Commun. Math. Phys. 401, 435–496 (2023)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-023-04641-9

Communications in
Mathematical Physics

Historical Lattice Trees

Manuel Cabezas[1], **Alexander Fribergh**[2], **Mark Holmes**[3], **Edwin Perkins**[4]

[1] Pontificia Universidad Católica de Chile, Santiago, Chile. E-mail: cabezas.mn@gmail.com
[2] Université de Montréal, Montreal, Canada. E-mail: alexander.fribergh@umontreal.ca
[3] School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia. E-mail: holmes.m@unimelb.edu.au
[4] Mathematics Department, The University of British Columbia, Vancouver, Canada. E-mail: perkins@math.ubc.ca

Received: 17 February 2022 / Accepted: 7 January 2023
Published online: 31 January 2023 – © The Author(s) 2023

Abstract: We prove that the rescaled historical processes associated to critical spread-out lattice trees in dimensions $d > 8$ converge to historical Brownian motion. This is a functional limit theorem for measure-valued processes that encodes the genealogical structure of the underlying random trees. Our results are applied elsewhere to prove that random walks on lattice trees, appropriately rescaled, converge to Brownian motion on super-Brownian motion.

1. Introduction and Main Results

In the past three decades, various critical high-dimensional spatial branching models have been conjectured or proved to converge to super-Brownian motion (SBM), which is a continuous Markov process taking values in the space of finite measures on $\mathbb{R}^d$. One of the settings in which significant progress has been made is that of critically weighted (and sufficiently spread-out) lattice trees (LT) above 8 dimensions [7,10,11,17–19]. In particular, convergence on path space has recently been proved in this setting (see [11]). For LT's convergence to SBM means weak convergence to SBM of the rescaled empirical measure process of the locations in the LT which are a given tree distance from the root. Hence the tree distance to the root plays the role of time for the stochastic processes. More recently, it has been proved in [20] that for LT's, and in fact for several lattice models, the rescaled ranges (for LT's the range is the compact set of vertices in the tree) converge weakly to the range of SBM. Convergence of genealogical observables is not forthcoming from the notions of weak convergence to SBM described thus far. Results of this kind can be obtained by proving convergence of the corresponding *"historical processes"* [6]. For LT's this would mean that instead of just having the convergence to SBM of the rescaled empirical measure process of the particles in the LT, as a function of the distance from the root, one establishes convergence to historical Brownian motion (HBM) of the rescaled empirical measure process for the entire paths in the LT to the endpoints, as a function of the distance from the root. HBM, constructed in [6], is a

process taking values in the space of finite measures on $\mathbb{R}^d$-valued paths, which at time t is the empirical measure of the past histories of the particles contributing to the SBM at time t. See Sect. 1.2.1 below for more about HBM, including the fact that is the weak limit of the rescaled historical processes associated with Branching Brownian Motion (Theorem 1.3). Our main result, Theorem 1.4 below, establishes this convergence of "historical processes" for LT's.

In Sect. 2.1 we give a set of general conditions that are sufficient for convergence of discrete-time historical processes to HBM in the sense of finite-dimensional distributions (Theorem 2.1). Most of these conditions are already known to hold for a range of lattice models above the critical dimension including lattice trees ($d > 8$) and oriented percolation ($d > 4$), as well as for the voter model ($d \geq 2$) and the contact process ($d > 4$), both of which are continuous-time models. The main condition that remains to be proved in each case is convergence of the joint characteristic functions of the increments of a finite dimensional subtree. These *detailed r-particle transforms* can be seen as enriched versions of the r-particle transforms studied e.g. in [13,16,17] (called Fourier transforms of $(r+1)$-point functions therein) that record genealogy. We prove that these conditions are satisfied for sufficiently spread-out lattice trees in high dimensions and so establish convergence to HBM in the sense of f.d.d.'s (Proposition 2.4). The required asymptotics of the detailed r-particle transforms are obtained via the *lace expansion* (see e.g. [25]) in Sect. 4. It is worth noting that these asymptotics can be understood from those of the usual r-particle transforms and the detailed 1-particle transform. In particular we do not require any new "diagrammatic estimates". We believe that all of the conditions can also be verified for the other models[1] mentioned above. For the voter model this is currently work in progress [1].

The second main ingredient in our proof is a novel tightness argument for historical processes which upgrades f.d.d. convergence to convergence on path space in a historical setting. This step is carried out in Sect. 3. We start with an abstract tightness result in a general historical setting (Theorem 3.6). For all of the lattice models mentioned above this reduces tightness of the approximating rescaled historical processes to that of the $\mathbb{R}$-valued processes obtained by integrating a test function (from an appropriate class) with respect to the rescaled historical processes. (Verification of the other conditions may be found in [20].) This key condition is then verified for LT's with some effort in Proposition 3.11. The main ingredients of this argument are tightness of the total mass process from [11] and a uniform modulus of continuity for the approximating historical paths from [20]. The latter is in fact verified for all of the other lattice models mentioned above, and so we have potentially reduced the problem of tightness for historical processes to that of the total mass process for a range of other lattice models.

A simple consequence of our results is that the unique path in the tree from the origin to a uniformly chosen vertex (called the *backbone* from the origin to that vertex) of distance n converges weakly to BM on path space (see [18, Theorem 1.3]). Another application of our results concerns the scaling limit of random walk on lattice trees. In particular, the historical convergence proved herein is used in [21] to verify certain conditions of Ben-Arous et al. [2] which imply that random walk on lattice trees converges to a BM on a SBM cluster.

[1] For oriented percolation, one should choose a notion of historical paths e.g. each site in the cluster of the origin chooses its parent uniformly at random from all possible parents in the cluster.

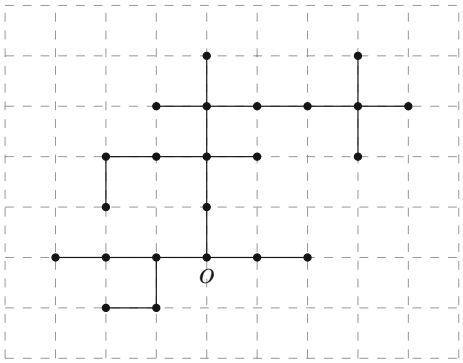

Fig. 1. A (nearest neighbour) lattice tree in 2 dimensions

1.1. Lattice trees and scaling limits. A lattice tree is a finite connected set of lattice bonds containing no cycles (see Fig. 1).

We will be considering lattice trees on $\mathbb{Z}^d$ with bonds connecting any two vertices that live in a common ball (in ℓ_∞) of sufficiently large radius $L \in \mathbb{N}$, and with $d > 8$. To be more precise, let $d > 8$ and let $D(\cdot)$ be the uniform distribution on a finite box $([-L, L]^d \setminus o) \cap \mathbb{Z}^d$, where $o = (0, \ldots, 0) \in \mathbb{Z}^d$. The assumption of uniformity of D is not essential. We expect that the results herein hold for D as in [17, Section 1].

For a lattice tree $T \ni o$ define $W_{z,D}(T) = z^{|T|} \prod_{e \in T} D(e)$, where the product is over the edges in T and $|T|$ is the number of edges in T.

Remark 1.1. If T is an edge-disjoint union of subtrees then $W_{z,D}(T)$ can be factored into a product over the weights of the subtrees. ★

It turns out (see e.g. [10,17]) that there exists a unique critical value z_D such that $\rho = \sum_{T \ni o} W_{z_D,D}(T) < \infty$ and $\mathbb{E}[|\mathcal{T}|] = \infty$, where $\mathbb{P}(\mathcal{T} = T) = \rho^{-1} W_{z_D,D}(T)$ for $T \ni o$. Hereafter we write $W(\cdot)$ for the critical weighting $W_{z_D,D}(\cdot)$ and suppose that we are selecting a random tree $\mathcal{T} \ni o$ according to this critical weighting.

Let T be a lattice tree containing $o \in \mathbb{Z}^d$, and for $m \in \mathbb{N}$, let T_m denote the set of vertices in T of tree distance m from o. In particular, $T_0 = \{o\}$, and for any $x \in T_m$ there is a unique path from o to x in the tree of length m. Roughly speaking, in this paper we consider the weak limit (as $m \to \infty$) of rescaled paths of this kind in high dimensions. For $t \in \mathbb{R}_+ \setminus \mathbb{Z}_+$ define $T_t = T_{\lfloor t \rfloor}$. For $t \geq 0$ and $x \in \mathbb{Z}^d$ we will write $(t, x) \in T$ to mean that $x \in T_t$. The notation (t, x) is consistent with that in [20], while in the oriented percolation and contact process literature often (x, t) is used instead.

Functional limit theorems. For our general discussion we require the notion of weak convergence of finite measures on Polish (i.e. complete, separable metric) spaces. We refer the reader to [8, Chapter 3] for further details on what we discuss below.

