue of $\mathbb{E}[\xi^p]$ are the moments of the coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$, which are now captured in $\mathsf{V}^J_{\lambda,\varepsilon}$ and $\mathsf{U}^J_{\lambda,\varepsilon}$.

As in (6.18), for a weight function $w_\varepsilon : \mathbb{Z}^2 \to \mathbb{R}$, see (4.7), we define the weighted operators

$$\widehat{\mathcal{Q}}^{I;J}_{\lambda,\varepsilon}(\mathsf{b},\mathsf{a}) := \frac{w_\varepsilon^{\otimes 4}(\mathsf{b})}{w_\varepsilon^{\otimes 4}(\mathsf{a})}\, \mathcal{Q}^{I;J}_{\lambda,\varepsilon}(\mathsf{b},\mathsf{a}),$$

and define $\widehat{\mathsf{U}}^J_{\lambda,\varepsilon}(\mathsf{a},\mathsf{b})$, $\widehat{\mathsf{V}}^J_{\lambda,\varepsilon}(\mathsf{a},\mathsf{b})$, and $\widehat{\mathsf{P}}^{I,J}_{\lambda,\varepsilon}$ similarly. For $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, we can then bound (8.16) via the following analogue of (6.20):

$$\begin{aligned}
\big|\mathcal{M}^{\varphi,\psi}_{\infty,\varepsilon}\big| \leqslant e^{\frac{2\lambda}{\varepsilon}} \|\psi\|^4_\infty \varepsilon^5 \sum_{r=1}^{\infty} C^r \sum_{I_1,\dots,I_r} \Big\| \frac{\varphi_\varepsilon^{\otimes 4}}{w_\varepsilon^{\otimes 4}} \Big\|_{\ell^p((\mathbb{Z}^2)^4_{I_1})} \Big\| \widehat{\mathsf{P}}^{*,I_1}_{\lambda,\varepsilon} \Big\|_{\ell^q\to\ell^q} \Big\| \widehat{\mathsf{P}}^{I_1;I_2}_{\lambda,\varepsilon} \Big\|_{\ell^q\to\ell^q} \cdots \\
\cdots \Big\| \widehat{\mathsf{P}}^{I_{r-1};I_r}_{\lambda,\varepsilon} \Big\|_{\ell^q\to\ell^q} \Big\| \widehat{\mathcal{Q}}^{I_r;*}_{\lambda,\varepsilon} \Big\|_{\ell^q\to\ell^q} \Big\| \mathbb{1}^{\otimes 4}_{B_\varepsilon} w_\varepsilon^{\otimes 4} \Big\|_{\ell^q((\mathbb{Z}^2)^4_{I_r})},
\end{aligned} \tag{8.20}$$

where we still have $(\mathbb{Z}^2)^4_I := \{\mathsf{b} \in (\mathbb{Z}^2)^4 : \mathsf{b} \sim I\}$, but the definition of the constraint $\mathsf{b} \sim I$ has changed as in (8.19). We still regard $\widehat{\mathcal{Q}}^{I;J}_{\lambda,\varepsilon}(\cdot,\cdot)$ and $\widehat{\mathsf{P}}^{I;J}_{\lambda,\varepsilon}(\cdot,\cdot)$ as operators from $\ell^q((\mathbb{Z}^2)^4_J) \to \ell^q((\mathbb{Z}^2)^4_I)$, and similarly for $\widehat{\mathsf{U}}^J_{\lambda,\varepsilon}$ and $\widehat{\mathsf{V}}^J_{\lambda,\varepsilon}$.

We choose $\lambda := \hat{\lambda}\varepsilon$ with $\hat{\lambda}$ large but fixed so that $e^{\lambda\varepsilon}$ remains bounded. We have the following analogue of Proposition 6.6, where we again omitted $(\mathbb{Z}^2)^4_I$ from $\|\cdot\|_{\ell^p(\cdot)}$.

Proposition 8.2. *For some c uniformly in $\lambda = \frac{\hat\lambda}{N} > 0$, $\varepsilon \in (0,1)$, and $I, J \subset \{1,\dots,4\}$ with $1 \leqslant |I|, |J| \leqslant 3$ and $I \neq J$ when $|I| = |J| = 3$, we have*

$$\big\|\widehat{\mathcal{Q}}^{I;J}_{\lambda,\varepsilon}\big\|_{\ell^q\to\ell^q} \leqslant c\mathbb{1}_{\{|I|=|J|=3\}} + cM_\varepsilon^8 \mathbb{1}_{\{|I|\wedge|J|\leqslant 2\}}; \tag{8.21}$$

$$\big\|\widehat{\mathcal{Q}}^{*,I}_{\lambda,\varepsilon}\big\|_{\ell^q\to\ell^q} \leqslant c\varepsilon^{-\frac{1}{q}}\big(\mathbb{1}_{\{|I|=3\}} + \mathbb{1}_{\{|I|\leqslant 2\}} M_\varepsilon^4\big); \tag{8.22}$$

$$\big\|\widehat{\mathcal{Q}}^{I;*}_{\lambda,\varepsilon}\big\|_{\ell^q} \leqslant c\varepsilon^{-\frac{1}{p}}\big(\mathbb{1}_{\{|I|=3\}} + \mathbb{1}_{\{|I|\leqslant 2\}} M_\varepsilon^4\big); \tag{8.23}$$

$$\big\|\widehat{\mathsf{U}}^{I}_{\hat\lambda\varepsilon,\varepsilon}\big\|_{\ell^q\to\ell^q} \leqslant \frac{c}{\log\hat\lambda} \quad \text{for } |I| = 3; \tag{8.24}$$

$$\big\|\widehat{\mathsf{V}}^{I}_{\lambda,\varepsilon}\big\|_{\ell^q\to\ell^q} \leqslant \frac{c}{(\log\frac{1}{\varepsilon})^{\frac{3}{4}}} \quad \text{for } |I| \leqslant 2. \tag{8.25}$$

We now substitute these bounds into (8.20) and note that when $|I| = |J| = 3$, each factor $\|\widehat{\mathcal{Q}}^{I;J}_{\lambda,\varepsilon}\|_{\ell^q\to\ell^q}$ can be controlled by $\|\widehat{\mathsf{U}}^{I}_{\hat\lambda\varepsilon,\varepsilon}\|_{\ell^q\to\ell^q}$, and when $|I|\wedge|J| \leqslant 2$, the powers of $M_\varepsilon = \log\log\frac{1}{\varepsilon}$ from (8.21)-(8.23) can be controlled by $\|\widehat{\mathsf{V}}^{I}_{\lambda,\varepsilon}\|^{1/2}_{\ell^q\to\ell^q}\|\widehat{\mathsf{V}}^{J}_{\lambda,\varepsilon}\|^{1/2}_{\ell^q\to\ell^q}$ (set $\|\widehat{\mathsf{V}}^{I}_{\lambda,\varepsilon}\|_{\ell^q\to\ell^q} := 1$ if $|I| = 3$). This leads to a convergent geometric series similar to (6.25), which gives

$$\big|\mathcal{M}^{\varphi,\psi}_{\infty,\varepsilon}\big| \leqslant C\varepsilon^{\frac{4}{p}}\Big\|\frac{\varphi_\varepsilon}{w_\varepsilon}\Big\|^4_{\ell^p}\|\psi\|^4_\infty\|w\mathbb{1}_B\|^4_{\ell^q} \tag{8.26}$$

for some C depending only on $\hat\lambda$. This proves Theorem 8.1. □

To conclude this section, we sketch the proof of Proposition 8.2.

Proof of Proposition 8.2. We will sketch how the proof of Proposition 6.6 can be adapted to the current setting.

Proof of (8.21). First note that it is equivalent to

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} f(\boldsymbol{x})\mathcal{Q}^{I,J}_{\lambda,\varepsilon}(\boldsymbol{x},\boldsymbol{y})\frac{w^{\otimes 4}_\varepsilon(\boldsymbol{x})}{w^{\otimes 4}_\varepsilon(\boldsymbol{y})}\,g(\boldsymbol{y}) \leqslant c\big(\mathbb{1}_{\{|I|=|J|=3\}} + M_\varepsilon^8\mathbb{1}_{\{|I|\wedge|J|<3\}}\big)\|f\|_{\ell^p}\|g\|_{\ell^q} \tag{8.27}$$

uniformly for all $f \in \ell^p((\mathbb{Z}^2)^4_I)$ and $g \in \ell^q((\mathbb{Z}^2)^4_J)$. We split the region of summation into $A_\varepsilon = \{|\boldsymbol{x}-\boldsymbol{y}| \leqslant C_0/\sqrt{\varepsilon|}\}$ and A_ε^c. Note that the analogue of Lemma 6.7 holds for $\mathcal{Q}_{\lambda,\varepsilon}$. Therefore following the same argument as in (6.35), the region A_ε^c gives the contribution

$$\begin{aligned}
&\sum_{\substack{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J\\(\boldsymbol{x},\boldsymbol{y})\in A_\varepsilon^c}} f(\boldsymbol{x})\mathcal{Q}^{I,J}_{\lambda,\varepsilon}(\boldsymbol{x},\boldsymbol{y})\frac{w^{\otimes 4}_\varepsilon(\boldsymbol{x})}{w^{\otimes 4}_\varepsilon(\boldsymbol{y})}\,g(\boldsymbol{y})\\
&\leqslant C\varepsilon^3\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J}|f(\boldsymbol{x})|^p\,e^{-|\boldsymbol{x}-\boldsymbol{y}|\sqrt{\varepsilon}}\Big)^{1/p}\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J}|g(\boldsymbol{y})|^q\,e^{-|\boldsymbol{x}-\boldsymbol{y}|\sqrt{\varepsilon}}\Big)^{1/q}\\
&\leqslant C\varepsilon^{3-\frac{|J|}{p}-\frac{|I|}{q}}\big(\mathbb{1}_{\{|I|=|J|=3\}} + M_\varepsilon^8\mathbb{1}_{\{|I|\wedge|J|<3\}}\big)\|f\|_{\ell^p}\,\|g\|_{\ell^q} \qquad (8.28)\\
&\leqslant C\|f\|_{\ell^p}\,\|g\|_{\ell^q},
\end{aligned}$$

where the spatial constraints in $\boldsymbol{x} \in (\mathbb{Z}^2)^4_I$ and $\boldsymbol{y} \in (\mathbb{Z}^2)^4_J$ (see (8.19)) led to the factor in the bracket in the third line.

In the region A_ε, the factor $w_\varepsilon^{\otimes 4}(\boldsymbol{x})/w_\varepsilon^{\otimes 4}(\boldsymbol{y})$ is bounded. By the analogue of Lemma 6.7 for $\mathcal{Q}_{\lambda,\varepsilon}$, it suffices to show

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} \frac{f(\boldsymbol{x})g(\boldsymbol{y})}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3} \leqslant c\big(\mathbb{1}_{\{|I|=|J|=3\}}+M_\varepsilon^8\mathbb{1}_{\{|I|\wedge|J|<3\}}\big)\|f\|_{\ell^p}\|g\|_{\ell^q} \tag{8.29}$$

When $|I|=|J|=3$, the proof is exactly the same as that of (6.27). When $|I|,|J|<3$, we can apply Hölder to bound the l.h.s. of (8.29) by

$$\begin{aligned}&\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} \frac{f(\boldsymbol{x})^p}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3}\Big)^{1/p}\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} \frac{g(\boldsymbol{y})^q}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3}\Big)^{1/q}\\ &\leqslant CM_\varepsilon^8\|f\|_{\ell^p}\|g\|_{\ell^q}.\end{aligned} \tag{8.30}$$

When $|I|<3$ and $|J|=3$ (the case $|I|=3$ and $|J|<3$ can be treated identically), we can find $k,l\in\{1,2,3,4\}$ that belong to the same partition element in I, but to different partition elements in J; in particular, $\boldsymbol{x}\in(\mathbb{Z}^2)^4_I$ implies $|x_k-x_l|\leqslant 20M_\varepsilon$. Fix any $a\in(0,1/q)$. We can then apply Hölder to bound the l.h.s. of (8.29) by

$$\begin{aligned}&\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} \frac{f(\boldsymbol{x})^p}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3}\cdot\frac{1}{(1+|y_k-y_l|^{2a})^p}\Big)^{1/p}\\ &\times\Big(\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4_J} \frac{g(\boldsymbol{y})^q}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3}\cdot(1+|y_k-y_l|^{2a})^q\Big)^{1/q} \leqslant CM_\varepsilon^8\|f\|_{\ell^p}\|g\|_{\ell^q},\end{aligned}$$

where in the first bracket, the sum over $\boldsymbol{y}$ is uniformly bounded by the same argument as in the bound for (6.28), while in the second bracket, we can distinguish between two cases: either $|y_k-y_l|\leqslant 40M_\varepsilon$, in which case we apply this bound and sum over $\boldsymbol{x}$ to get a bound of $M_\varepsilon^{\frac{4}{q}+2a}\|g\|_{\ell^q}$; or $|y_k-y_l|>40M_\varepsilon\geqslant 2|x_k-x_l|$, in which case we apply the triangle inequality

$$|x_k-y_k|+|x_l-y_l|\geqslant\frac{|x_k-y_k|+|x_l-y_l|}{2}+\frac{|y_k-y_l|}{4}$$

and follow the same argument as for (6.30) to get a bound of $M_\varepsilon^{4/q}\|g\|_{\ell^q}$, where $M_\varepsilon^{4/q}$ comes from summing over the redundant components of $\boldsymbol{x}$ after selecting one component of $\boldsymbol{x}$ for each partition element of I. This concludes the proof of (8.21).