For a Polish space $\mathfrak{P}$, let $\mathcal{M}_F(\mathfrak{P})$ (resp. $\mathcal{M}_1(\mathfrak{P})$) denote the space of finite (resp. probability) measures on the Borel sets of $\mathfrak{P}$. For a sequence $\nu_n \in \mathcal{M}_F(\mathfrak{P})$ we say that ν_n converges weakly to $\nu \in \mathcal{M}_F(\mathfrak{P})$ and write $\nu_n \xrightarrow{w} \nu$ if for every $f : \mathfrak{P} \to \mathbb{R}$ bounded and continuous,

$$\int f(x)\nu_n(dx) \to \int f(x)\nu(dx), \qquad \text{as } n \to \infty.$$

Equipped with the Vasershtein metric, which generates the topology of weak convergence, $\mathcal{M}_F(\mathfrak{P})$ is also Polish (see e.g., [24, Ch. II]). We will use the notation $E_\nu[f(X)]$ for $\int f(x)\nu(dx)$, with the understanding that $X \in \mathfrak{P}$. This will be particularly convenient when X is a $\mathfrak{P}$-valued random variable defined on an underlying probability space and $\nu(\cdot) = c \cdot \mathbb{P}(X \in \cdot)$ for some $c > 0$.

Let S_n denote the location of a nearest-neighbour simple symmetric random walk on $\mathbb{Z}^d$ after n steps (starting from the origin $o \in \mathbb{Z}^d$). Then $\mathbb{E}[S_n^2] = n$ (here and elsewhere, for $x, y \in \mathbb{R}^d$ we abuse notation and write xy to mean $x \cdot y$, and hence x^2 to mean $|x|^2$) and the central limit theorem (CLT) states that $n^{-1/2}S_n$ converges in distribution to a random vector Z that is (multivariate-) normally distributed with mean $0 \in \mathbb{R}^d$ and covariance matrix $\mathrm{diag}(1/d)$. Define probability measures ν_n, ν on (the Borel sets of) $\mathbb{R}^d$ by

$$\nu_n(\cdot) = \mathbb{P}\big(n^{-1/2}S_n \in \cdot\big), \quad \text{and} \quad \nu(\cdot) = \mathbb{P}(Z \in \cdot).$$

Phrased in the language of weak convergence of (finite) measures, the CLT says that $\nu_n \xrightarrow{w} \nu$. The statement $\nu_n \xrightarrow{w} \nu$ in $\mathcal{M}_F(\mathbb{R}^d)$ is well known to be equivalent to pointwise convergence of the characteristic functions (Fourier transforms), so for ν_n, ν as above

$$\int \mathrm{e}^{\mathrm{i}kx}\nu_n(dx) \to \int \mathrm{e}^{\mathrm{i}kx}\nu(dx) = \mathrm{e}^{-\frac{k^2}{2d}}, \text{ for } k \in \mathbb{R}^d.$$

For a Polish space $\mathfrak{P}$ let $\mathcal{D}_t(\mathfrak{P})$ (resp. $\mathcal{D}(\mathfrak{P})$) denote the space of càdlàg paths (paths that are continuous from the right with limits existing from the left) mapping $[0, t]$ (resp. $[0, \infty)$) to $\mathfrak{P}$. Let $\mathcal{C}_t(\mathfrak{P})$ (resp. $\mathcal{C}(\mathfrak{P})$) denote the corresponding subspace of continuous paths. It is well known that there are complete metrics on these spaces (generating the Skorokhod J_1 topology) for which $\mathcal{D}_t(\mathfrak{P})$ and $\mathcal{D}(\mathfrak{P})$ are also Polish (see [8, Chapter 3.5]). The *functional* central limit theorem (FCLT) concerns the entire path $(W_t^{(n)})_{t\geq 0}$ defined by

$$W_t^{(n)} = n^{-1/2}S_{\lfloor nt \rfloor}.$$

Defined in this way, for each n, $W^{(n)}$ jumps at times $t = i/n$ for $i \in \mathbb{N}$ and is constant on intervals $[i/n, i + 1/n)$ for $i \in \mathbb{Z}_+$. In particular the process $W^{(n)}$ is a random element of the space $\mathcal{D}(\mathbb{R}^d)$ of càdlàg paths from $\mathbb{R}_+ = [0, \infty)$ to $\mathbb{R}^d$. The FCLT states that the sequence of rescaled random walks $(W_t^{(n)})_{t\geq 0}$ converges to a d-dimensional Brownian motion $(B_t)_{t\geq 0}$ (with $B_1 \sim \mathcal{N}(0, \mathrm{diag}(1/d))$). Phrased in the language of weak convergence of (probability) measures this FCLT says that $\nu_n \xrightarrow{w} \nu$, where $\nu_n, \nu \in \mathcal{M}_1(\mathcal{D}(\mathbb{R}^d))$ are defined by

$$\nu_n(\cdot) = \mathbb{P}\big((W_t^{(n)})_{t\geq 0} \in \cdot\big), \qquad \nu(\cdot) = \mathbb{P}\big((B_t)_{t\geq 0} \in \cdot\big).$$

Note that ν puts all its mass on continuous paths.

Paths and measure-valued processes for lattice trees. For $(m, x) \in \mathcal{T}$ let $w(m, x) = (w_0(m, x) = o, w_1(m, x), \ldots, w_m(m, x) = x)$ denote the unique path in $\mathcal{T}$ from o to x in the tree. It is convenient to extend this to a function on $\mathbb{Z}_+$ and then to a function in $\mathcal{D}$ by writing

$$w_n(m, x) := w_m(m, x) = x, \text{ for } n \geq m, \qquad w_s(m, x) = w_{\lfloor s \rfloor}(m, x), \quad \text{for } s \in [0, \infty). \tag{1.1}$$

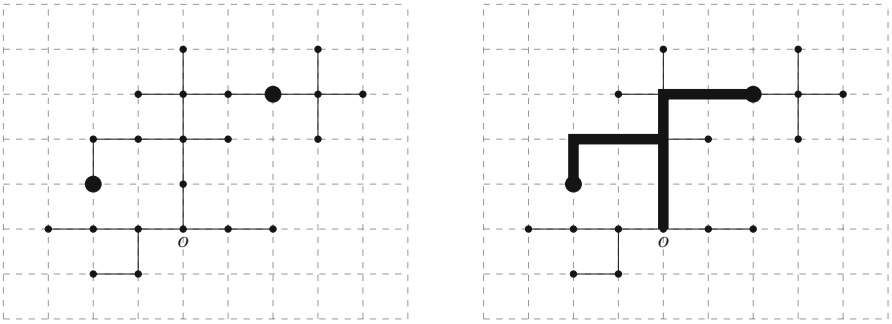

Fig. 2. The MVP $X_5^{(1)}$ assigns masses to points in the tree at distance 5 from the root, while $H_5^{(1)}$ assigns the same masses to paths in the tree leading to these points

Thus every $(m, x) \in \mathcal{T}$ has associated to it an infinite càdlàg path $w(m, x)$ that is constant after time m. Denote the collection of ancestral paths for $\mathcal{T}$ by $\mathcal{W} = (w(m, x))_{(m,x)\in\mathcal{T}}$. For $t \geq 0$ and $x \in \mathbb{Z}^d/\sqrt{n}$ such that $\sqrt{n}x \in \mathcal{T}_{nt}$ we define $w^{(n)}(t, x) \in \mathcal{D}$ by

$$w_s^{(n)}(t, x) = \frac{w_{ns}(\lfloor nt \rfloor, \sqrt{n}x)}{\sqrt{n}}, \quad \text{for } s \in [0, \infty). \tag{1.2}$$

By [10,17] there exist constants $C_A, C_V > 0$[2] such that

$$\lim_{n\to\infty} \mathbb{E}[|\mathcal{T}_n|] = C_A, \text{ and } \lim_{n\to\infty} n\mathbb{P}(|\mathcal{T}_n| > 0) = 2/(C_A C_V). \tag{1.3}$$

Let $C_0 = C_A^2 C_V$, and let

$$X_t^{(n)} = \frac{1}{C_0 n} \sum_{\sqrt{n}x\in\mathcal{T}_{nt}} \delta_x \in \mathcal{M}_F(\mathbb{R}^d), \qquad \text{and}$$
$$H_t^{(n)} = \frac{1}{C_0 n} \sum_{\sqrt{n}x\in\mathcal{T}_{nt}} \delta_{w^{(n)}(t,x)} \in \mathcal{M}_F(\mathcal{D}(\mathbb{R}^d)) \tag{1.4}$$

denote the (rescaled) measure-valued "process" and historical "process" (see e.g. [6]) associated with the random lattice tree $\mathcal{T}$ respectively. Note that $X_t^{(n)}$ assigns mass to certain particles in the tree (but does not encode the genealogy) whereas $H_t^{(n)}$ assigns mass to genealogical paths leading to those particles. See e.g. Fig. 2.

For $\phi : \mathfrak{P} \to \mathbb{C}$ and $Y_t \in \mathcal{M}_F(\mathfrak{P})$ write $Y_t(\phi) = \int \phi dY_t$. Then for $\phi : \mathbb{R}^d \to \mathbb{C}$ we have

$$\int \phi(w_t) dH_t^{(n)}(dw) = X_t^{(n)}(\phi), \tag{1.5}$$

and in particular

$$H_t^{(n)}(1) \equiv X_t^{(n)}(1).$$

Define the survival/extinction time as

$$S^{(n)} := \inf\{t > 0 : X_t^{(n)}(1) = 0\} = \inf\{t > 0 : H_t^{(n)}(1) = 0\}.$$

[2] Our constant C_A is equal to $A' = A/\rho$ from [17] and our constant C_V is $V\rho^2$ from [17].

Let $C_1 = C_A C_V$ so that from (1.3),

$$n\mathbb{P}(H_t^{(n)}(1) > 0) = n\mathbb{P}(S^{(n)} > t) \to \frac{2}{C_1 t}, \qquad \text{as } n \to \infty. \tag{1.6}$$

Then we define $\nu_n^{\mathrm{LT}} \in \mathcal{M}_F(\mathcal{D}(\mathcal{M}_F(\mathbb{R}^d)))$ by

$$\nu_n^{\mathrm{LT}}(\bullet) = nC_1\mathbb{P}(X^{(n)} \in \bullet),$$

and $\mu_n^{\mathrm{LT}} \in \mathcal{M}_F(\mathcal{D}(\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))))$ by

$$\mu_n^{\mathrm{LT}}(\bullet) = nC_1\mathbb{P}(H^{(n)} \in \bullet). \tag{1.7}$$

Due to the survival probability asymptotics (1.6), multiplying by n and working on the event that the process survives until time n is asymptotically the same (up to a constant) as conditioning on survival until time n (or rescaled time 1).

According to [24, Section II.7], for any $\gamma, \sigma^2 > 0$ (representing the branching rate and diffusion parameter respectively) there exists a σ-finite measure $\mathbb{N} = \mathbb{N}^{\gamma,\sigma^2}$ on $\mathcal{C}(\mathcal{M}_F(\mathbb{R}^d))$, with $\mathbb{N}(X_t(1) > 0) = 2/(\gamma t)$ such that $\mathbb{N}$ is the canonical measure associated to the $((B_t)_{t\geq 0}, \gamma, 0)$-superprocess. Here $(B_t)_{t\in[0,\infty)}$ is a d-dimensional BM with $B_1 \sim \mathcal{N}(0, \sigma^2 I_{d\times d})$, which is a (time-homogeneous) Markov process. The superprocess in question (called super-Brownian motion) is a measure-valued process that can be thought of as the empirical measures of an infinitesimal critical branching process whose spatial dispersion is governed by the $\mathbb{R}^d$-valued process $(B_t)_{t\geq 0}$. If $S = \inf\{t > 0 : X_t(1) = 0\}$, then $\mathbb{N}$ is supported on $\{X \in \mathcal{C}(\mathcal{M}_F(\mathbb{R}^d)) : X_0 = 0, S > 0, X_t = 0\ \forall t \geq S\}$, and so the above implies that

$$\mathbb{N}(S > t) = 2/(\gamma t). \tag{1.8}$$

By replacing the Markov process $(B_t)_{t\geq 0}$ with the path-valued (time-inhomogeneous) Markov process $(B_{[0,t]})_{t\geq 0} \equiv \big((B_s)_{s\in[0,t]}\big)_{t\geq 0}$, and using the general theory of superprocesses, there also exists a σ-finite measure $\mathbb{N}_H = \mathbb{N}_H^{\gamma,\sigma^2}$ on $\mathcal{C}(\mathcal{M}_F(\mathcal{C}(\mathbb{R}^d)))$ with $\mathbb{N}_H(H_t(1) > 0) = \mathbb{N}(S > t)$ such that $\mathbb{N}_H$ is the canonical measure associated to the $((B_{[0,t]})_{t\geq 0}, \gamma, 0)$-superprocess. The latter (as well as the process H underlying $\mathbb{N}_H$) is called historical Brownian motion (HBM). The general construction of canonical measures for superprocesses may be found in [24, Section II.7], while Section II.8 therein shows how to consider the historical processes in this general framework. One can also construct $\mathbb{N}_H$ from the canonical measure of Le Gall's Brownian snake since the historical process is a functional of the snake. See [22, pages 34, 64] for details.

It is proved in [11,17] that for lattice trees in dimensions $d > 8$ (with L sufficiently large) $\nu_n^{\mathrm{LT}} \xrightarrow{w} \mathbb{N}$, where the parameters of $\mathbb{N}$ are $\gamma = 1$ and $\sigma_0^2 = \sigma_0^2(L, d)$, which is to be discussed later. Since the limit is a σ-finite measure, $\nu_n \xrightarrow{w} \mathbb{N}$ is defined in terms of weak convergence of a family of finite measures (indexed by t) on $\mathcal{D}(\mathcal{M}_F(\mathbb{R}^d))$ as

$$\nu_n(\bullet, S^{(n)} > t) \xrightarrow{w} \mathbb{N}(\bullet, S > t), \qquad \text{for each } t > 0, \tag{1.9}$$

or equivalently in terms of weak convergence of their conditional (on $S > t$) counterparts, which are probability measures. (The equivalence holds by (1.6), (1.7) and (1.8).) Similar results have been proved for other self-interacting branching systems such as the voter model [3,4] ($d \geq 2$), oriented percolation (OP) [16] ($d > 4$), and the contact

process (CP) [13] ($d > 4$)), although for OP and CP only convergence of the finite-dimensional distributions has been established and tightness remains an open problem. The corresponding result for the historical processes ($\mu_n \xrightarrow{w} \mathbb{N}_H$) was an open problem in all of the above contexts. Here we resolve this problem for lattice trees ($d > 8$, and L sufficiently large[3]), and, as was suggested above, our general approach may well also help in the other contexts above. A discussion of possible extensions and challenges for other models, including these, may be found in Sect. 1.3.

1.2. Main results. In this section we state our main result (Theorem 1.4 below). For this, we first introduce some notation and present the relevant notions of weak convergence. We then introduce critical branching Brownian motion (BBM) as a simpler process from which one can understand the limiting historical Brownian motion through a corresponding historical limit theorem for rescaled BBM's, see Theorem 1.3. The latter follows easily from results in the literature as we will describe. Following this, we state our main result. Theorem 1.3 is also used in the proof of our main result by identifying the joint characteristic functions of the general moment measures for the limiting HBM in Proposition 2.6.

For a Polish space $\mathfrak{P}$, and $\boldsymbol{x} = (x_t)_{t\geq 0} \in \mathcal{D}(\mathcal{M}_F(\mathfrak{P}))$, let $S(\boldsymbol{x}) = \inf\{t > 0 : x_t(\mathfrak{P}) = 0\}$. Let $\mathcal{M}^{\mathrm{EX}}(\mathfrak{P})$ (resp. $\mathcal{M}_1^{\mathrm{EX}}(\mathfrak{P})$) denote the set of σ-finite (resp. probability) measures μ on $\mathcal{D}(\mathcal{M}_F(\mathfrak{P}))$ such that

(1) $\mu\big(\{\boldsymbol{x} : S(\boldsymbol{x}) > s\}\big) \in (0, \infty)$ for each $s > 0$ and $\mu(\{\boldsymbol{x} : S(\boldsymbol{x}) = \infty\}) = 0$, and
(2) $\mu\big(\{\boldsymbol{x} : x_t(\mathfrak{P}) > 0 \text{ for some } t > S(\boldsymbol{x})\}\big) = 0$.

One should think of $\mathcal{M}_1^{\mathrm{EX}}(\mathfrak{P})$ as the space of excursion measures for càdlàg measure-valued paths where the measures are on $\mathfrak{P}$. For $\mu \in \mathcal{M}^{\mathrm{EX}}(\mathfrak{P})$, and $s > 0$ define the (probability) measure μ^s on $\mathcal{D}(\mathcal{M}_F(\mathfrak{P}))$ to be μ conditional on $S > s$, i.e.

$$\mu^s(\bullet) = \frac{\mu(\bullet, \{\boldsymbol{x} : S(\boldsymbol{x}) > s\})}{\mu(\{\boldsymbol{x} : S(\boldsymbol{x}) > s\})}.$$

For $r \in \mathbb{N}$ and $\vec{t} = (t_1, \ldots, t_r) \in [0, \infty)^r$ and a finite measure κ on $\mathcal{D}(\mathcal{M}_F(\mathfrak{P}))$, let $\kappa_{\vec{t}}$ denote the (finite) measure on $(\mathcal{M}_F(\mathfrak{P}))^r$ defined by

$$\kappa_{\vec{t}}(\bullet) = \kappa(\{\boldsymbol{x} : (x_{t_i})_{i=1}^r \in \bullet\}).$$

Definition 1.2 (*Weak convergence*). Fix a sequence $(\mu_n)_{n\in\mathbb{N}\cup\{\infty\}}$ in $\mathcal{M}^{\mathrm{EX}}(\mathfrak{P})$.