Proof of (8.22)-(8.23). Similar to (6.36), we need to show

$$\sum_{\boldsymbol{x}\in(\mathbb{Z}^2)^4_I,\boldsymbol{y}\in(\mathbb{Z}^2)^4} f(\boldsymbol{x})\mathcal{Q}^{I,*}_{\lambda,\varepsilon}(\boldsymbol{x},\boldsymbol{y})\frac{w_\varepsilon^{\otimes 4}(\boldsymbol{x})}{w_\varepsilon^{\otimes 4}(\boldsymbol{y})}g(\boldsymbol{y}) \leqslant c\,\varepsilon^{-\frac{1}{p}}\big(\mathbb{1}_{\{|I|=3\}}+\mathbb{1}_{\{|I|<3\}}M_\varepsilon^4\big)\|f\|_{\ell^p}\|g\|_{\ell^q} \tag{8.31}$$

uniformly in $f\in\ell^p((\mathbb{Z}^2)^4_I)$ and $g\in\ell^q((\mathbb{Z}^2)^4)$. Restricted to $(\boldsymbol{x},\boldsymbol{y})\in A_\varepsilon^c$, we note that the bound (8.28) can now be replaced by

$$C\varepsilon^{3-\frac{4}{p}}\Big(\mathbb{1}_{\{|I|=3\}}\varepsilon^{-\frac{3}{q}}+\mathbb{1}_{\{|I|<3\}}M_\varepsilon^{\frac{8-2|I|}{q}}\varepsilon^{-\frac{|I|}{q}}\Big)\|f\|_{\ell^p}\,\|g\|_{\ell^q}\leqslant C\varepsilon^{-\frac{1}{p}}\|f\|_{\ell^p}\,\|g\|_{\ell^q}.$$

Restricted to A_ε, it suffices to bound the following analogue of (6.38):

$$\sum_{\substack{\boldsymbol{x}\in(\mathbb{Z}^4)^h_I,\,\boldsymbol{y}\in(\mathbb{Z}^2)^4\\(\boldsymbol{x},\boldsymbol{y})\in A_\varepsilon}} \frac{f(\boldsymbol{x})g(\boldsymbol{y})}{\big(1+\sum_{i=1}^4|x_i-y_i|^2\big)^3} \leqslant C\varepsilon^{-\frac{1}{p}}\big(\mathbb{1}_{\{|I|=3\}}+\mathbb{1}_{\{|I|<3\}}M_\varepsilon^4\big)\|f\|_{\ell^p}\|g\|_{\ell^q}. \tag{8.32}$$

When $|I| = 3$, this follows by exactly the same proof as that of (6.38). When $|I| \leqslant 2$, the estimate is simpler and we can apply the Hölder inequality to bound the l.h.s. of (8.32) by

$$\Bigg(\sum_{\substack{\boldsymbol{x} \in (\mathbb{Z}^2)^4_I, \boldsymbol{y} \in (\mathbb{Z}^2)^4 \\ (\boldsymbol{x}, \boldsymbol{y}) \in A_\varepsilon}} \frac{f(\boldsymbol{x})^p}{\big(1 + \sum_{i=1}^4 |x_i - y_i|^2\big)^3} \Bigg)^{\frac{1}{p}} \Bigg(\sum_{\substack{\boldsymbol{x} \in (\mathbb{Z}^2)^4_I, \boldsymbol{y} \in (\mathbb{Z}^2)^4 \\ (\boldsymbol{x}, \boldsymbol{y}) \in A_\varepsilon}} \frac{g(\boldsymbol{y})^q}{\big(1 + \sum_{i=1}^4 |x_i - y_i|^2\big)^3} \Bigg)^{\frac{1}{q}}. \quad (8.33)$$

Since $|I| \leqslant 2$, the second factor can be bounded by $C M_\varepsilon^4 \|g\|_{\ell^q}$. For the first factor, summing over $\boldsymbol{y}$ gives

$$\sum_{\substack{\boldsymbol{y} \in (\mathbb{Z}^2)^4 \\ |\boldsymbol{y} - \boldsymbol{x}| < C_0/\sqrt{\varepsilon}}} \frac{1}{\big(1 + \sum_{i=1}^4 |x_i - y_i|^2\big)^3} \leqslant C \varepsilon^{-1}$$

uniformly in $\boldsymbol{x} \in (\mathbb{Z}^2)^4_I$. Collecting all the bounds obtained so far then gives (8.31).

Proof of (8.24). Assume w.l.o.g. that $I \vdash \{1,2,3,4\}$ consists of the partition elements $\{1,2\}, \{3\}, \{4\}$. Recall from (8.18) and (8.12) that for $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{Z}^2)^4_I$,

$$\mathsf{U}^I_{\hat{\lambda}\varepsilon, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon n} \, \overline{U}^{(\mathrm{cg})}_{\infty, \varepsilon}(n, y_1 - x_1) \prod_{i=3,4} g_{80n}(x_i, y_i).$$

We then have the following analogues of (6.43) and (6.44):

$$\sum_{\boldsymbol{y} \in (\mathbb{Z}^2)^4_I} \mathsf{U}^I_{\hat{\lambda}\varepsilon, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) \leqslant 4 \sum_{n=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon n} \sum_{z \in \mathbb{Z}^2} \overline{U}^{(\mathrm{cg})}_{\infty, \varepsilon}(n, z) \leqslant \frac{C}{\log \hat{\lambda}},$$

$$\sum_{\boldsymbol{y} \in (\mathbb{Z}^2)^4_I} \mathsf{U}^I_{\hat{\lambda}\varepsilon, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) e^{C|\boldsymbol{x} - \boldsymbol{y}|\sqrt{\varepsilon}} \leqslant C \sum_{n=0}^{2/\varepsilon} e^{-\hat{\lambda}\varepsilon n} \sum_{z \in \mathbb{Z}^2} \overline{U}^{(\mathrm{cg})}_{\infty, \varepsilon}(n, z) e^{C|z|\sqrt{\varepsilon}} \leqslant \frac{C}{\log \hat{\lambda}},$$

both of which follows from Lemma 5.9 by Fatou's Lemma (recall $\overline{U}^{(\mathrm{cg})}_{\infty, \varepsilon}$ from (8.10)). The rest of the proof is exactly the same as that of (6.24).

Proof of (8.25). Given a partition $I \vdash \{1,2,3,4\}$ with $|I| \leqslant 2$, recall the definition of $\widehat{\mathsf{V}}^I_{\lambda, \varepsilon}$ from (8.18). We need to show that

$$\sum_{\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{Z}^2)^4_I} f(\boldsymbol{x}) \mathsf{V}^{(\mathrm{cg}), I}_{\infty, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) g(\boldsymbol{y}) \frac{w_\varepsilon^{\otimes 4}(\boldsymbol{x})}{w_\varepsilon^{\otimes 4}(\boldsymbol{y})} \leqslant \frac{C}{(\log \frac{1}{\varepsilon})^{\frac{3}{4}}} \|f\|_{\ell^p} \|g\|_{\ell^q} \quad (8.34)$$

uniformly for all $f \in \ell^p((\mathbb{Z}^2)^4_I)$ and $g \in \ell^q((\mathbb{Z}^2)^4_I)$. As before, we consider the sum of $\boldsymbol{x}, \boldsymbol{y}$ over $A_\varepsilon = \{|\boldsymbol{x} - \boldsymbol{y}| \leqslant C_0/\sqrt{\varepsilon|}\}$ and A_ε^c separately and apply Hölder's inequality. The bound (8.25) will follow if we show that uniformly in $\boldsymbol{x} \in (\mathbb{Z}^2)^4_I$,

$$\sum_{\boldsymbol{y} \in (\mathbb{Z}^2)^4_I} \mathsf{V}^{(\mathrm{cg}), I}_{\infty, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \frac{C}{(\log \frac{1}{\varepsilon})^{\frac{3}{4}}} \quad \text{and} \quad \sum_{\boldsymbol{y} \in (\mathbb{Z}^2)^4_I} \mathsf{V}^{(\mathrm{cg}), I}_{\infty, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) e^{C|\boldsymbol{y} - \boldsymbol{x}|\sqrt{\varepsilon}} \leqslant \frac{C}{(\log \frac{1}{\varepsilon})^{\frac{3}{4}}}. \quad (8.35)$$

These bounds hold because (8.13) and (8.11) imply that

$$\mathsf{V}^{(\mathrm{cg}), I}_{\infty, \varepsilon}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \frac{C}{(\log \frac{1}{\varepsilon})^{\frac{3}{4}}} \prod_{i=1}^4 \mathbb{1}_{\{|y_i - x_i| \leqslant M_\varepsilon\}} e^{-|y_i - x_i|}.$$

This concludes the proof of Proposition 8.2. □

9. Proof of the main results: Theorems 1.1 and 1.2

9.1. Proof of Theorem 1.1.

We can rephrase Theorem 1.1 as follows.

Theorem 9.1. *Assume the same setup as in Theorem 1.1. Let $k \in \mathbb{N}$. For $i = 1, \dots, k$, assume $0 \leqslant s_i \leqslant t_i < \infty$, $\varphi_i \in C_c(\mathbb{R}^2)$ has compact support, and $\psi_i \in C_b(\mathbb{R}^2)$ is bounded. Then the following convergence in distribution holds as $N \to \infty$:*

$$\big(\mathscr{Z}^{\beta_N}_{N,s_i,t_i}(\varphi_i,\psi_i)\big)_{i=1,\dots,k} \implies \big(\mathscr{Z}^{\vartheta}_{s,t}(\varphi_i,\psi_i)\big)_{i=1,\dots,k}\,, \tag{9.1}$$

where $\mathscr{Z}^{\beta_N}_{N;\,s,t}(\varphi,\psi) := \iint \varphi(x)\,\psi(y)\,\mathscr{Z}^{\beta_N}_{N;\,s,t}(\mathrm{d}x,\mathrm{d}y)$.

We will prove Theorem 9.1 by showing that the random vector in the l.h.s. of (9.1) converges in distribution as $N \to \infty$ to a unique random limit. This in turn implies that $(\mathscr{Z}^{\beta_N}_{N;\,s,t}(\mathrm{d}x,\mathrm{d}y))_{0\leqslant s\leqslant t<\infty}$ converges to a unique limit, denoted $\mathscr{Z}^{\vartheta} = (\mathscr{Z}^{\vartheta}_{s,t}(\mathrm{d}x,\mathrm{d}y))_{0\leqslant s\leqslant t<\infty}$.

The convergence of the one point distribution in Theorem 9.1 follows from the following result. We will explain how this can be adapted to finite dimensional distributions in Remark 9.5.

Proposition 9.2. *Given $\varphi \in C_c(\mathbb{R}^2)$ and $\psi \in C_b(\mathbb{R}^2)$, let $\mathcal{Z}_N(\varphi,\psi) := \mathscr{Z}^{\beta_N}_{N,0,1}(\varphi,\psi) = \iint \varphi(x)\psi(y)\mathscr{Z}^{\beta_N}_N(\mathrm{d}x,\mathrm{d}y)$ be as in Theorem 9.1 with $N \in 2\mathbb{N}$. Then $\mathcal{Z}_N(\varphi,\psi)$ converges in distribution to a unique limit as $N \to \infty$.*

Proof. Since

$$\mathrm{E}[|\mathcal{Z}_N(\varphi,\psi)|] \leqslant \frac{1}{N}\sum_{x,y\in\mathbb{Z}^2} \big|\varphi\big(\tfrac{x}{\sqrt{N}}\big)\big|\big|\psi\big(\tfrac{y}{\sqrt{N}}\big)\big| q_N(y-x) \to \iint |\varphi(x)||\psi(y)| g_{\frac12}(y-x)\mathrm{d}x\mathrm{d}y\,,$$

it follows that $\mathrm{E}[|\mathcal{Z}_N(\varphi,\psi)|]$ is uniformly bounded in N and hence $(\mathcal{Z}_N(\varphi,\psi))_{N\in\mathbb{N}}$ is a tight family and admits subsequential weak limits.

To show that the limit is unique, it then suffices to show that for every bounded $f : \mathbb{R} \to \mathbb{R}$, with uniformly bounded first three derivatives, the limit

$$\lim_{N\to\infty} \mathrm{E}[f(\mathcal{Z}_N(\varphi,\psi))]$$

exists. To this end, we will show that $(\mathrm{E}[f(\mathcal{Z}_N(\varphi,\psi))])_{N\in\mathbb{N}}$ is a Cauchy sequence.

Theorem 4.7 allows us to approximate $\mathcal{Z}_N(\varphi,\psi)$ by the coarse-grained model $\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta)$ with coarse-grained disorder variables $\Theta = \Theta^{(\mathrm{cg})}_{N,\varepsilon}$, with an L^2 error that is arbitrarily small, uniformly in large N, if $\varepsilon > 0$ is chosen sufficiently small. Therefore it only remains to show

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}\sup_{m,n\geqslant N}\big|\mathrm{E}[f(\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta^{(\mathrm{cg})}_{m,\varepsilon}))] - \mathrm{E}[f(\mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta^{(\mathrm{cg})}_{n,\varepsilon}))]\big| = 0. \tag{9.2}$$

We will prove (9.2) by applying the Lindeberg principle for multilinear polynomials of dependent random variables formulated in Lemmas A.2–A.4.

Let us set $\Phi(\Theta) := \mathscr{Z}^{(\mathrm{cg})}_\varepsilon(\varphi,\psi|\Theta)$, and note from its definition in (4.8) that $\Phi(\Theta)$ is a multilinear polynomial in the variables $\Theta := (\Theta(\vec{\mathsf{i}},\vec{\mathsf{a}}))_{(\vec{\mathsf{i}},\vec{\mathsf{a}})\in\mathbb{T}_\varepsilon}$, where recall from (4.5) that

$$\mathbb{T}_\varepsilon := \Big\{(\vec{\mathsf{i}},\vec{\mathsf{a}}) = ((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}')) :\ |\vec{\mathsf{i}}| = \mathsf{i}'-\mathsf{i}+1 \leqslant K_\varepsilon,\ |\vec{\mathsf{a}}| = |\mathsf{a}-\mathsf{a}'| \leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}\Big\}.$$

We write Θ_n for the coarse-grained disorder variables $\Theta^{(\mathrm{cg})}_{n,\varepsilon} := (\Theta^{(\mathrm{cg})}_{n,\varepsilon}(\vec{\mathsf{i}},\vec{\mathsf{a}}))_{(\vec{\mathsf{i}},\vec{\mathsf{a}})\in\mathbb{T}_\varepsilon}$, see (4.11). These satisfy Assumption A.1 with the following dependency neighborhoods:

- for each $\mathsf{z}_1 := (\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1) \in \mathbb{T}_\varepsilon$, its dependency neighbourhood is given by

$$A_{\mathsf{z}_1} = \Big\{ \mathsf{z}_2 = (\vec{\mathsf{i}}_2, \vec{\mathsf{a}}_2) \in \mathbb{T}_\varepsilon : \{\mathsf{i}_2, \mathsf{i}_2'\} \cap \{\mathsf{i}_1, \mathsf{i}_1'\} \neq \varnothing, \\ \operatorname{dist}(\{\mathsf{a}_2\}, \{\mathsf{a}_1, \mathsf{a}_1'\}) \wedge \operatorname{dist}(\{\mathsf{a}_2'\}, \{\mathsf{a}_1, \mathsf{a}_1'\}) \leqslant 2M_\varepsilon \Big\};$$

- given $\mathsf{z}_1 \in \mathbb{T}_\varepsilon$ and $\mathsf{z}_2 \in A_{\mathsf{z}_1}$, the dependency neighbourhood of $\{\mathsf{z}_1, \mathsf{z}_2\}$ is given by

$$A_{\mathsf{z}_1\mathsf{z}_2} = A_{\mathsf{z}_1} \cup A_{\mathsf{z}_2}.$$

Recalling the definition of $\mathbb{T}_\varepsilon$, we see that, uniformly in $\varepsilon > 0$ and $\mathsf{z} \in \mathbb{T}_\varepsilon$,

$$|A_{\mathsf{z}}| \leqslant C\, M_\varepsilon^2\, K_\varepsilon\, (M_\varepsilon \sqrt{K_\varepsilon})^2 = C\, M_\varepsilon^4\, K_\varepsilon^2\,. \tag{9.3}$$

In order to apply Lemma A.4, we first verify that condition (A.8) is satisfied by $\Phi(\Theta)$.

Lemma 9.3. *The multilinear polynomial* $\Phi(\Theta) := \mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi, \psi|\Theta)$ *satisfies condition* (A.8).

Proof. Condition (A.8) reads as

$$\forall\, \mathsf{z}_1 \in \mathbb{T}_\varepsilon,\ \forall\, \mathsf{z}_2 \in A_{\mathsf{z}_1},\ \forall\, \mathsf{z}_3 \in A_{\mathsf{z}_1\mathsf{z}_2} = A_{\mathsf{z}_1} \cup A_{\mathsf{z}_2} : \qquad \partial^2_{\mathsf{z}_i\mathsf{z}_j} \Phi = 0 \text{ for all } 1 \leqslant i, j \leqslant 3,$$

where ∂_{z} denotes derivative w.r.t. $\Theta(\mathsf{z})$. Since Φ is multilinear in $(\Theta(\mathsf{z}))_{\mathsf{z}\in\mathbb{T}_\varepsilon}$, this condition is equivalent to the claim that no term in the expansion of Φ (recall its definition from (4.8)) contains more than one of the factors $\Theta(\mathsf{z}_i)$, $1 \leqslant i \leqslant 3$. From the definition of Φ, clearly the product $\Theta(\mathsf{z}_1)\Theta(\mathsf{z}_2)$ cannot appear because $\mathsf{z}_1 = ((\mathsf{i}_1, \mathsf{i}_1'), (\mathsf{a}_1, \mathsf{a}_1'))$ and $\mathsf{z}_2 = ((\mathsf{i}_2, \mathsf{i}_2'), (\mathsf{a}_2, \mathsf{a}_2'))$ have an overlapping time index. Similarly, if $\mathsf{z}_3 \in A_{\mathsf{z}_i}$ for either $i = 1$ or 2, then the factor $\Theta(\mathsf{z}_i)\Theta(\mathsf{z}_3)$ cannot appear. The last case is if $\mathsf{z}_3 \in A_{\mathsf{z}_2}$, but $\mathsf{z}_3 \notin A_{\mathsf{z}_1}$ (the case $\mathsf{z}_3 \in A_{\mathsf{z}_1}$, $\mathsf{z}_3 \notin A_{\mathsf{z}_2}$ is the same by symmetry, since $\mathsf{z}_2 \in A_{\mathsf{z}_1}$ if and only if $\mathsf{z}_1 \in A_{\mathsf{z}_2}$): let us show that $\Theta(\mathsf{z}_1)\Theta(\mathsf{z}_3)$ does not appear in $\Phi(\Theta)$. Both $\vec{\mathsf{i}}_1$ and $\vec{\mathsf{i}}_3$ have an overlapping time index with $\vec{\mathsf{i}}_2$, hence $\operatorname{dist}(\vec{\mathsf{i}}_1, \vec{\mathsf{i}}_3) \leqslant \mathsf{i}_2' - \mathsf{i}_2 = |\vec{\mathsf{i}}_2| - 1 \leqslant K_\varepsilon - 1$, which contradicts the constraint imposed by $\vec{\boldsymbol{\mathcal{A}}}_\varepsilon^{(\text{no triple})}$ in (4.4), that for $\Theta(\mathsf{z}_1)\Theta(\mathsf{z}_3)$ to appear, we must have $\operatorname{dist}(\vec{\mathsf{i}}_1, \vec{\mathsf{i}}_3) \geqslant K_\varepsilon$. This verifies condition (A.8). □

We can then apply Lemmas A.2–A.4 to bound

$$\big|\mathrm{E}[f(\Phi(\Theta_m))] - \mathrm{E}[f(\Phi(\Theta_n))]\big| \leqslant I_1^{(m)} + I_2^{(m)} + I_1^{(n)} + I_2^{(n)} + I_3^{(m,n)}, \tag{9.4}$$

where $I_1^{(m)}$ and $I_2^{(m)}$ are the terms from applying Lemma A.2 to $h(\cdot) = f(\Phi(\cdot))$ and $X = \Theta_m$, see (A.4) and (A.5), similarly for $I_1^{(n)}$ and $I_2^{(n)}$, while $I_3^{(m,n)}$ is the term from applying Lemma A.3 to two Gaussian families $Z = \Theta_m^{(\mathrm{G})}$ and $\tilde{Z} = \Theta_n^{(\mathrm{G})}$ with the same mean and covariance structure as Θ_m and Θ_n, respectively, but independent of them, see (A.6).[†]

We are now ready to prove (9.2) exploiting (9.4). It suffices to prove that

$$\lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} I_1^{(n)} = 0\,, \qquad \lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} I_2^{(n)} = 0\,, \qquad \lim_{\varepsilon\downarrow 0} \limsup_{n,m\to\infty} I_3^{(m,n)} = 0\,, \tag{9.5}$$

We will prove these relations separately, exploiting (A.9), (A.10) and (A.11) from Lemma A.4. This will conclude the proof of Proposition 9.2.

[†]Since $(\Theta_n(\mathsf{z}))_{\mathsf{z}\in\mathbb{T}_\varepsilon}$ are uncorrelated, $(\Theta_n^{(\mathrm{G})}(\mathsf{z}))_{\mathsf{z}\in\mathbb{T}_\varepsilon}$ are in fact independent Gaussian random variables.

Bound on $\limsup_{n\to\infty} I_1^{(n)}$. By (A.9), we have

$$\begin{aligned} |I_1^{(n)}| \leqslant \frac{1}{2}\|f'''\|_\infty \sup_{\mathsf{z}_1\in\mathbb{T}_\varepsilon} \mathrm{E}\big[|\Theta_n(\mathsf{z}_1)|^3\big] \sum_{\mathsf{z}_1\in\mathbb{T}_\varepsilon,\, \mathsf{z}_2\in A_{\mathsf{z}_1},\, \mathsf{z}_3\in A_{\mathsf{z}_1\mathsf{z}_2}} \\ \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_1}\Phi\big(W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_2}\Phi\big(W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_3}\Phi\big(W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}\big)\big|^3\Big]^{\frac{1}{3}}, \end{aligned} \tag{9.6}$$

where for $s,t,u\in[0,1]$,

$$W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2} := su\sqrt{t}\Theta_n^{A_{\mathsf{z}_1}} + u\sqrt{t}\Theta_n^{A_{\mathsf{z}_1\mathsf{z}_2}\setminus A_{\mathsf{z}_1}} + \sqrt{t}\Theta_n^{A_{\mathsf{z}_1\mathsf{z}_2}^c} + \sqrt{1-t}\Theta_n^{(\mathrm{G})}, \tag{9.7}$$

with $\Theta_n^A(\mathsf{z}) := \Theta_n(\mathsf{z})\mathbb{1}_{\{\mathsf{z}\in A\}}$.

First note that by the assumption $\varphi\in C_c(\mathbb{R}^2)$ and the definition of the multilinear polynomial $\Phi(\Theta)$ in (4.8), $\Phi(\Theta)$ depends only on $\Theta(\mathsf{z})$ for a finite set of $\mathsf{z}\in\mathbb{T}_\varepsilon$. In particular, the sum in (9.6) is finite, and we can pass $\limsup_{n\to\infty}$ inside the sum.