- We write $\mu_n \xrightarrow{w} \mu_\infty$ as $n \to \infty$ if for every $s > 0$, $\mu_n(S > s) \to \mu_\infty(S > s)$ and
$$\mu_n^s \xrightarrow{w} \mu_\infty^s, \quad \text{in } \mathcal{M}_1\big(\mathcal{D}(\mathcal{M}_F(\mathfrak{P}))\big).$$
- We write $\mu_n \xrightarrow{f.d.d.} \mu_\infty$ if for every $s > 0, r \in \mathbb{N}$ and $\vec{t} \in \mathbb{R}_+^r$, we have $\mu_n(S > s) \to \mu_\infty(S > s)$, and
$$\mu_{n,\vec{t}}^s \xrightarrow{w} \mu_{\infty,\vec{t}}^s, \quad \text{in } \mathcal{M}_1\big((\mathcal{M}_F(\mathfrak{P}))^r\big).$$

◀

[3] Super-Brownian motion is not expected to arise as the scaling limit for $d < 8$.

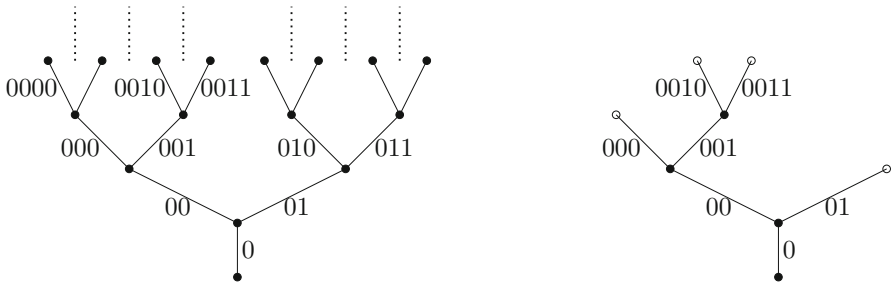

Fig. 3. On the left is the index set I drawn (with labels as edges) up to and including generation 3. On the right is an example of a Galton–Watson tree (with edge labels α), where $e^\alpha = 0$ for all $\alpha \in \{000, 0010, 0011, 01\}$, while $e^\alpha = 2$ for $\alpha \in \{0, 00, 001\}$. Note that we have dropped the parentheses and commas in the notation for elements of I to declutter the pictures

1.2.1. Branching Brownian motion. A good way to understand historical Brownian motion is as a limit of critical branching Brownian motions. Recall that branching Brownian motion may be viewed as a system of Brownian motions run along the edges of a critical Galton–Watson tree. The notation introduced below is presented in [24] at a more leisurely pace. We start by defining a Brownian motion on a full binary tree. Let

$$I = \{\alpha = (\alpha_0, \ldots, \alpha_n) : \alpha_0 = 0, \alpha_i \in \{0, 1\} \text{ for } i > 0, n \in \mathbb{Z}_+\}, \tag{1.10}$$

and for α as above set $|\alpha| = n$, $\alpha|i = (\alpha_0, \ldots, \alpha_i)$, $i \le n$, and say β is an ancestor of α iff $\beta = \alpha|i$ for some $i < |\alpha|$. If $\alpha, \beta \in I$, the greatest common antecedent (gca) of α and β is $\alpha \wedge \beta = \alpha|i$, where i is the maximal integer such that $\alpha|i = \beta|i$. If $|\alpha| > 0$, the parent of α is $\pi\alpha := \alpha(|\alpha| - 1)$.

Let $\{W^\alpha : \alpha \in I\}$ be iid d-dimensional Brownian motions with variance parameter σ^2. For a fixed $n \in \mathbb{N}$ (dependence on n is suppressed) and for $\alpha \in I$, let

$$\hat{B}_t^\alpha = \sum_{i=0}^{|\alpha|} \int_0^t \mathbb{1}_{\{s \in [i/n, (i+1)/n)\}} dW_s^{\alpha|i},$$

and note that $(\hat{B}_t^\alpha)_{t \ge 0}$ is a d-dimensional Brownian motion, starting at 0, that runs until time $(|\alpha|+1)/n$ (after which it stays constant). We can view $\{\hat{B}_t^\alpha : t < (|\alpha|+1)/n, \alpha \in I\}$ as a Brownian motion run on a rescaled binary tree with edge lengths $1/n$. We next prune the binary tree to make it a critical Galton–Watson (G–W) tree. Let $\{e^\alpha : \alpha \in I\}$ be a collection of iid random variables with (critical) binary offspring law $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$ that is independent of $\{W^\alpha : \alpha \in I\}$. For a fixed $n \in \mathbb{N}$ (dependence on n is suppressed) and for $\alpha \in I$, let

$$\tau^\alpha = \min\left\{\frac{i+1}{n} : e^{\alpha|i} = 0\right\} \quad (\min \emptyset = \frac{|\alpha|+1}{n}),$$

and also define

$$B_t^\alpha = \begin{cases} \hat{B}_t^\alpha, & \text{if } t < \tau^\alpha, \\ \Delta, & \text{if } t \ge \tau^\alpha. \end{cases}$$

Here Δ is added to $\mathbb{R}^d$ as a cemetery point. In this way $GW = \{\alpha : \tau^\alpha = \frac{|\alpha|+1}{n}\}$ labels the points (drawn as edges in Fig. 3) on a G–W tree with a critical binary offspring law

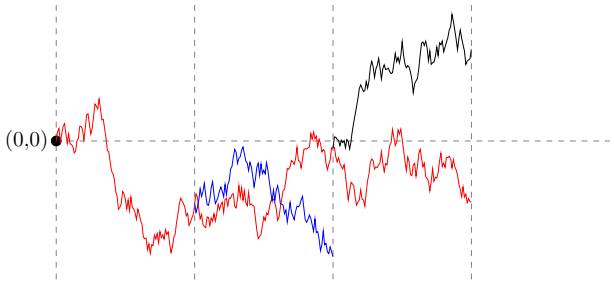

Fig. 4. A (binary) branching Brownian motion in 1-dimension, with time on the x axis, drawn up to the third branch time, $3/n$. In the corresponding G–W tree, the root 0 has two children, exactly one of which has 2 children

that does not depend on n. We have scaled the edge lengths of the tree to be n^{-1} and write $\alpha \sim t$ iff $\alpha \in GW$ and $\frac{|\alpha|}{n} \leq t < \frac{|\alpha|+1}{n}$. Therefore $\alpha \sim t$ means that α labels an edge in the Galton–Watson tree which is alive at time $t \geq 0$. In particular, $0 \sim t$ for every $t < 1/n$, see Fig. 3. Finally $\{B_t^\alpha : \alpha \sim t\}$ for $t \geq 0$ is a system of Brownian motions, starting with a single particle at the origin, and run along these edges while undergoing critical binary branching at times $\{j/n : j \in \mathbb{N}\}$, with the motions being independent along the disjoint scaled edges in the G–W tree. Figure 4 gives a depiction of the system of Brownian motions in 1-dimension.

We define the scaled empirical measures $X_\cdot^{(n)} \in \mathcal{D}(\mathcal{M}_F(\mathbb{R}^d))$ and $H_\cdot^{(n)} \in \mathcal{D}(\mathcal{M}_F(\mathcal{C}(\mathbb{R}^d)))$ associated with these locations and historical paths, respectively, by

$$X_t^{(n)} = \frac{1}{n}\sum_{\alpha\sim t} \delta_{B_t^\alpha}, \quad H_t^{(n)} = \frac{1}{n}\sum_{\alpha\sim t} \delta_{B_{\cdot\wedge t}^\alpha}, \quad t \geq 0.$$

It is easy to extend the above definitions to the setting of a general mean 1 finite variance γ offspring law in place of the critical binary branching law above where we have $\gamma = 1$ (see [24, Section II.3]). In this setting let $\nu_n^{\mathrm{BBM}} = n\mathbb{P}(X^{(n)} \in \cdot)$ and $\mu_n^{\mathrm{BBM}} = n\mathbb{P}(H^{(n)} \in \cdot)$. We believe that the following limit result was first proved in [24], although part (b) was not stated explicitly there. The original construction of $\mathbb{N} = \mathbb{N}^{\gamma,\sigma^2}$ was done by Le Gall using his Brownian snake (see [22, Ch. IV] and the references therein) from which the result below was clear enough.

Theorem 1.3. *As $n \to \infty$,*

(a) $\nu_n^{\mathrm{BBM}} \xrightarrow{w} \mathbb{N}^{\gamma,\sigma^2}$;

(b) $\mu_n^{\mathrm{BBM}} \xrightarrow{w} \mathbb{N}_H^{\gamma,\sigma^2}$.

Proof. (a) is a special case of [24, Theorem II.7.3]. We also use Kolmogorov's classical result on survival asymptotics for critical branching processes (eg. [24, Theorem II.1.1]).