Note that $\|f'''\|_\infty$ is bounded by assumption, and by (7.11),

$$\limsup_{n\to\infty} \sup_{\mathsf{z}_1\in\mathbb{T}_\varepsilon} \mathrm{E}\big[|\Theta_n(\mathsf{z}_1)|^3\big] \leqslant \limsup_{n\to\infty} \sup_{\mathsf{z}_1\in\mathbb{T}_\varepsilon} \mathrm{E}\big[|\Theta_n(\mathsf{z}_1)|^4\big]^{3/4} \leqslant \frac{C}{(\log\frac{1}{\varepsilon})^{3/4}}. \tag{9.8}$$

The sum in (9.6) can be bounded by

$$\sum_{\mathsf{z}_1\in\mathbb{T}_\varepsilon,\, \mathsf{z}_2\in A_{\mathsf{z}_1},\, \mathsf{z}_3\in A_{\mathsf{z}_1\mathsf{z}_2}} \frac{1}{3}\sum_{i=1}^3 \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_i}\Phi\big(W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}\big)\big|^4\Big]^{3/4}. \tag{9.9}$$

Given $\mathsf{z} = (\vec{\mathsf{i}},\vec{\mathsf{a}}) = ((\mathsf{i},\mathsf{i}'),(\mathsf{a},\mathsf{a}'))\in\mathbb{T}_\varepsilon$, by the definition of $\Phi(\Theta) = \mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta)$ in (4.8),

$$\partial_{\mathsf{z}}\Phi(\Theta) = \partial_{\Theta(\mathsf{z})}\mathscr{Z}_\varepsilon^{(\mathrm{cg})}(\varphi,\psi|\Theta) = \frac{2}{\varepsilon}\,\bar{\mathscr{Z}}_{[0,\mathsf{i}]}^{(\mathrm{cg})}(\varphi,\mathbb{1}_{\mathcal{S}_\varepsilon(\mathsf{a})}|\Theta)\bar{\mathscr{Z}}_{[\mathsf{i}',1/\varepsilon]}^{(\mathrm{cg})}(\mathbb{1}_{\mathcal{S}_\varepsilon(\mathsf{a}')},\psi|\Theta), \tag{9.10}$$

where $\bar{\mathscr{Z}}_{[0,\mathsf{i}]}^{(\mathrm{cg})}(\varphi,\mathbb{1}_{\{\mathcal{S}_\varepsilon(\mathsf{a})\}}|\Theta)$ denotes the centred coarse-grained model with initial condition φ at time 0 and terminal condition $\mathbb{1}_{\mathcal{S}_\varepsilon(\mathsf{a})}$ at time i: this is just the original coarse-grained model in (4.8) with time horizon i instead of $\lfloor 1/\varepsilon\rfloor$ and with the constant term $\frac{1}{2}g_{\frac{1}{2}}(\varphi,\psi)$ omitted. (We recall that $\mathcal{S}_\varepsilon(\mathsf{a})$ is a square of side length $\sqrt{\varepsilon}$ defined as in (5.35).) The definition of $\bar{\mathscr{Z}}_{[\mathsf{i}',1/\varepsilon]}^{(\mathrm{cg})}(\mathbb{1}_{\mathcal{S}_\varepsilon(\mathsf{a}')},\psi|\Theta)$ is similar, which is independent of $\bar{\mathscr{Z}}_{[0,\mathsf{i}]}^{(\mathrm{cg})}(\varphi,\mathbb{1}_{\{\mathcal{S}_\varepsilon(\mathsf{a})\}}|\Theta)$. Each of $\partial_{\Theta(\mathsf{z})}\mathscr{Z}_\varepsilon^{(\mathrm{cg})}$, $\bar{\mathscr{Z}}_{[0,\mathsf{i}]}^{(\mathrm{cg})}$ and $\bar{\mathscr{Z}}_{[\mathsf{i}',1/\varepsilon]}^{(\mathrm{cg})}$ contains a factor of $\varepsilon/2$ by the defintion of the coarse-grained model in (4.8), which is why there is a prefactor of $2/\varepsilon$ in (9.10). We then have

$$\begin{aligned} &\mathrm{E}\Big[\big|\partial_{\mathsf{z}}\Phi\big(W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}\big)\big|^4\Big] \\ =&\frac{16}{\varepsilon^4}\,\mathrm{E}\Big[\Big(\bar{\mathscr{Z}}_{[0,\mathsf{i}]}^{(\mathrm{cg})}(\mathbb{1}_{\mathcal{S}_\varepsilon(0)},\varphi(\cdot-\sqrt{\varepsilon}\mathsf{a})|W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2})\Big)^4\Big]\mathrm{E}\Big[\Big(\bar{\mathscr{Z}}_{[0,1/\varepsilon-\mathsf{i}']}^{(\mathrm{cg})}(\mathbb{1}_{\mathcal{S}_\varepsilon(0)},\psi(\cdot-\sqrt{\varepsilon}\mathsf{a}')|W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2})\Big)^4\Big], \end{aligned} \tag{9.11}$$

where we interchanged initial and terminal conditions by symmetry and used translation invariance and independence.

We can bound the two factors in the r.h.s. of (9.11) by applying slight variants of Theorem 8.1, which was formulated for the original coarse-grained model. Let us focus on the first factor in the r.h.s. of (9.11), for which we need to take into account two differences: the time range $[0,\mathsf{i}]$ instead of $[0,\varepsilon^{-1}]$ and the disorder $W_{s,t,u}^{\mathsf{z}_1,\mathsf{z}_2}$ instead of Θ_n.

The first difference is immaterial, because our moment estimate Theorem 8.1 is monotone increasing in the time length $\mathsf{i} \in [1, \varepsilon^{-1}]$ (see the argument leading to (8.4)). As a consequence, we can apply the bound in Theorem 8.1 with $w(x) = e^{-|x|}$ to $\bar{\mathscr{Z}}^{(\mathrm{cg})}_{[0,\mathsf{i}]}(\varphi, \psi|\Theta_n)$, which yields

$$\begin{aligned} \mathrm{E}\Big[\Big(\bar{\mathscr{Z}}^{(\mathrm{cg})}_{[0,\mathsf{i}]}(\mathbb{1}_{\mathcal{S}_\varepsilon(0)}, \varphi(\cdot - \sqrt{\varepsilon}\mathsf{a})|\Theta_n)\Big)^4\Big] &\leqslant C\varepsilon^{\frac{4}{p}} \Big\| \frac{\mathbb{1}_{\mathcal{S}_1(0)}}{w} \Big\|^4_{\ell^p} \|\varphi\|^4_\infty \|w \mathbb{1}_{\{B_\varphi - \sqrt{\varepsilon}\mathsf{a}\}}\|^4_q \\ &\leqslant C\|\varphi\|^4_\infty e^{-\mathrm{dist}(\sqrt{\varepsilon}\mathsf{a}, B_\varphi)} \varepsilon^{\frac{4}{p}}, \end{aligned} \tag{9.12}$$

where C depends on the choice of p, q, but is uniform in i and a and in n sufficiently large, while B_φ is a ball that contains the support of φ.

The second difference is also immaterial, that is we can replace Θ_n by $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ in (9.12). This is recorded in the following result, which we prove later.

Lemma 9.4. *The bound* (9.12) *also holds if* Θ_n *is replaced by* $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$*, uniformly in* $\mathsf{z}_1, \mathsf{z}_2 \in \mathbb{T}_\varepsilon$ *and* $s, t, u \in [0, 1]$*, and* n *large.*

Similarly, we can also bound the second factor in the r.h.s. (9.11) by

$$\mathrm{E}\Big[\Big(\bar{\mathscr{Z}}^{(\mathrm{cg})}_{[0,1/\varepsilon - \mathsf{i}']}(\mathbb{1}_{\mathcal{S}_\varepsilon(0)}, \psi(\cdot - \sqrt{\varepsilon}\mathsf{a}')|W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u})\Big)^4\Big] \leqslant C\varepsilon^{\frac{4}{p}} \Big\| \frac{\mathbb{1}_{\mathcal{S}_1(0)}}{w} \Big\|^4_{\ell^p} \|\psi\|^4_\infty \|w\|^4_q \leqslant C\|\psi\|^4_\infty \varepsilon^{\frac{4}{p}}. \tag{9.13}$$

Substituting these bounds into (9.11) and then (9.9) gives (for n sufficiently large)

$$\begin{aligned} \sum_{\substack{\mathsf{z}_1 \in \mathbb{T}_\varepsilon,\, \mathsf{z}_2 \in A_{\mathsf{z}_1} \\ \mathsf{z}_3 \in A_{\mathsf{z}_1\mathsf{z}_2}}} \frac{1}{3} \sum_{i=1}^3 \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_i} \Phi\big(W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}\big)\big|^4\Big]^{3/4} &\leqslant C\|\varphi\|^3_\infty \|\psi\|^3_\infty \varepsilon^{\frac{6}{p}-3} \sum_{\substack{\mathsf{z}_1 = (\vec{\mathsf{i}}_1, \vec{\mathsf{a}}_1) \in \mathbb{T}_\varepsilon,\, \mathsf{z}_2 \in A_{\mathsf{z}_1} \\ \mathsf{z}_3 \in A_{\mathsf{z}_1\mathsf{z}_2}}} e^{-\frac{3}{4}\mathrm{dist}(\sqrt{\varepsilon}\mathsf{a}_1, B_\varphi)} \\ &\leqslant C\|\varphi\|^3_\infty \|\psi\|^3_\infty \varepsilon^{\frac{6}{p}-5} M^{10}_\varepsilon K^6_\varepsilon, \end{aligned}$$

where we used the symmetry in the dependency structure between $\mathsf{z}_1, \mathsf{z}_2, \mathsf{z}_3$ ($\mathsf{z}_2 \in A_{\mathsf{z}_1}$ if and only if $\mathsf{z}_1 \in A_{\mathsf{z}_2}$) to reduce $\sum_{i=1}^3$ to the case $i = 1$, and in the last inequality, we first summed out $\mathsf{z}_2, \mathsf{z}_3$ and applied the bounds $|A_{z_1}|, |A_{\mathsf{z}_1\mathsf{z}_2}| \leqslant C M^4_\varepsilon K^2_\varepsilon$ from (9.3), then summed out $(\mathsf{i}'_1, \mathsf{a}'_1)$ in $\mathsf{z}_1 = ((\mathsf{i}_1, \mathsf{i}'_1), (\mathsf{a}_1, \mathsf{a}'_1))$ where the sum over i'_1 gives another factor of K_ε and the sum over a'_1 gives another factor of $M^2_\varepsilon K_\varepsilon$, and lastly we summed out $(\mathsf{i}_1, \mathsf{a}_1)$, noting that $\mathsf{i}_1 \in \{1, \dots, \varepsilon^{-1}\}$ while the sum over a_1 gives a factor $O((\sqrt{\varepsilon})^{-2}) = O(\varepsilon^{-1})$ because of the exponential decay on the scale $(\sqrt{\varepsilon})^{-1}$ (we recall that φ has compact support).

We recall from (4.2) that $K_\varepsilon = (\log \frac{1}{\varepsilon})^6$ and $M_\varepsilon = \log\log \frac{1}{\varepsilon}$. Choose $p > 1$ sufficiently close to 1 and substitute the above bound into (9.9) and then (9.6), together with (9.8), we then obtain that for any $\delta \in (0, 1)$, there exists C_δ such that

$$\limsup_{n\to\infty} |I^{(n)}_1| \leqslant C_\delta\, \varepsilon^\delta\,.$$

This proves the first relation in (9.5).

Bound on $\limsup_{n\to\infty} I^{(n)}_2$**.** By (A.10),

$$\begin{aligned} |I^{(n)}_2| \;\leqslant\; \frac{1}{2}\|f'''\|_\infty \sup_{\mathsf{z}_1 \in \mathbb{T}_\varepsilon} \mathrm{E}\big[|\Theta_n(\mathsf{z}_1)|^3\big] \sum_{\mathsf{z}_1 \in \mathbb{T}_\varepsilon,\, \mathsf{z}_2 \in A_{\mathsf{z}_1},\, \mathsf{z}_3 \in A_{\mathsf{z}_1\mathsf{z}_2}} \\ \sup_{t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_1} \Phi\big(W^{\mathsf{z}_1,\mathsf{z}_2}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_2} \Phi\big(W^{\mathsf{z}_1,\mathsf{z}_2}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{t,u} \mathrm{E}\Big[\big|\partial_{\mathsf{z}_3} \Phi\big(W^{\mathsf{z}_1,\mathsf{z}_2}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}}, \end{aligned}$$

where $W^{\mathsf{z}_1,\mathsf{z}_2}_{t,u} := u\sqrt{t}\Theta_n^{A_{\mathsf{z}_1\mathsf{z}_2}} + \sqrt{t}\Theta_n^{A^c_{\mathsf{z}_1\mathsf{z}_2}} + \sqrt{1-t}\Theta_n^{(\mathrm{G})}$ for $t,u\in[0,1]$. The bounds are exactly the same as for $|I_n^{(1)}|$, which gives $\limsup_{n\to\infty}|I_2^{(n)}| \leqslant C_\delta\varepsilon^\delta$ for any $\delta\in(0,1)$. This proves the second relation in (9.5).