(b) also follows from the same results, where [24, Section II.8] explains how to put the historical setting into the general framework of [24, Theorem II.7.3]. □

An easy consequence of the above and the obvious analogue of (1.5) for branching Brownian motion is that H projects down to super-Brownian motion,

$$X_t(\cdot) = H_t(\{y \in \mathcal{C}(\mathbb{R}^d) : y_t \in \cdot\}) \quad \forall t \geq 0 \ \ \mathbb{N}_H - \text{a.e.}.$$

1.2.2. Lattice trees in high dimensions. Our main result is that the functional limit theorem for historical processes in (b) above, continues to hold for lattice trees in high dimensions (the analogue of (a) was already noted in (1.9)). Recall the definition of μ_n from (1.7).

Theorem 1.4. *For each $d > 8$ there exists $L_0 \geq 1$ such that: for every $L \geq L_0$, there exists $\sigma_0^2 = \sigma_0^2(L, d) > 0$ such that $\mu_n^{\mathrm{LT}} \xrightarrow{w} \mathbb{N}_H^{1,\sigma_0^2}$.*

Here, and throughout this work, the constant σ_0^2 is equal to $v\sigma^2/d$ in [17, Theorem 3.7].

1.3. Discussion. We finish this section with a brief discussion of extensions and applications of our results, and commentary on possible extensions to other models.

Our results are extended in [21] and used in [2] to prove weak convergence of rescaled random walk on lattice trees to a Brownian motion on a Super-Brownian motion cluster, the latter as defined in [5]. [2] reduces this latter result to the verification of two conditions. Roughly speaking, the first of these conditions is that if one chooses K points at random in the lattice tree, then the spatial tree generated by these K points, and suitably rescaled, converges (as the scaling parameter becomes large) to the random tree in $\mathbb{R}^d$ generated by choosing K paths independently at random according to $\int_0^\infty H_t(\cdot)\,dt$ (normalized by its total mass). One interprets this convergence in an appropriate metric space. The weak convergence in Theorem 1.4 is extended in [21] to joint convergence with K independently chosen paths as above, and moreover one can include the branch times and path lengths, to eventually obtain the required spatial tree convergence. The second condition states that in a certain precise sense the vertices of the rescaled tree generated by the K points become dense in the full rescaled lattice trees, uniformly in the scaling parameter, as K becomes large. This is also verified in [21] by using one of the inputs of our tightness argument, namely the modulus of continuity from [20] as stated in Condition 3.4 below.

One may ask about historical convergence in other contexts. This is most natural in cases where there are existing notions of time and ancestry in the model. Such notions exist in the voter model, where the parent of (t, x) is the corresponding point (t', x') from which (t, x) most recently updated its vote, and also in the contact process where the parent of an infected particle is the infected particle which most recently infected it. In his PhD thesis, Tim Banova is using the methodology of Sect. 2 to prove historical convergence of the voter model in dimensions $d > 2$ (for both nearest-neighbour and spread-out (finite range) models). We believe the methodology of Sect. 2 is also relevant for historical convergence of sufficiently spread-out contact processes for $d > 4$. Results for convergence of empirical measures associated with high-dimensional contact processes (but not in the historical context) have relied on a time-discretisation argument and analysis of oriented percolation (OP) (see [13]).

In the context of OP, there is a natural notion of time, but ancestral paths are not unique because there can be multiple connections between vertices. One possible "remedy" is for each site (n, x) of generation n in the cluster of the origin to choose a parent uniformly at random from among sites of generation $n - 1$ in the cluster that are connected to (n, x). We expect that the resulting historical process of sufficiently spread-out OP does converge to historical Brownian motion in dimensions $d > 4$, but note that this process does not encode every connection in the cluster of the origin.

Another approach that one could take (which would also be relevant for percolation and lattice animals) is to define ancestral paths only in terms of pivotal bonds for connections. Pivotal bonds for a connection from $(0, o)$ to (n, x) in oriented percolation,

and from o to x in percolation and lattice animals (if such a connection exists) have a natural temporal ordering, as all paths from point to point must pass through these pivotal bonds in the same order. One could then define historical paths by e.g. linearly interpolating between these pivotal bonds. After appropriate scaling we expect that these historical processes would converge to historical Brownian motion in dimensions larger than the respective critical dimension. Section 2 below would be relevant in each of these contexts.

As has already been noted, except for the voter model [3,4], tightness for any of these models has been a challenging problem even in the context of convergence of empirical measures to SBM, where it has only been established for high-dimensional lattice trees [11] with considerable effort. The proof of tightness for our historical lattice trees uses some bounds on the total mass of the rescaled LT's from [11], and Conditions 2.3 and 2.4 which have also been shown in [20] for OP and the contact process. The additional special property of LT's we use is a sub-Markov property, Lemma 3.15. It would be interesting to see if the proof of tightness can be carried out without this property. The reason is that then control of the total mass process should suffice to prove tightness, even in the historical context, for both the contact process and OP. For percolation and lattice animals, tightness through this historical approach, without even a uniform modulus continuity (Condition 2.3), still seems to be out of reach.

Finally, note that in this paper we have assumed that the step kernel $D(\cdot)$ is uniform on a large box. As noted earlier, the uniformity assumption is not essential. We suspect that D with unbounded support but > 2 finite moments and with $d > d_c = 8$ suffices for convergence to historical Brownian motion. In particular this ought to be true in the nearest-neighbour setting, but at present it would seem to be a considerable challenge (see e.g. [9]) to quantify some dimension d_0 above which this holds.

2. Finite-Dimensional Distributions

2.1. A general theorem. In what follows we write $\mathbb{N}_H$ for $\mathbb{N}_H^{\gamma,\sigma^2}$ where the branching variance $\gamma > 0$ and the diffusion parameter $\sigma^2 > 0$ are fixed throughout.

A collection of $\mathcal{G}$ of bounded continuous functions from $\mathfrak{P}$ to $\mathbb{C}$ is a *determining class* for $\mathcal{M}_F(\mathfrak{P})$ if whenever $\mu, \mu' \in \mathcal{M}_F(\mathfrak{P})$ satisfy $\int g \mathrm{d}\mu = \int g \mathrm{d}\mu'$ for all $g \in \mathcal{G}$, then $\mu = \mu'$. The following is the path-valued analogue of [19, Theorem 2.6]:

Theorem 2.1 (F.d.d. convergence to historical BM). *Let $\mu_n \in \mathcal{M}^{\mathrm{EX}}(\mathfrak{P})$, where $\mathfrak{P} = \mathcal{D}(\mathbb{R}^d)$, and let $\mathcal{G}$ be a determining class for $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$ that contains 1 and is closed under complex conjugation. Assume*

(i) *for every $n \in \mathbb{N}$, $\mu_n\big(\sup_{t\geq 0} H_t(\{h : h_0 \neq o\}) \neq 0\big) = 0$ (paths start at o)*
(ii) *for every $t > 0$, $\mu_n(S > t) \to \mathbb{N}_H(S > t)$ (convergence of survival measures)*
(iii) *for every $t > 0$, $E_{\mu_n}[H_t(\bullet)] \xrightarrow{w} E_{\mathbb{N}_H}[H_t(\bullet)]$ (weak convergence of finite mean measures on $\mathcal{D}$), and for every $\varepsilon > 0$, $\mu_n(H_0(1) > \varepsilon) \to 0$.*
(iv) *for every $\ell \in \mathbb{Z}_+$ and every $\vec{t} \in (0,\infty)^\ell$, and every $\phi_1, \ldots, \phi_\ell \in \mathcal{G}$,*

$$\lim_{n\to\infty} E_{\mu_n}\left[\prod_{j=1}^{\ell} H_{t_j}(\phi_j)\right] = E_{\mathbb{N}_H}\left[\prod_{j=1}^{\ell} H_{t_j}(\phi_j)\right] < \infty. \tag{2.1}$$

Then $\mu_n \xrightarrow{f.d.d.} \mathbb{N}_H$.

Note that $\mathbb{N}_H\big(\sup_{t\geq 0} H_t(\{h : h_0 \neq o\}) \neq 0\big) = 0$. The following elementary consequence of [24, (II.8.6)(a)] states that both the mean measure at time t under $\mathbb{N}_H$, and the mean measure to a uniformly chosen point at time t conditional on survival until time t, are Wiener measure (i.e. the law of Brownian motion) for paths on $[0, t]$:

Lemma 2.2. *The historical canonical measure $\mathbb{N}_H = \mathbb{N}_H^{1,\sigma^2}$ satisfies*

$$E_{\mathbb{N}_H}[H_t(\bullet)] = P(B_{[0,t]} \in \bullet) = E_{\mathbb{N}_H^t}\left[\frac{H_t(\bullet)}{H_t(1)}\right], \quad \forall t > 0,$$

where under P, $B_{[0,t]} = (B_s)_{s\in[0,t]}$ is a d-dimensional BM on $[0, \infty)$ (with $B_1 \sim \mathcal{N}(0, diag(\sigma^2))$) stopped at time $t > 0$.

The proof of Theorem 2.1 is a simple adaption of the proof of [19, Theorem 2.6].

Sketch proof of Theorem 2.1. The only substantial change to the proof of [19, Theorem 2.6] is in the proof of tightness [19, Lemma 3.3].

If $t, \eta > 0$, by (iii) there exists a compact set $K = K_{t,\eta} \subset \mathcal{D}$ such that

$$\sup_n E_{\mu_n}\left[H_t(K^c)\right] < \eta^2,$$

and so by Markov's inequality

$$\sup_n \mu_n(H_t(K^c) > \eta) < \eta. \tag{2.2}$$

Fix $s > 0$. Since $\mu_n(H_s(1) > 0) \to 2/s$ we may find $n_s \in \mathbb{N}$ and $c_s > 0$ so that $\inf_{n\geq n_s} \mu_n(H_s(1) > 0) > c_s$. If $\varepsilon > 0$ we may now use (2.2) and argue as in the proof of [19, Lemma 3.3] to find a compact set $\vec{\mathbf{K}} = \vec{\mathbf{K}}_{t,\varepsilon} \in \mathcal{M}_F(\mathcal{D})$ such that

$$\sup_n \mu_n(H_t \in \vec{\mathbf{K}}^c) < \varepsilon c_s,$$

and hence (working now with the conditional measures) for $t > 0$,

$$\sup_{n\geq n_s} \mu_n^s(H_t \in \vec{\mathbf{K}}^c) < \varepsilon.$$

It follows that for any $\vec{t} \in (0, \infty)^\ell$, $\left(\mu^s_{n,\vec{t}}\right)_{n\in\mathbb{N}}$ is tight in $\mathcal{M}_F(\mathcal{D})^\ell$. Assume $\mu \in \mathcal{M}_F\left(M_F(\mathcal{D})^\ell\right)$ is a limit point of $(\mu^s_{n,\vec{t}})_{n\in\mathbb{N}}$. Then it follows from (2.1) and Dominated Convergence that

$$E_\mu\left[\prod_{i=1}^\ell H_{t_i}(\phi_i)\right] = \mathbb{N}^s_{H,\vec{t}}\left[\prod_{i=1}^\ell H_{t_i}(\phi_i)\right] \qquad \forall \phi_1, \ldots, \phi_\ell \in \mathcal{G}.$$

By [8, Proposition 3.4.6] it follows that $\mu = \mathbb{N}^s_{H,\vec{t}}$. Although this result is stated in [8] for $\mathcal{G}$ a set of real-valued functions, the fact that $\mathcal{G}$ is closed under complex conjugation allows one to see it is also a determining class for complex-valued measures and the proof in [8] then adapts easily to the complex-valued set of functions $\mathcal{G}$. It follows that $\mu^s_{n,\vec{t}} \xrightarrow{w} \mathbb{N}^s_{H,\vec{t}}$ for all $\vec{t} \in (0, \infty)^\ell$. Hypothesis (iii) implies that under μ^s_n, H_0 converges to the zero measure, which is also equal to H_0 under $\mathbb{N}^s_H$. Thus, $\mu^s_{n,\vec{t}} \xrightarrow{w} \mathbb{N}^s_{H,\vec{t}}$ for all $\vec{t} \in [0, \infty)^\ell$, as required. □

For $\vec{s} = (s_0, \ldots, s_m)$, where $0 = s_0 < \cdots < s_m$ and $\vec{k} = (k_0, k_1, \ldots, k_m) \in \mathbb{R}^{d(m+1)}$ define $\phi_{\vec{s},\vec{k}} : \mathcal{D} \to \mathbb{C}$ by

$$\phi_{\vec{s},\vec{k}}(w) = \mathrm{e}^{\mathrm{i}k_0 w_{s_0}} \prod_{j=1}^{m} \mathrm{e}^{\mathrm{i}k_j (w_{s_j} - w_{s_{j-1}})}, \tag{2.3}$$

and let $\mathcal{G}^* = \left\{\phi_{\vec{s},\vec{k}} : \vec{s}, \vec{k} \text{ as above for some } m \in \mathbb{N}\right\}$. Note that $\mathcal{G}^*$ is a determining class for $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$ since finite measures on $\mathcal{D}(\mathbb{R}^d)$ are determined by their finite-dimensional distributions, and the laws of these finite-dimensional random vectors are determined by the characteristic functions of appropriate dimension. The elements of $\mathcal{G}^*$ are precisely those which correspond to the characteristic function of the increments of the path at all finite sets of times. Setting $\vec{k} = \vec{0}$ we see that $1 \in \mathcal{G}^*$ and by replacing k_j with $-k_j$ we observe that $\mathcal{G}^*$ is closed under complex conjugation. So we see that $\mathcal{G}^*$ satisfies the conditions on $\mathcal{G}$ in Theorem 2.1.

Remark 2.3. Under $\mathbb{N}_H$, H_t assigns mass only to paths that are constant from time t onwards and start at o at time 0. The same holds for $H_t^{(n)}$ for all n for LT and BBM. Therefore, when applying Theorem 2.1 in these settings, with $\mathcal{G} = \mathcal{G}^*$ as above, we may restrict our attention to $\phi_{\vec{s}^{(1)},\vec{k}^{(1)}}, \ldots, \phi_{\vec{s}^{(\ell)},\vec{k}^{(\ell)}} \in \mathcal{G}^*$ in part (iv) of the theorem satisfying $s_j^{(i)} \le t_i$ for each i, j and $k_0^{(i)} = 0$. The latter means we can set $\vec{k}^{(i)} = (k_1^{(i)}, \ldots, k_m^{(i)}) \in \mathbb{R}^{dm}$ and ignore the first factor in (2.3). Moreover we can without loss of generality assume that that the largest element of $\vec{s}^{(i)}$ is t_i for each i (i.e., if not we can append an extra component $t^{(i)}$ to $\vec{s}^{(i)}$ and set the corresponding $k_j^{(i)}$ equal to zero without changing $\phi_{\vec{s}^{(i)},\vec{k}^{(i)}}$). ★

In the context of Theorem 1.4, we will use Theorem 2.1 with the determining class $\mathcal{G}^*$ at the end of this section to first establish the following result:

Proposition 2.4. *For $d > 8$ there is an $L_0 \ge 1$ so that for $L \ge L_0(d)$ there is a $\sigma_0^2(L, d) > 0$ for which*

$$\mu_n^{\mathrm{LT}} \xrightarrow{f.d.d.} \mathbb{N}_H.$$

Indeed, condition (i) of Theorem 2.1 trivially holds for lattice trees rooted at the origin. Condition (ii) of the Theorem is (1.6). The first part of Condition (iii) holds by [18, Theorem 2.1], and the second part is obvious because under μ_n^{LT}, $H_0(1) = \frac{1}{C_0 n}$. Condition (iv) of the Theorem (for the determining class $\mathcal{G}^*$) will follow immediately from Proposition 2.6 and Theorem 2.7 below. In order to state these results we need to introduce various notation, which we proceed to do now.

The *degree* of a vertex in a graph is the number of incident edges. Vertices of degree 1 are called *leaves*. Vertices of degree ≥ 3 are called *branch points*.

Definition 2.5. A *non-degenerate shape* is an isomorphism class of finite connected rooted tree graphs whose vertices all have degree 1 or 3, and whose $r + 1$ leaves (for some $r \ge 1$) are labelled $0, 1, 2, \ldots, r$: the root 0 is always one of the leaves. To be more precise, two such graphs are considered to be the same shape if there is a graph isomorphism which preserves the labelling of the leaves (thus there is exactly one shape with 3 leaves and exactly 3 shapes with 4 leaves).

We let Σ_r denote the set of non-degenerate shapes with $r+1$ leaves. For any $\digamma \in \Sigma_r$, we know that α has $r - 1$ branch points, $2r$ vertices and $2r - 1$ edges. Label the branch

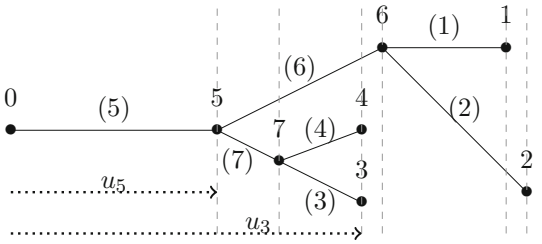

Fig. 5. A depiction of a shape $F \in \Sigma_4$ with vertex labels above vertices and edge labels in brackets. The set of edges in the path from vertex 0 to vertex 1 is $\mathcal{E}_1(F) = \{1, 5, 6\}$. Variables u_i are associated to each of the vertices i, describing a 'length' from 0 to i, to form $\mathcal{T}(F, \vec{u})$. Differences in these u_i are then the "edge lengths"

points as $r + 1, \ldots, 2r - 1$ in order, as you encounter them as you move from the root to vertex 1, then continue to label new internal vertices in the order that you encounter them as you move from the root to vertex 2 and so on up to vertex r. See e.g. Fig. 5. This is just a convenient arbitrary but fixed order. For $i, j \in \{0, \ldots, 2r - 1\}$, we abuse the notation for the usual order and let $i \wedge j \in \{0, \ldots, 2r - 1\}$ denote the greatest common antecedent (gca) of i and j. The edges e of $F \in \Sigma_r$ are labelled as $\mathcal{E}(F) = \{1, \ldots, 2r-1\}$ corresponding to the vertex labelling of the endvertex of e that is farthest from the root. For $e, f \in \mathcal{E}(F)$, write $e \prec f$ if e is an ancestor of f in F.
For leaves $\ell \in 1, \ldots, r$, let $\mathcal{E}_\ell(F)$ be the set of edges in the unique path in F from o to ℓ.