Bound on $\limsup_{m,n\to\infty} I_3^{(m,n)}$. By (A.11) and the fact that Θ_n is a family of mean zero uncorrelated random variables,

$$|I_3^{(m,n)}| \leqslant \frac{1}{2}\|f'''\|_\infty \sum_{\mathsf{z}\in\mathbb{T}_\varepsilon} \big(\mathbb{E}\big[\Theta_n^2(\mathsf{z})\big] - \mathbb{E}\big[\Theta_m^2(\mathsf{z})\big]\big) \sup_{t\in[0,1]} \mathrm{E}\big[\big|\partial_{\mathsf{z}}\Phi(W_t)\big|^2\big] \tag{9.14}$$

where $W_t := \sqrt{t}\Theta_n^{(\mathrm{G})} + \sqrt{1-t}\Theta_m^{(\mathrm{G})}$ for $t\in[0,1]$. Note that by definition, $\Phi(\Theta)$ depends only on a finite set of Θ_{z}, $\mathsf{z}\in\mathbb{T}_\varepsilon$. Therefore the sum in (9.14) contains finitely many terms. For each $\mathsf{z}\in\mathbb{T}_\varepsilon$,

$$\lim_{m,n\to\infty} \big(\mathbb{E}\big[\Theta_n^2(\mathsf{z})\big] - \mathbb{E}\big[\Theta_m^2(\mathsf{z})\big]\big) = 0$$

by Lemma 7.1. On the other hand, uniformly in $t\in[0,1]$, $\mathrm{E}\big[\big|\partial_{\mathsf{z}}\Phi(W_t)\big|^2\big]$ converges to a finite limit as $m,n\to\infty$ because $\partial_{\mathsf{z}}\Phi(W_t)$ is a multilinear polynomial in $W_t(\mathsf{z})$ for finitely many $\mathsf{z}\in\mathbb{T}_\varepsilon$, while its second moment is a multilinear polynomial of $\mathrm{E}[W_t^2(\mathsf{z})]$, $\mathsf{z}\in\mathbb{T}_\varepsilon$, each of which converges by Lemma 7.1. It follows that $\limsup_{m,n\to\infty}|I_3^{(m,n)}| = 0$, which is stronger than the third relation in (9.5).

Conclusion. Assuming Lemma 9.4, we have proved (9.5). This implies (9.2) and finally completes the proof of Proposition 9.2. □

Remark 9.5 (Extension to finite-dimensional distribution). *Finally, to prove the finite-dimensional distribution convergence in Theorem 9.1, we argue as in the proof of Proposition 9.2. First we approximate the components $\mathscr{Z}^{\beta_N}_{N,s_i,t_i}(\varphi_i,\psi_i)$ of the random vector in the l.h.s. of* (9.1) *by coarse-grained models $\mathscr{Z}^{(\mathrm{cg})}_{\varepsilon,s_i,t_i}(\varphi_i,\psi_i|\Theta)$, with the same coarse-grained disorder variables $\Theta = \Theta^{(\mathrm{cg})}_{N,\varepsilon}$, which we can do with a small L^2 error, uniformly in large N, provided we choose $\varepsilon>0$ small enough, by Theorem 4.7.*

It remains to apply a Lindeberg principle for a vector of multilinear polynomials, which is given in Remark A.5. The estimates needed are exactly the same as in the Lindeberg principle for a single multilinear polynomial. This concludes the proof of Theorem 9.1.

Proof of Lemma 9.4. We re-examine the proof of Theorem 8.1. First note that, similar to the L^2 orthogonal decomposition of $\Theta_n(\mathsf{i},\mathsf{a})$ with $|\vec{\mathsf{i}}|\geqslant 1$ as in (8.5), we can write

$$W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}(\vec{\mathsf{i}},\vec{\mathsf{a}}) = \sum_{\substack{\mathsf{b}:\,|\mathsf{b}-\mathsf{a}|\leqslant M_\varepsilon,\,\mathsf{b}':\,|\mathsf{b}'-\mathsf{a}'|\leqslant M_\varepsilon\\ |\mathsf{b}'-\mathsf{b}|\leqslant M_\varepsilon\sqrt{|\vec{\mathsf{i}}|}}} W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}')), \tag{9.15}$$

where $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))$ are defined through the same mixture as in (9.7) between $\Theta_n(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))$ and $\Theta_n^{(\mathrm{G})}(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))$, with the latter being independent normals with mean 0 and the same variance as $\Theta_n(\vec{\mathsf{i}},(\mathsf{a},\mathsf{b}),(\mathsf{b}',\mathsf{a}'))$. We can then carry out the same expansion as for (8.4), and note that: whenever a product of coarse-grained disorder variables Θ_n has zero expectation because of the presence of some $\Theta_n(\mathsf{i},\mathsf{a})$ with either (i,a) or $(\mathsf{i}',\mathsf{a}')$ unmatched by any other Θ_n in the product, the same is true if the family Θ_n is replaced by

$W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$; similarly, whenever two collections of Θ_n variables are independent of each other, the same is true if Θ_n is replaced by $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$. This implies that the expansion and re-summation carried out for the r.h.s. of (8.4), as well as the accompanying constraints on summation indices, also apply when Θ_n is replaced by $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$.

Next we claim that, for $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$ that visits two distinct blocks $\mathcal{I}_1,\mathcal{I}_2$ (see the exposition leading to (8.6)), although we no longer have a chaos expansion representation as in (8.6) due to the Gaussian component of $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$, we can still make a replacement similar to (8.8) in order to bound the r.h.s. of (8.4), with Θ_n replaced by $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$:

$$\begin{aligned} &W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2)) \\ \overset{\leqslant}{\rightsquigarrow}\ &C W^{\mathsf{z}_1,\mathsf{z}_2}_{1,1/2,1}(\mathsf{i}_1;\mathsf{a}_1,\mathsf{b}_1)\, g_{10(\mathsf{i}_2-\mathsf{i}_1)}(\mathsf{a}_2-\mathsf{b}_1)\, W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}(\mathsf{i}_2;\mathsf{a}_2,\mathsf{b}_2). \end{aligned} \tag{9.16}$$

To see this, note that $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ is a mixture of Θ_n and $\Theta^{(\mathrm{G})}_n$ with mixture coefficients given in (9.7). When we expand the r.h.s. of (8.4) with $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ in place of $\Theta^{(\mathrm{cg})}_{N,\varepsilon}$, we can further expand $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ in terms of its mixture. Each term in the expansion then consists of a product of Θ_n variables and $\Theta^{(\mathrm{G})}_n$ variables, whose expectation factorises due to the independence between Θ_n and $\Theta^{(\mathrm{G})}_n$. For terms that contain the factor $\Theta_n((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$, the decomposition (8.6) applies, and the same argument justifying the replacement (8.8) can be applied here. For terms that contain the Gaussian factor $\Theta^{(\mathrm{G})}_n((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$, the expectation will be non-zero only if this factor appears exactly twice or four times. The resulting contribution will be either the second or the fourth moment of $\Theta^{(\mathrm{G})}_n((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$. Its second moment coincide with that of $\Theta_n((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$, while its fourth moment can be bounded in terms of its second moment by Gaussianity. Then as in (8.8), we make the following replacement in the expansion:

$$\Theta^{(\mathrm{G})}_n((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2)) \overset{\leqslant}{\rightsquigarrow} C\Theta^{(\mathrm{G})}_n(\mathsf{i}_1;\mathsf{a}_1,\mathsf{b}_1)\, g_{10(\mathsf{i}_2-\mathsf{i}_1)}(\mathsf{a}_2-\mathsf{b}_1)\, \Theta^{(\mathrm{G})}_n(\mathsf{i}_2;\mathsf{a}_2,\mathsf{b}_2). \tag{9.17}$$

Therefore in the mixture of $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}((\mathsf{i}_1,\mathsf{i}_2),(\mathsf{a}_1,\mathsf{b}_1),(\mathsf{a}_2,\mathsf{b}_2))$ (recall (9.7)), we can replace the Θ_n and $\Theta^{(\mathrm{G})}_n$ components each by its factorisation as in (9.17). The mixture coefficient of the term $C\Theta_n(\mathsf{i}_1;\mathsf{a}_1,\mathsf{b}_1)g_{10(\mathsf{i}_2-\mathsf{i}_1)}(\mathsf{a}_2-\mathsf{b}_1)\Theta_n(\mathsf{i}_2;\mathsf{a}_2,\mathsf{b}_2)$ is equal to $\alpha\sqrt{t}$ with either $\alpha = su$, u or 1, while the mixture coefficient of the term $C\Theta^{(\mathrm{G})}_n(\mathsf{i}_1;\mathsf{a}_1,\mathsf{b}_1)g_{10(\mathsf{i}_2-\mathsf{i}_1)}(\mathsf{a}_2-\mathsf{b}_1)\Theta^{(\mathrm{G})}_n(\mathsf{i}_2;\mathsf{a}_2,\mathsf{b}_2)$ equals $\sqrt{1-t}$. This mixture can be further replaced by the r.h.s. of (9.16), which just contains extra terms and larger coefficients, where the choice of $s = u = 1$ and $t = 1/2$ in the first factor $W^{\mathsf{z}_1,\mathsf{z}_2}_{1,1/2,1}(\mathsf{i}_1;\mathsf{a}_1,\mathsf{b}_1)$ helps to bound the mixture coefficients.

The remaining parts of the proof of Theorem 8.1 depends on the coarse-grained disorder variables $\Theta_n = \Theta^{(\mathrm{cg})}_{n,\varepsilon}$ only through their second and fourth moments. Note that uniformly in $s,t,u \in [0,1]$ and $\mathsf{z}_1,\mathsf{z}_2$, the second moment of $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ is bounded by that of Θ_n, and modulo a constant multiple, the fourth moment of $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$ can be bounded by that of Θ_n. Therefore the remaining parts of the proof of Theorem 8.1 carries through without change if Θ_n is replaced by $W^{\mathsf{z}_1,\mathsf{z}_2}_{s,t,u}$. In particular, the bound on $U^{(\mathrm{cg})}_{N,\varepsilon}$ in (8.10) still holds since it only depends on the second moment of Θ_n, and the bound on $V^{(\mathrm{cg})}_{N,\varepsilon}$ in (8.11) still holds because it only depends on the fourth moment of Θ_n. □

9.2. Proof of Theorem 1.2. The translation invariance of the law of $\mathscr{Z}^{\vartheta}$ is obvious.

To prove the scaling relation (1.12), let us write $\beta_N(\vartheta)$ to emphasize that β_N as specified in (1.11) depends on a parameter ϑ. Given $\mathsf{a} > 0$, let $\tilde{N} = N/\mathsf{a}$. Then using (1.11) and the

fact that $R_N = \frac{1}{\pi}(\log N + \alpha + o(1))$ given later in (3.2), we have

$$\sigma_N^2 = \frac{1}{R_N}\Big(1 + \frac{\vartheta + o(1)}{\log N}\Big) = \frac{1}{R_{\tilde{N}}}\Big(1 + \frac{\vartheta - \log \mathsf{a} + o(1)}{\log \tilde{N}}\Big), \tag{9.18}$$

so that $\beta_N(\vartheta) = \beta_{\tilde{N}}(\vartheta - \log \mathsf{a})$, or equivalently, $\beta_N(\vartheta + \log a) = \beta_{\tilde{N}}(\vartheta)$.