For $F \in \Sigma_r$ we assign edge lengths by letting $\vec{u} = (u_1, \ldots, u_{2r-1}) \in (0, \infty)^{2r-1}$ give the distances from the vertices to the root. That is, u_i is the distance from the root to vertex i, and the edge lengths can be found by differencing. We let $\mathbb{T}(F, \vec{u})$ denote the resulting tree with shape F and edge lengths $\vec{u}$. See Fig. 5. We often will specify the distances $\vec{t} = (t_1, \ldots, t_r) \in (0, \infty)^r$ of the r leaves to the root in advance. In this case we let $\mathcal{M}(\vec{t}, F)$ denote the set of possible vertex distances from the root. That is, $\mathcal{M}(\vec{t}, F)$ denotes the set of $\vec{u} = (u_1, \ldots, u_{2r-1}) \in (0, \infty)^{2r-1}$ such that:

$$u_i = t_i, \text{ for } i = 1, \ldots, r; \tag{2.4}$$

$$\text{if } k \text{ and } j \text{ are vertices of } F \text{ and } k \text{ is an ancestor of } j \text{ in } F, \text{ then } u_k < u_j. \tag{2.5}$$

◀

Consider a given (non-degenerate) shape $F \in \Sigma_r$, $\vec{t} \in (0, \infty)^r$, and $\vec{u} \in \mathcal{M}(\vec{t}, F)$ as above. Let $\boldsymbol{s} = (\vec{s}^{(1)}, \ldots, \vec{s}^{(r)})$, where $\vec{s}^{(\ell)} = (s_0^{(\ell)}, \ldots, s_{m^{(\ell)}}^{(\ell)})$, and $0 = s_0^{(\ell)} < s_1^{(\ell)} < \cdots < s_{m^{(\ell)}}^{(\ell)} = t_\ell$ for each $\ell \in [r] := \{1, \ldots, r\}$. If $e \notin \mathcal{E}_\ell(F)$ then set $\mathcal{I}(e, \vec{s}^{(\ell)}) = \varnothing$. If $e \in \mathcal{E}_\ell(F)$, then let $\mathcal{I}(e, \vec{s}^{(\ell)})$ denote those elements of $\vec{s}^{(\ell)}$ that fall in the interval $(u_-(e), u_+(e))$, where $u_-(e), u_+(e)$ are the elements of $\vec{u}$ corresponding to the endvertices of e (if e is adjacent to the root, then set $u_-(e) = 0$). Let $\mathcal{I}(e, \boldsymbol{s}) = \cup_{\ell=1}^r \mathcal{I}(e, \vec{s}^{(\ell)})$. The $j(e) := |\mathcal{I}(e, \boldsymbol{s})|$ elements of $\mathcal{I}(e, \boldsymbol{s})$ divide the interval $[u_-(e), u_+(e)]$ into $j_e := j(e) + 1$ subintervals - denote their lengths by $(\check{s}_{e,k})_{k=1,\ldots,j_e}$ and set $\check{\boldsymbol{s}} = (\check{s}_{e,k})_{e \in \mathcal{E}(F); k=1,\ldots,j_e}$. If $j(e) = 0$ then $\check{s}_{e,1} = u_+(e) - u_-(e)$. Note that $j(e)$ and $\check{\boldsymbol{s}}$ depend on $F, \vec{u}, \boldsymbol{s}$. See Fig. 6.

For $\ell \in [r]$, $e \in \mathcal{E}_\ell(F)$, and $a \in \{1, \ldots, j_e\}$, let

$$\zeta^{[e]}(a, \ell) = \min \left\{ i \le m^{(\ell)} : s_i^{(\ell)} \ge u_-(e) + \sum_{i_e=1}^{a} \check{s}_{e,i_e} \right\}.$$

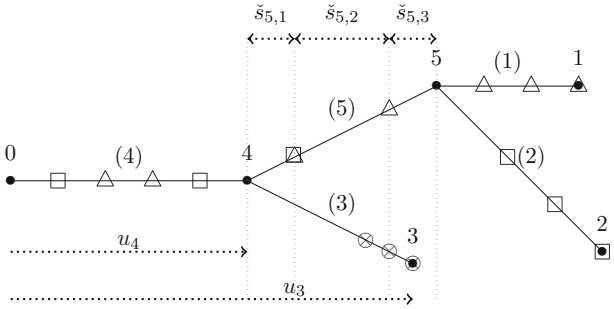

Fig. 6. The tree $\mathbb{T}(F, \vec{u})$ together with times $\boldsymbol{s}$. The ($m^{(1)} = 7$) $\triangle$ symbols represent times $s_1^{(1)}, \ldots, s_7^{(1)}$. Similarly $\square$ symbols represent times $s_j^{(2)}$ (with $m^{(2)} = 6$) and $\otimes$ symbols represent times $s_j^{(3)}$ (with $m^{(3)} = 3$) respectively. In this example there is one point (on edge 5) that is both square and triangle simultaneously. The 'subinterval' lengths $\check{s}_{5,i}$ are indicated for edge 5

Given $\vec{k}^{(\ell)} = (k_1^{(\ell)}, \ldots, k_{m^{(\ell)}}^{(\ell)}) \in (\mathbb{R}^d)^{m^{(\ell)}}$, for each $\ell \in [r]$, and for $e \in \{1, \ldots, 2r-1\}$ and $a \le j_e$, let

$$\check{k}_{e,a} = \sum_{\ell : e \in \mathcal{E}_\ell(F)} k^{(\ell)}_{\zeta^{[e]}(a,\ell)}.$$

For given $\sigma^2 > 0$, $r \in \mathbb{N}$, $F \in \Sigma_r$, $\vec{u} \in \mathcal{M}(\vec{t}, F)$ for some $\vec{t} \in (0,\infty)^r$, and for given $\boldsymbol{k} = (\vec{k}^{(1)}, \ldots, \vec{k}^{(r)})$ and $\boldsymbol{s} = (s_0^{(\ell)}, \ldots, s_{m^{(\ell)}}^{(\ell)})$ (where, for $\ell \in [r]$, $m^{(\ell)} \in \mathbb{N}$, $\vec{s}^{(\ell)} = (0 = s_0^{(\ell)}, s_1^{(\ell)}, \ldots, s_{m^{(\ell)}}^{(\ell)} = t_\ell)$, $(s_i^{(\ell)} < s_{i+1}^{(\ell)})$, $\vec{k}^{(\ell)} \in \mathbb{R}^{m(\ell)}$), define

$$\Phi_{\sigma^2}(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) = \prod_{e=1}^{2r-1} \prod_{i=1}^{j_e} \exp\Big(\frac{-\sigma^2 |\check{k}_{e,i}(F, \vec{u}, \boldsymbol{s})|^2 \check{s}_{e,i}(F, \vec{u}, \boldsymbol{s})}{2} \Big). \tag{2.6}$$

The following proposition (proved in Sect. 2.2) gives an explicit formula for the right hand side of (2.1). The integral over $\mathcal{M}(\vec{t}, F)$ is actually an $(r-1)$-dimensional integral over $(u_{r+1}, \ldots, u_{2r-1})$ as the first r components are fixed.

Proposition 2.6. *For any $r \in \mathbb{N}$, $\vec{t} \in (0,\infty)^r$ and $\phi^{(1)}, \ldots, \phi^{(r)} \in \mathcal{G}^*$ (with $\phi^{(\ell)} = \phi_{\vec{s}^{(\ell)}, \vec{k}^{(\ell)}}$ where $\vec{s}^{(\ell)} = (s_0^{(\ell)} = 0, s_1^{(\ell)}, \ldots, s_{m^{(\ell)}}^{(\ell)} = t_\ell)$, $(s_i^{(\ell)} < s_{i+1}^{(\ell)})$ and $\vec{k}^{(\ell)} = (k_1^{(\ell)}, \ldots, k_{m^{(\ell)}}^{(\ell)})$) as in Remark 2.3,*

$$E_{\mathbb{N}_H^{1,\sigma^2}}\Big[\prod_{\ell=1}^{r} H_{t_\ell}(\phi^{(\ell)}) \Big] = \sum_{F \in \Sigma_r} \int_{\vec{u} \in \mathcal{M}(\vec{t}, F)} \Phi_{\sigma^2}(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) d\vec{u}.$$

The following result is proved in Sect. 2.4 below.