By (9.1), for $\varphi, \psi \in C_c(\mathbb{R}^2)$, we have

$$\begin{aligned}\iint \varphi\Big(\frac{x}{\sqrt{a}}\Big)\psi\Big(\frac{y}{\sqrt{a}}\Big)\mathscr{Z}_{\tilde{N};\,\mathsf{a}s,\mathsf{a}t}^{\beta_{\tilde{N}}(\vartheta),\omega}(\mathrm{d}x,\mathrm{d}y) &\underset{\tilde{N}\to\infty}{\Longrightarrow} \iint \varphi\Big(\frac{x}{\sqrt{a}}\Big)\psi\Big(\frac{y}{\sqrt{a}}\Big)\mathscr{Z}_{\mathsf{a}s,\mathsf{a}t}^{\vartheta}(\mathrm{d}x,\mathrm{d}y)\\ &= \iint \varphi(x)\psi(y)\mathscr{Z}_{\mathsf{a}s,\mathsf{a}t}^{\vartheta}(\mathrm{d}(\sqrt{\mathsf{a}}x),\mathrm{d}(\sqrt{\mathsf{a}}y)).\end{aligned} \tag{9.19}$$

On the other hand, recall (1.9), we can rewrite the l.h.s. of (9.19) as

$$\begin{aligned}&\tilde{N}\iint \varphi\Big(\frac{x}{\sqrt{a}}\Big)\psi\Big(\frac{y}{\sqrt{a}}\Big)Z_{[\![\tilde{N}\mathsf{a}s]\!],[\![\tilde{N}\mathsf{a}t]\!]}^{\beta_{\tilde{N}}(\vartheta),\,\omega}([\![\sqrt{\tilde{N}}x]\!],[\![\sqrt{\tilde{N}}y]\!])\,\mathrm{d}x\,\mathrm{d}y\\ =\;& \frac{1}{\tilde{N}}\sum_{\tilde{x},\tilde{y}\in\mathbb{Z}^2_{\mathrm{even}}} Z_{[\![\tilde{N}\mathsf{a}s]\!],[\![\tilde{N}\mathsf{a}t]\!]}^{\beta_{\tilde{N}}(\vartheta),\,\omega}(\tilde{x},\tilde{y}) \int\limits_{|u-\tilde{x}|_1\leqslant 1}\varphi\Big(\frac{u}{\sqrt{\mathsf{a}\tilde{N}}}\Big)\mathrm{d}u \int\limits_{|v-\tilde{y}|_1\leqslant 1}\psi\Big(\frac{v}{\sqrt{\mathsf{a}\tilde{N}}}\Big)\mathrm{d}v\\ =\;& \frac{\mathsf{a}}{N}\sum_{\tilde{x},\tilde{y}\in\mathbb{Z}^2_{\mathrm{even}}} Z_{[\![Ns]\!],[\![Nt]\!]}^{\beta_N(\vartheta+\log\mathsf{a}),\,\omega}(\tilde{x},\tilde{y}) \int\limits_{|u-\tilde{x}|_1\leqslant 1}\varphi\Big(\frac{u}{\sqrt{N}}\Big)\mathrm{d}u \int\limits_{|v-\tilde{y}|_1\leqslant 1}\psi\Big(\frac{v}{\sqrt{N}}\Big)\mathrm{d}v\\ \underset{N\to\infty}{\Longrightarrow}&\iint \varphi(x)\psi(y)\,\mathsf{a}\mathscr{Z}_{s,t}^{\vartheta+\log\mathsf{a}}(\mathrm{d}x,\mathrm{d}y),\end{aligned}$$

where we again applied (9.1). Combined with (9.19), this implies (1.12).

The first and second moments of $\mathscr{Z}^{\vartheta}$ can be identified from the limits in (3.53) and (3.54), since for $\varphi \in C_c(\mathbb{R}^2)$ and $\psi \in C_b(\mathbb{R}^2)$, the averaged partition function $\mathcal{Z}_{N,t-s}^{\beta_N}(\varphi,\psi)$ has a finite fourth moment that is uniformly bounded in N, see Theorem 6.1.

Appendix A. Enhanced Lindeberg principle

In this appendix, we prove a Lindeberg principle for multilinear polynomials of dependent random variables with a local form of dependence. This extends [Rol13], which requires the function to have bounded first three derivatives and is not applicable to multilinear polynomials, and it extends [MOO10], which considers multilinear polynomials of independent random variables. We first introduce the necessary setup. Let $\mathbb{T}$ be a finite index set.

Assumption A.1 (Local dependence). *Let $X = (X_i)_{i\in\mathbb{T}}$ be random variables satisfying:*

- $\mathrm{E}[X_i] = 0$ *and* $\mathrm{E}[X_iX_j] = \sigma_{ij}$*;*
- *for every $k \in \mathbb{T}$ there is $A_k \subseteq \mathbb{T}$ such that X_k is independent of $(X_i)_{i\in A_k^c}$;*
- *for all $k \in \mathbb{T}$, $l \in A_k$ there is $A_{kl} \subseteq \mathbb{T}$ such that (X_k, X_l) is independent of $(X_i)_{i\in A_{kl}^c}$.*

The sets $(A_k)_{k\in\mathbb{T}}$ and $(A_{kl})_{k\in\mathbb{T},l\in A_k}$ will be called dependency neighbourhoods *of $X = (X_i)_{i\in\mathbb{T}}$.*

Let $Z = (Z_i)_{i\in\mathbb{T}} \sim N(0,\sigma_\cdot)$ be a Gaussian vector independent of X, but with the same covariance matrix as X. For $u,t \in [0,1]$, $k \in \mathbb{T}$ and $l \in A_k$, we define

$$W_{t,u}^{k,l} := u\sqrt{t}X^{A_{kl}} + \sqrt{t}X^{A_{kl}^c} + \sqrt{1-t}Z, \tag{A.1}$$

where for any subset of indices $B \subseteq \mathbb{T}$ we write

$$X_i^B := X_i \mathbb{1}_{\{i \in B\}}. \tag{A.2}$$

For $s, u, t \in [0,1]$, $k \in \mathbb{T}$ and $l \in A_k$, we define

$$W_{s,t,u}^{k,l} := su\sqrt{t}X^{A_k} + u\sqrt{t}X^{A_{kl}\setminus A_k} + \sqrt{t}X^{A_{kl}^c} + \sqrt{1-t}Z. \tag{A.3}$$

We have the following Lindeberg type result, which controls the distributional distance between X and Z through smooth test functions.

Lemma A.2 (Lindeberg principle for dependent random variables). *Let X, Z, $W_{t,u}^{k,l}$ and $W_{s,t,u}^{k,l}$ be defined as above. Let $h : \mathbb{R}^{|\mathbb{T}|} \to \mathbb{R}$ be bounded and thrice differentiable. Then*

$$\mathrm{E}[h(X)] - \mathrm{E}[h(Z)] = I_1 + I_2$$

where

$$I_1 := \frac{1}{2}\int_{[0,1]^3} \sum_{k\in\mathbb{T},\, l\in A_k,\, m\in A_{kl}} \mathrm{E}\Big[X_k X_l X_m \big(s\mathbb{1}_{\{m\in A_k\}} + \mathbb{1}_{\{m\in A_{kl}\setminus A_k\}}\big)\sqrt{t}\,\partial_{klm}^3 h\big(W_{s,t,u}^{k,l}\big)\Big]\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}u\,, \tag{A.4}$$

$$I_2 := -\frac{1}{2}\int_{[0,1]^2} \sum_{k\in\mathbb{T},\, l\in A_k,\, m\in A_{kl}} \sigma_{kl}\,\mathrm{E}\Big[X_m\sqrt{t}\,\partial_{klm}^3 h\big(W_{t,u}^{k,l}\big)\Big]\mathrm{d}t\,\mathrm{d}u\,, \tag{A.5}$$

assuming that the integrals and expectations above are all finite.

To control the distributional distance between X and another random vector $\widetilde{X} = (\widetilde{X}_i)_{i\in\mathbb{T}}$ with slightly perturbed covariance matrix $(\tilde{\sigma}_{ij})_{i,j\in\mathbb{T}}$ and dependency neighbourhoods $\tilde{A}_k, \tilde{A}_{kl}$, we will apply Lemma A.2 to X and $\widetilde{X}$ separately, plus the following result that compares two Gaussian vectors.

Lemma A.3. *Let $Z = (Z_i)_{i\in\mathbb{T}} \sim N(0, \sigma_\cdot)$ and $\widetilde{Z} = (\widetilde{Z}_i)_{i\in\mathbb{T}} \sim N(0, \tilde{\sigma}_\cdot)$ be centred Gaussian random vectors with covariance matrices $(\sigma_{ij})_{i,j\in\mathbb{T}}$ and $(\tilde{\sigma}_{ij})_{i,j\in\mathbb{T}}$ respectively. Let $h : \mathbb{R}^{|\mathbb{T}|} \to \mathbb{R}$ be bounded and twice differentiable. Denote $W_t := \sqrt{t}\widetilde{Z} + \sqrt{1-t}Z$. Then we have*

$$\mathrm{E}[h(\widetilde{Z})] - \mathrm{E}[h(Z)] =: I_3 = \frac{1}{2}\sum_{k,l\in\mathbb{T}} (\tilde{\sigma}_{kl} - \sigma_{kl}) \int_0^1 \mathrm{E}\Big[\partial_{kl}^2 h(W_t)\Big]\mathrm{d}t, \tag{A.6}$$

assuming that the integrals and expectations above are all finite.

We now specialise Lemmas A.2 and A.3 to our case of interest, where h is a function of a multi-linear polynomial $\Phi(X)$.

Lemma A.4. *Assume that $h(X) := f(\Phi(X))$ for some bounded $f : \mathbb{R} \to \mathbb{R}$ with bounded first three derivatives, and*

$$\Phi(X) = \sum_{I\subseteq\mathbb{T}} c_I \prod_{i\in I} X_i \tag{A.7}$$

for some fixed coefficients $c_I \in \mathbb{R}$. Furthermore, assume that

$$\forall\, k \in \mathbb{T},\, l \in A_k,\, m \in A_{kl}, \qquad \partial_{km}^2\Phi = \partial_{lm}^2\Phi = \partial_{kl}^2\Phi = 0\,. \tag{A.8}$$

Then for I_1, I_2 and I_3 as in (A.4), (A.5) *and* (A.6), *we have*

$$|I_1| \leqslant \frac{1}{2}\|f'''\|_\infty \sup_{k\in\mathbb{T}} \mathrm{E}\big[|X_k|^3\big] \sum_{k\in\mathbb{T},\, l\in A_k,\, m\in A_{kl}} \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_k\Phi\big(W^{k,l}_{s,t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_l\Phi\big(W^{k,l}_{s,t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{s,t,u} \mathrm{E}\Big[\big|\partial_m\Phi\big(W^{k,l}_{s,t,u}\big)\big|^3\Big]^{\frac{1}{3}}, \quad \text{(A.9)}$$

$$|I_2| \leqslant \frac{1}{2}\|f'''\|_\infty \sup_{k\in\mathbb{T}} \mathrm{E}\big[|X_k|^3\big] \sum_{k\in\mathbb{T},\, l\in A_k,\, m\in A_{kl}} \sup_{t,u} \mathrm{E}\Big[\big|\partial_k\Phi\big(W^{k,l}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{t,u} \mathrm{E}\Big[\big|\partial_l\Phi\big(W^{k,l}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}} \sup_{t,u} \mathrm{E}\Big[\big|\partial_m\Phi\big(W^{k,l}_{t,u}\big)\big|^3\Big]^{\frac{1}{3}}, \quad \text{(A.10)}$$

$$|I_3| \leqslant \frac{1}{2}\|f''\|_\infty \sum_{k\in\mathbb{T},\, l\in A_k} (\tilde\sigma_{kl} - \sigma_{kl}) \sup_t \mathrm{E}\Big[\big(\partial_k\Phi(W_t)\big)^2\Big]^{\frac{1}{2}} \sup_t \mathrm{E}\Big[\big(\partial_l\Phi(W_t)\big)^2\Big]^{\frac{1}{2}}. \quad \text{(A.11)}$$

Remark A.5. *We can extend Lemma A.4 to the vector setting, i.e. for a function $h(X) = f(\Phi^{(1)}(X), \dots, \Phi^{(k)}(X))$ of a finite number k of multi-linear polynomials $\Phi^{(i)}(X)$ as in* (A.7), *each satisfying* (A.8), *where $f : \mathbb{R}^k \to \mathbb{R}$ has bounded partial derivatives of order up to three. The bounds* (A.9)-(A.10) *are simply modified replacing $\|f'''\|_\infty$ by $\max_{1\leqslant i,j,l\leqslant k} \|\partial_{i,j,l} f\|_\infty$ and the three occurrences of Φ by $\Phi^{(i)}$, $\Phi^{(j)}$, $\Phi^{(l)}$, and then summing over $1 \leqslant i, j, l \leqslant k$; similar modifications apply to the bound* (A.11). *The adaptation of the proof is straightforward.*

We now give the proofs of Lemmas A.2–A.4.

Proof of Lemma A.2. Let $Y_t := \sqrt{t}X + \sqrt{1-t}Z$, $t \in [0,1]$. Then we can write

$$\begin{aligned}\mathrm{E}[h(X)] - \mathrm{E}[h(Z)] &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}[h(Y_t)]\,\mathrm{d}t \\ &= \frac{1}{2}\int_0^1 \mathrm{E}\bigg[\sum_{k\in\mathbb{T}} \frac{X_k}{\sqrt{t}}\,\partial_k h(Y_t) - \sum_{k\in\mathbb{T}} \frac{Z_k}{\sqrt{1-t}}\,\partial_k h(Y_t)\bigg]\mathrm{d}t\,.\end{aligned}$$

Given $C \subseteq \mathbb{T}$, let us denote

$$U^C_t := \sqrt{t}X^C + \sqrt{1-t}Z\,, \qquad \text{where} \qquad X^C_i := X_i\,\mathbb{1}_{\{i\in C\}}\,.$$

In particular, $Y_t = U^{\mathbb{T}}_t$. Observe that $\mathrm{E}[X_k\,\partial_k h(U^{A_k^c}_t)] = \mathrm{E}[X_k]\mathrm{E}[\partial_k h(U^{A_k^c}_t)] = 0$ by independence. We recall Gaussian integration by parts: for smooth functions φ,

$$\mathrm{E}[Z_k\,\varphi(Z)] = \sum_{l\in\mathbb{T}} \sigma_{kl}\,\mathrm{E}[\partial_l\varphi(Z)] = \sum_{l\in A_k} \sigma_{kl}\,\mathrm{E}[\partial_l\varphi(Z)]\,,$$

since $\sigma_{kl} = 0$ for $l \notin A_k$. Then

$$\mathrm{E}[h(X)] - \mathrm{E}[h(Z)] = \frac{1}{2}\int_0^1 \mathrm{E}\bigg[\sum_{k\in\mathbb{T}} \frac{X_k}{\sqrt{t}}\,\big(\partial_k h(U^{\mathbb{T}}_t) - \partial_k h(U^{A_k^c}_t)\big) - \sum_{k\in\mathbb{T},\, l\in A_k} \sigma_{kl}\,\partial^2_{kl} h(U^{\mathbb{T}}_t)\bigg]\mathrm{d}t\,.$$

Let us expand the first term. We can interpolate between $U^{\mathbb{T}}_t$ and $U^{A_k^c}_t$ by

$$U^{A_k^c}_t + s\sqrt{t}X^{A_k} \qquad \text{for } s \in [0,1]\,.$$

Then

$$\begin{aligned}\partial_k h(U_t^{\mathbb{T}}) - \partial_k h(U_t^{A_k^c}) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} \partial_k h(U_t^{A_k^c} + s\sqrt{t} X^{A_k})\,\mathrm{d}s \\ &= \sqrt{t} \int_0^1 \sum_{l\in\mathbb{T}} X_l^{A_k}\, \partial_{kl}^2 h\big(U_t^{A_k^c} + s\sqrt{t} X^{A_k}\big)\,\mathrm{d}s\,.\end{aligned}$$

Note that we can restrict the sum to $l \in A_k$ because $X_l^{A_k} = 0$ otherwise. This leads to

$$\begin{aligned}\mathrm{E}[h(X)] - \mathrm{E}[h(Z)] = \frac{1}{2}\int_0^1\int_0^1 \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k} X_k\, X_l\, \partial_{kl}^2 h\big(U_t^{A_k^c} + s\sqrt{t} X^{A_k}\big) \\ - \sum_{k\in\mathbb{T},\, l\in A_k} \sigma_{kl}\, \partial_{kl}^2 h(U_t^{\mathbb{T}})\Big]\mathrm{d}s\,\mathrm{d}t\,.\end{aligned}$$

Note that $\mathrm{E}[X_k\, X_l\, \partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big)] = \mathrm{E}[X_k\, X_l]\mathrm{E}[\partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big)] = \sigma_{kl}\mathrm{E}[\partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big)]$ by the independence assumption. By adding and subtracting this term, we get

$$\mathrm{E}[h(X)] - \mathrm{E}[h(Z)] = I_1 + I_2$$

where

$$\begin{aligned}I_1 &:= \frac{1}{2}\int_0^1\int_0^1 \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k} X_k\, X_l \Big(\partial_{kl}^2 h\big(U_t^{A_k^c} + s\sqrt{t} X^{A_k}\big) - \partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big)\Big)\Big]\,\mathrm{d}s\,\mathrm{d}t \\ I_2 &:= \frac{1}{2}\int_0^1\int_0^1 \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k} \sigma_{kl} \Big(\partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big) - \partial_{kl}^2 h(U_t^{\mathbb{T}})\Big)\Big]\mathrm{d}s\,\mathrm{d}t \\ &= \frac{1}{2}\int_0^1 \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k} \sigma_{kl} \Big(\partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big) - \partial_{kl}^2 h(U_t^{\mathbb{T}})\Big)\Big]\mathrm{d}t\,,\end{aligned}$$

where we performed the integration on s in I_2 (whose integrand does not depend on s).

Let us deal with I_1. Note that

$$U_t^{A_k^c} + s\sqrt{t} X^{A_k} = s\sqrt{t} X^{A_k} + \sqrt{t} X^{A_k^c} + \sqrt{1-t} Z \quad \text{and} \quad U_t^{A_{kl}^c} = \sqrt{t} X^{A_{kl}^c} + \sqrt{1-t} Z.$$

We can therefore interpolate between $U_t^{A_{kl}^c}$ and $U_t^{A_k^c} + s\sqrt{t} X^{A_k}$ by

$$W_{s,t,u}^{k,l} := U_t^{A_{kl}^c} + u\sqrt{t}(sX^{A_k} + X^{A_{kl}\setminus A_k})\,, \qquad u \in [0,1]\,.$$

Note that $(sX^{A_k} + X^{A_{kl}\setminus A_k})_m = 0$ for $m \notin A_{kl}$. We can then write

$$\begin{aligned}\partial_{kl}^2 h\big(U_t^{A_k^c} + s\sqrt{t} X^{A_k}\big) - \partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}u} \partial_{kl}^2 h\big(W_{s,t,u}^{k,l}\big)\,\mathrm{d}u \\ = \sqrt{t} \int_0^1 \sum_{m\in A_{kl}} X_m (s\mathbb{1}_{\{m\in A_k\}} + \mathbb{1}_{\{m\in A_{kl}\setminus A_k\}})\, \partial_{klm}^3 h\big(W_{s,t,u}^{k,l}\big)\,\mathrm{d}u\,.\end{aligned}$$

This yields the final form of I_1:

$$I_1 := \frac{1}{2}\int_{[0,1]^3} \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k,\, m\in A_{kl}} X_k\, X_l\, X_m (s\mathbb{1}_{\{m\in A_k\}} + \mathbb{1}_{\{m\in A_{kl}\setminus A_k\}})\, \sqrt{t}\, \partial_{klm}^3 h\big(W_{s,t,u}^{k,l}\big)\Big]\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}u.$$

Let us now deal with I_2. We can interpolate between $U_t^{A_{kl}^c}$ and $U_t^{\mathbb{T}}$ by

$$W_{t,u}^{k,l} := U_t^{A_{kl}^c} + u\sqrt{t}X^{A_{kl}}, \qquad u \in [0,1].$$

Then

$$\partial_{kl}^2 h\big(U_t^{A_{kl}^c}\big) - \partial_{kl}^2 h(U_t^{\mathbb{T}}) = -\int_0^1 \frac{\mathrm{d}}{\mathrm{d}u}\partial_{kl}^2 h\big(W_{t,u}^{k,l}\big)\,\mathrm{d}u = -\sqrt{t}\int_0^1 \sum_{m\in A_{kl}} X_m\,\partial_{klm}^3 h\big(W_{t,u}^{k,l}\big)\,\mathrm{d}u.$$

This yields the final form of I_2:

$$I_2 := -\frac{1}{2}\int_{[0,1]^2} \mathrm{E}\Big[\sum_{k\in\mathbb{T},\, l\in A_k,\, m\in B_{kl}} \sigma_{kl}\,X_m\,\sqrt{t}\,\partial_{klm}^3 h\big(W_{t,u}^{k,l}\big)\Big]\mathrm{d}t\,\mathrm{d}u.$$

This concludes the proof of Lemma A.2. □

Proof of Lemma A.3. Let $W_t := \sqrt{t}\widetilde{Z} + \sqrt{1-t}Z$, $t\in[0,1]$. Using Gaussian integration by parts as in the proof of Lemma A.2, we have

$$\begin{aligned}
\mathrm{E}[h(\widetilde{Z})] - \mathrm{E}[h(Z)] &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}[h(W_t)]\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^1 \mathrm{E}\Big[\sum_{k\in\mathbb{T}} \frac{\widetilde{Z}_k}{\sqrt{t}}\,\partial_k h(W_t) - \sum_{k\in\mathbb{T}} \frac{Z_k}{\sqrt{1-t}}\,\partial_k h(W_t)\Big]\mathrm{d}t \\
&= \frac{1}{2}\sum_{k,l\in\mathbb{T}} (\tilde{\sigma}_{kl} - \sigma_{kl}) \int_0^1 \mathrm{E}\Big[\partial_{kl}^2 h(W_t)\Big]\mathrm{d}t,
\end{aligned}$$

which proves the lemma. □

Proof of Lemma A.4. We can easily compute

$$\begin{aligned}
\partial_k h(x) &= f'(\Phi(x))\,\partial_k\Phi(x), \\
\partial_{kl}^2 h(x) &= f''(\Phi(x))\,\partial_k\Phi(x)\,\partial_l\Phi(x) + f'(\Phi(x))\,\partial_{kl}^2\Phi(x), \\
\partial_{klm}^3 h(x) &= f'''(\Phi(x))\,\partial_k\Phi(x)\,\partial_l\Phi(x)\,\partial_m\Phi(x) \\
&\quad + f''(\Phi(x))\big\{\partial_{km}^2\Phi(x)\,\partial_l\Phi(x) + \partial_{lm}^2\Phi(x)\,\partial_k\Phi(x) + \partial_{kl}^2\Phi(x)\,\partial_m\Phi(x)\big\} \\
&\quad + f'(\Phi(x))\,\partial_{klm}^3\Phi(x).
\end{aligned}$$

which by assumption (A.8), gives

$$\partial_{klm}^3 h(x) = f'''(\Phi(x))\,\partial_k\Phi(x)\,\partial_l\Phi(x)\,\partial_m\Phi(x). \tag{A.12}$$

We can then substitute this into (A.4) to bound

$$\begin{aligned}
&\Big|\mathrm{E}\Big[X_k\,X_l\,X_m(1 - s\mathbb{1}_{\{m\in A_k\}})\,\sqrt{t}\,\partial_{klm}^3 h\big(W_{s,t,u}^{A_k,B_{kl}}\big)\Big]\Big| \\
&\leqslant \|f'''\|_\infty \left(\mathrm{E}\Big[\big|X_k\,\partial_k\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\,\mathrm{E}\Big[\big|X_l\,\partial_l\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\,\mathrm{E}\Big[\big|X_m\,\partial_m\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\right)^{1/3} \\
&\leqslant \|f'''\|_\infty \sup_{k\in\mathbb{T}} \mathrm{E}[|X_k|^3] \left(\mathrm{E}\Big[\big|\partial_k\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\,\mathrm{E}\Big[\big|\partial_l\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\,\mathrm{E}\Big[\big|\partial_m\Phi\big(W_{s,t,u}^{k,l}\big)\big|^3\Big]\right)^{1/3},
\end{aligned}$$

where we used the fact that $\partial_k\Phi(\dots)$ is independent of X_k since assumption (A.8) implies that $\partial_k\Phi(\dots)$ does not depend on $(X_m)_{m\in A_{kl}}$. This immediately implies (A.9).

The proof of (A.10) is the same if in (A.5), we write $\sigma_{kl} = \mathrm{E}[\widetilde{X}_k \widetilde{X}_l]$ for $(\widetilde{X}_k, \widetilde{X}_l)$ with the same distribution as (X_k, X_l) but independent of X. The proof of (A.11) is even simpler. □

Acknowledgements

We especially thank Te LI for showing us how Hardy-Littlewood-Sobolev type inequalities can be proved without using Fourier transform. R.S. is supported by NUS grant R-146-000-288-114. N.Z. is supported by EPRSC through grant EP/R024456/1.

References

[AKQ14a] T. Alberts, K. Khanin, J. Quastel. The intermediate disorder regime for directed polymers in dimension $1 + 1$. *Ann. Probab.* 42, 1212–1256, 2014.

[AKQ14b] T. Alberts, K. Khanin, J. Quastel. The continuum directed random polymer. *J. Stat. Phys.* 154, 305-326, 2014.

[AFH+92] S. Albeverio, J.E. Fenstad, R. Hoegh-Krohn, W. Karwowski, T. Lindstrom. Schrödinger operators with potentials supported by null sets. *Ideas and methods in quantum and statistical physics*, 63-95, Cambridge Univeristy Press, 1992.

[AGH+05] S. Albeverio, F. Gesztesy, R. Hoegh-Krohn, H. Holden. *Solvable models in quantum mechanics.* AMS Chelsea Publishing, 2005.

[Ba21] E. Bates. Full-path localization of directed polymers. *Electron. J. Probab.* 26, 1-24, 2021.

[Bo89] E. Bolthausen. A note on the diffusion of directed polymers in a random environment. *Comm. Math. Phys.* 123, 529-534, 1989.

[BC98] L. Bertini, N. Cancrini. The two-dimensional stochastic heat equation: renormalizing a multiplicative noise. *J. Phys. A: Math. Gen.* 31, 615, 1998.

[BC20a] E. Bates, S. Chatterjee. The endpoint distribution of directed polymers. *Ann. Probab.* 48, 817-871, 2020.

[BC20b] E. Bates, S. Chatterjee. Localization in Gaussian disordered systems at low temperature. *Ann. Probab.* 48, 2755-2806, 2020.

[BL17] Q. Berger, H. Lacoin. The high-temperature behavior for the directed polymer in dimension $1 + 2$. *Ann. Institut Henri Poincaré, Prob. et Statistiques* 53, 430-450, 2017.

[CES21] G. Cannizzaro, D. Erhard, P. Schönbauer. $2D$ anisotropic KPZ at stationarity: Scaling, tightness and nontriviality. *Ann. Prob.* 49, 122-156, 2021.

[CET20a] G. Cannizzaro, D. Erhard, F.L. Toninelli. Logarithmic superdiffusivity of the 2−dimensional anisotropic KPZ equation. ArXiv:2009.12934, 2020.

[CET20b] G. Cannizzaro, D. Erhard, F.L. Toninelli. The stationary AKPZ equation: logarithmic superdiffusivity. ArXiv:2007.12203, 2020.

[CET21] G. Cannizzaro, D. Erhard, F.L. Toninelli. Weak coupling limit of the Anisotropic KPZ equation. ArXiv:2108.09046, 2021.

[CSZ16] F. Caravenna, R. Sun, N. Zygouras. The continuum disordered pinning model. *Probab. Theory Related Fields* 164, 17–59, 2016.

[CSZ17a] F. Caravenna, R. Sun, N. Zygouras. Polynomial chaos and scaling limits of disordered systems. *J. Eur. Math. Soc.* 19, 1-65, 2017.

[CSZ17b] F. Caravenna, R. Sun, N. Zygouras. Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* 27, 3050-3112, 2017.

[CSZ19a] F. Caravenna, R. Sun, N. Zygouras. The Dickman subordinator, renewal theorems, and disordered systems. *Electron. J. Probab.* 24, paper no. 101, 2019.

[CSZ19b] F. Caravenna, R. Sun, N. Zygouras. On the moments of the (2+1)-dimensional directed polymer and stochastic heat equation in the critical window. *Commun. Math. Phys.* 372, 385-440, 2019.

[CSZ20] F. Caravenna, R. Sun, N. Zygouras. The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Prob.* 48, 1086-1127, 2020.

[CSZ22] F. Caravenna, R. Sun, N. Zygouras. The critical 2d Stochastic Heat Flow is not a Gaussian Multiplicative Chaos. ArXiv:2206.08766, 2022.

[CH02] P. Carmona, Y. Hu, On the partition function of a directed polymer in a Gaussian random environment. *Prob. Th. Rel. Fields*, 124(3), 431-457, (2002).

[Cha06] S. Chatterjee. A generalization of the Lindeberg principle. *Ann. Probab.* 34, 2061-2076, 2006.

[Cha19] S. Chatterjee. Proof of the path localization conjecture for directed polymers. *Comm. Math. Phys.* 370, 703-717, 2019.

[CD20] S. Chatterjee, A. Dunlap. Constructing a solution of the $(2+1)$-dimensional KPZ equation. *Ann. Prob.* 48, 1014-1055, 2020.

[Che21] Y.-T. Chen. The critical 2D delta-Bose gas as mixed-order asymptotics of planar Brownian motion. arXiv:2105.05154, 2021.

[Cla19b] J.T. Clark. The conditional Gaussian multiplicative chaos structure underlying a critical continuum random polymer model on a diamond fractal. ArXiv:1908.08192, 2019.

[Cla21] J. T. Clark. Weak-disorder limit at criticality for directed polymers on hierarchical graphs. *Comm. Math. Phys.* 386, 651-710, 2021.

[Cla22] J.T. Clark. Continuum models of directed polymers on disordered diamond fractals in the critical case. *Ann. Appl. Prob.* 32, 4186-4250, 2022.

[Com17] F. Comets. *Directed Polymers in Random Environments.* Lecture Notes in Mathematics, 2175. Springer, Cham, 2017.

[CCM20] F. Comets, C. Cosco, C.Mukherjee. Renormalizing the Kardar-Parisi-Zhang Equation in $d \geqslant 3$ in weak disorder. *J. Stat. Physics* 179, 713-728, 2020.

[CSY03] F. Comets, T. Shiga, N. Yoshida, Directed polymers in a random environment: path localization and strong disorder. *Bernoulli* 9, 705-723, 2003.

[CY06] F. Comets, N. Yoshida. Directed polymers in random environment are diffusive at weak disorder. *Ann. Probab.* 34, 1746–1770, 2006.

[Cor12] I. Corwin. The Kardar-Parisi-Zhang equation and universality class. *Random Matrices Theory Appl.* 1, 1130001, 2012.

[Cor16] I. Corwin. Kardar-Parisi-Zhang universality. *Notices of the AMS* 63, 230-239, 2016.

[CH16] I. Corwin, A. Hammond. KPZ Line Ensemble. *Probab. Theory Relat. Fields* 166, 67–185, 2016.

[CN21] C. Cosco, S. Nakajima. Gaussian fluctuations for the directed polymer partition function in dimension $d \geqslant 3$ and in the whole L^2-region. *Ann. Inst. H. Poincaré Prob. Stat.* 57, 872-889, 2021.

[CNN22] C. Cosco, S. Nakajima, M. Nakashima. Law of large numbers and fluctuations in the sub-critical and L^2 regions for SHE and KPZ equation in dimension $d \geqslant 3$. *Stochastic Process. Appl.* 151, 127-173, 2022.

[DFT94] G. F. Dell'Antonio, R. Figari, A. Teta. Hamiltonians for systems of N particles interacting through point interactions. *Ann. Inst. H. Poincaré Phys. Théor.* 60, 253-290, 1994.

[DR04] J. Dimock, S. Rajeev. Multi-particle Schrödinger operators with point interactions in the plane. *J. Phys. A: Math. Gen.* 37(39):9157, 2004.

[DGRZ20] A. Dunlap, Y. Gu, L. Ryzhik, O. Zeitouni. Fluctuations of the solutions to the KPZ equation in dimensions three and higher. *Probab. Th. Rel. Fields* 176, 1217-1258, 2020.

[ET60] P. Erdös, S. J. Taylor. Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hungar.* 11, 137-162, 1960.

[F16] Z. S. Feng. Rescaled Directed Random Polymer in Random Environment in Dimension 1+2. Ph.D. thesis, Ann Arbor, MI, 2016. Available at `https://www.proquest.com/docview/1820736587`.

[Ga21] S. Gabriel. Central limit theorems for the (2+1)-dimensional directed polymer in the weak disorder limit. ArXiv:2104.07755, 2021.

[GaSt12] C. Garban, J. Steif. Lectures on noise sensitivity and percolation. Proceedings of the Clay Mathematics Institute Summer School (Buzios, Brazil), Clay Mathematics Proceedings 15, 49-154, 2012.

[GaSu09] J. Gärtner, R. Sun. A quenched limit theorem for the local time of random walks on $\mathbb{Z}^2$. *Stochastic Process. Appl.* 119 , 1198-1215, 2009.

[Gi11] G. Giacomin. *Disorder and critical phenomena through basic probability models.* Lecture Notes in Mathematics, 2025, Springer, Heidelberg, 2011.

[GLT10] G. Giacomin, H. Lacoin, F.L. Toninelli. Marginal relevance of disorder for pinning models. *Comm. Pure Appl. Math* 63, 233-265, 2010.

[GJ14] P. Goncalves, M. Jara, Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.* 212, 597-644, 2014.

[Gu20] Y. Gu. Gaussian fluctuations from the 2D KPZ equation. *Stoch. Partial Differ. Equ. Anal. Comput.* 8, 150-185, 2020.

[GQT21] Y. Gu, J. Quastel, L.-C. Tsai. Moments of the 2D SHE at criticality. *Prob. Math. Phys.* 2, 179-219, 2021.

[GRZ18] Y. Gu, L. Ryzhik, O. Zeitouni. The Edwards-Wilkinson limit of the random heat equation in dimensions three and higher. *Comm. Math. Phys.* 363, 351–388, 2018.

[GIP15] M. Gubinelli, P. Imkeller, N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi* 3, e6, 2015.

[GP17] M. Gubinelli, N. Perkowski. KPZ reloaded. *Comm. Math.Phys.* 349, 165-269, 2017.

[H13] M. Hairer. Solving the KPZ equation. *Ann. of Math.* 178, 559-664, 2013.

[H14] M. Hairer. A theory of regularity structures. *Inventiones Math.* 198, 269-504, 2014.

[HH12] T. Halpin-Healy. $(2+1)$-dimensional directed polymer in a random medium: scaling phenomena and universal distributions. *Phys. Rev. Lett.* 109, 170602, 2012.

[HH13] T. Halpin-Healy. Extremal paths, the stochastic heat equation, and the three-dimensional Kardar-Parisi-Zhang universality class. *Physical Review E* 88, 042118, 2013.

[HH85] D.A. Huse, C.L. Henley. Pinning and roughening of domain walls in Ising systems due to random impurities. *Phys. Rev. Lett.* 54, 2708-2711, 1985.

[IS88] J. Z. Imbrie, T. Spencer. Diffusion of directed polymers in a random environment. *J. Stat. Physics* 52, 609-626, 1988.

[J00] K. Johansson. Transversal fluctuations for increasing subsequences on the plane. *Probab. Theory Related Fields* 116, 445-456, 2000.

[Kal97] O. Kallenberg. *Foundations of modern probability.* Springer, 1997.

[Koz07] G. Kozma. The scaling limit of loop-erased random walk in three dimensions. *Acta Math.* 199, 29-152, 2007.

[Kup14] A. Kupiainen. Renormalization Group and Stochastic PDEs. *Ann. H. Poincaré* 17, 497-535, 2016.

[L10] H. Lacoin. New bounds for the free energy of directed polymers in dimension $1+1$ and $1+2$. *Comm. Math. Physics* 294, 471-503, 2010.

[LaLi10] G.F. Lawler, V. Limic. *Random walk: a modern introduction.* Cambridge University Press, 2010.

[LiLo01] E.H. Lieb, M. Loss. *Analysis*, 2nd ed. Graduate Studies in Mathematics 14, American Mathematical Society, 2001.

[LZ22] D. Lygkonis, N. Zygouras. Edwards-Wilkinson fluctuations for the directed polymer in the full L^2-regime for dimensions $d \geqslant 3$. *Ann. Inst. Henri Poincaré Probab. Stat.* 58, 65-104, 2022.

[MU18] J. Magnen, J. Unterberger. The scaling limit of the KPZ equation in space dimension 3 and higher. *J. Stat. Physics* 171, 543-598, 2018.

[MOO10] E. Mossel, R. O'Donnell, K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.* 171, 295-341, 2010.

[MSZ16] C. Mukherjee, A. Shamov, O. Zeitouni. Weak and strong disorder for the stochastic heat equation and continuous directed polymers in $d \geqslant 3$. *Electron. Comm. Prob.* 21, 1-12, 2016.

[N19] M. Nakashima. Free energy of directed polymers in random environment in $1+1$-dimension at high temperature. *Electron. J. Probab.* 24, 1-43, 2019.

[QS15] J. Quastel, H. Spohn. The One-dimensional KPZ equation and tts universality class. *J. Stat. Physics* 160, 965-984, 2015.

[Raj99] S. G. Rajeev, A condensation of interacting Bosons in two dimensional space, arXiv preprint hep-th/9905120, (1999).

[Rol13] A. Röllin. Stein's method in high dimensions with applications. *Ann. Inst. Henri Poincaré Probab. Stat.* 49, 529-549, 2013.

[Rot79] V.I. Rotar. Limit theorems for polylinear forms. *J. Multivariate Anal.* 9, 511-530, 1979.

[V07] V. Vargas. Strong localization and macroscopic atoms for directed polymers. *Probab. Theory Related Fields* 138, 391-410, 2007.

Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, via Cozzi 55, 20125 Milano, Italy
E-mail address: francesco.caravenna@unimib.it

Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, 119076 Singapore
E-mail address: matsr@nus.edu.sg

Department of Mathematics, University of Warwick, Coventry CV4 7AL, UK
E-mail address: N.Zygouras@warwick.ac.uk