Theorem 2.7. *Let $d > 8$. There exists L_0 such that for all $L \ge L_0$, and $r \in \mathbb{N}$, $\vec{t}$, and $\phi^{(1)}, \ldots, \phi^{(r)} \in \mathcal{G}^*$ as in Proposition 2.6,*

$$E_{\mu_n^{\mathrm{LT}}}\Big[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)}) \Big] \to \sum_{F \in \Sigma_r} \int_{\vec{u} \in \mathcal{M}(\vec{t}, F)} \Phi_{\sigma_0^2}(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) d\vec{u} \quad \text{as } n \to \infty. \tag{2.7}$$

Proof of Proposition 2.4. As noted after the statement of the Proposition, we only need verify condition (2.1) in Theorem 2.1 with $\mathcal{G} = \mathcal{G}^*$, and this is immediate from Proposition 2.6 and Theorem 2.7. □

2.2. Branching Brownian motion f.d.d. and proof of Proposition 2.6.

Definition 2.8. Let $r \in \mathbb{N}$, $F \in \Sigma_r$, $\vec{t} \in (0,\infty)^r$, and $\vec{u} \in \mathcal{M}(\vec{t}, F)$. For each edge e we let $\ell(e) \in \{1,\dots,r\}$ be the minimal leaf such that $e \in \mathcal{E}_\ell(F)$. Let $(W^i_s)_{s\le t_i}$ for $i \in [r]$ be (dependent) d-dimensional Brownian motions with variance parameter σ^2, such that for any distinct $i, j \in \{1,\dots,r\}$,

$$W^i_s = W^j_s \quad \text{for all } s \le u_{i\wedge j}, \tag{2.8}$$

(recall $u_{i\wedge j}$ is the distance from the root to the gca of i and j) and

$$\begin{aligned}&\big\{(W^{\ell(e)}_{u_-(e)+s} - W^{\ell(e)}_{u_-(e)})_{s\le u_+(e)-u_-(e)} : e \text{ an edge of } \mathbb{T}(F,\vec{u})\big\}\\ &\quad\text{are independent } d\text{-dimensional Brownian motions with variance } \sigma^2.\end{aligned} \tag{2.9}$$

We call $(W^1,\dots,W^r)$ a *tree-indexed BM* with variance parameter σ^2 on $\mathbb{T}(F,\vec{u})$. ◀

(2.9) simply says that the collection of Brownian motions run along the disjoint edges of $\mathbb{T}(F,\vec{u})$ are independent. Note that in (2.9) we could choose any ℓ such that $e \in \mathcal{E}_\ell(F)$ by (2.8). We remark that the law of $(W^1,\dots,W^r)$ is uniquely specified by the above (note it is mean zero Gaussian with $\mathrm{Cov}(W^i(s_i), W^j(s_j)) = \sigma^2 \min(u_{i\wedge j}, s_i, s_j)$).

Proposition 2.9. *Let $r \in \mathbb{N}$, $F \in \Sigma_r$, $\vec{t} \in (0,\infty)^r$, $\vec{u} \in \mathcal{M}(\vec{t}, F)$, and $(W^i_s)_{s\le t_i}$ for $i \in [r]$ be a tree-indexed BM with variance parameter σ^2 on $\mathbb{T}(F,\vec{u})$. If $\ell \in [r]$, $m^{(\ell)} \in \mathbb{N}$, $\vec{s}^{(\ell)} = (0 = s^{(\ell)}_0, s^{(\ell)}_1, \dots, s^{(\ell)}_{m^{(\ell)}} = t_\ell)$, $(s^{(\ell)}_i < s^{(\ell)}_{i+1})$, $\vec{k}^{(\ell)} \in \mathbb{R}^{m(\ell)}$, and $\phi^{(\ell)} = \phi_{\vec{s}^{(\ell)},\vec{k}^{(\ell)}}$, then*

$$\mathbb{E}\Big[\prod_{\ell=1}^{r} \phi^{(\ell)}(W^\ell)\Big] = \Phi_{\sigma^2}(F,\vec{u},\boldsymbol{s},\boldsymbol{k}).$$

Proof. This is an elementary calculation which divides the dependent Brownian increments on the left-hand side into smaller non-overlapping independent increments and keeps track of the Fourier coefficients multiplying each increment. The details are left for the reader. □

Notation. For $t \ge 0$, let

$$[t]_n = \max\{k/n \in [0,t] : k \in \mathbb{Z}_+\}. \tag{2.10}$$

Proof of Proposition 2.6. We will work with the measures μ^{BBM}_n for branching Brownian motion where the variance parameter is $\sigma^2 > 0$ and the offspring distribution is critical binary branching, i.e., $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$, and so $\gamma = 1$. In this case, [23, Proposition 2.6(a)(i)] with $\phi = 1$, and Doob's strong L^p inequality for martingales imply

$$\forall p > 1 \ \text{ there is a } C_p \text{ such that } \forall K \in \mathbb{N} \ \sup_n E_{\mu_n}\Big[\sup_{t\le K} H^{(n)}_t(1)^p\Big] \le C_p K^{p-1}. \tag{2.11}$$

Theorem 1.3 and the continuity of $t \to H_t$ under $\mathbb{N}_H = \mathbb{N}^{1,\sigma^2}_H$ imply weak convergence of $(H^{(n)}_{t_1}(\phi^{(1)}),\dots,H^{(n)}_{t_r}(\phi^{(r)}))$ under $(\mu^{\mathrm{BBM}}_n)^{t_1}$ to $(H_{t_1}(\phi^{(1)}),\dots,H_{t_r}(\phi^{(r)}))$ under $\mathbb{N}^{t_1}_H$ (see, e.g., [8, Theorem 10.2 in Ch. 3]). Note also that for K large enough,

$\left|\prod_{l=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})\right|^2 \le \sup_{t\le K} H_t^{(n)}(1)^{2r}$. Therefore, the above together with (2.11) and Dominated Convergence imply that

$$\begin{aligned} E_{\mathbb{N}_H}\left[\prod_{\ell=1}^{r} H_{t_\ell}(\phi^{(\ell)})\right] &= E_{\mathbb{N}_H^{t_1}}\left[\prod_{\ell=1}^{r} H_{t_\ell}(\phi^{(\ell)})\right]\mathbb{N}_H(S > t_1) \\ &= \lim_{n\to\infty} E_{(\mu_n^{\mathrm{BBM}})^{t_1}}\left[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})\right]\mu_n^{\mathrm{BBM}}(S^{(n)} > t_1) \\ &= \lim_{n\to\infty} E_{\mu_n^{\mathrm{BBM}}}\left[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})\right]. \end{aligned} \tag{2.12}$$

A moment calculation for branching Brownian motion which uses Proposition 2.9 and is much simpler than that for lattice trees in Theorem 2.7, shows that the limit on the right-hand side of the above equals the right-hand side of the equality in the proposition. We sketch the proof as it explains how the right-hand side of (2.7) arises. Let $\mathbb{Z}_+/n = \{j/n : j \in \mathbb{Z}_+\}$. Recall (1.10), and let $I_t = \{\beta \in I : |\beta| = \lfloor t\rfloor\}$. Fix $t_1, \dots, t_r > 0$ and consider only large enough n so that

$$\lfloor nt_i\rfloor \ge 2, \quad i = 1, \dots, r.$$

Recall the random subset GW of indices in I defined in Sect. 1.2.1. A simple expansion of the sum defining $H_{t_\ell}^{(n)}$ shows that

$$\begin{aligned} E_{\mu_n^{\mathrm{BBM}}}\left[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})\right] &= \frac{1}{n^{r-1}} \sum_{\beta^1\in I_{nt_1}} \cdots \sum_{\beta^r\in I_{nt_r}} \mathbb{E}\left[\mathbb{1}_{\{\{\beta^1,\dots,\beta^r\}\subset GW\}} \prod_{\ell=1}^{r} \phi^{(\ell)}(B_{\cdot\wedge t_\ell}^{\beta^\ell})\right] \\ &= \frac{1}{n^{r-1}} \sum_{\beta^1\in I_{nt_1}} \cdots \sum_{\beta^r\in I_{nt_r}} \mathbb{P}\big(\{\beta^1,\dots,\beta^r\}\subset GW\big)\mathbb{E}\Big[\prod_{\ell=1}^{r} \phi^{(\ell)}(\hat{B}_{\cdot\wedge t_\ell}^{\beta^\ell})\Big], \end{aligned} \tag{2.13}$$

where in the last we used the independence of the branching variables $\{e^\beta : \beta \in I\}$ and the spatial motions $\{\hat{B}^\beta : \beta \in I\}$ as well as the fact that $B_{\cdot\wedge t_\ell}^{\beta^\ell} = \hat{B}_{\cdot\wedge t_\ell}^{\beta^\ell}$ if $\beta^\ell \in GW$. It is easy to see that the contribution to the above sum from $\beta^1, \dots, \beta^r$ such that for some $i \ne j$: $\pi\beta^i$ is an ancestor of β^j, is bounded by $C(r, K)/n$ for $\max\{t_i : i \in [r]\} \le K$. To see this, note that if $\pi\beta^i$ is an ancestor of β^j, then $\pi\beta^i$ is determined by β^j since its length is $\lfloor nt_i\rfloor - 1$. This means there are only two possible values of β^i and so we can bound this contribution by twice the $(r-1)$-fold sum with each $\vec{k}^{(\ell)} = \vec{0}$ (so each $\phi^{(\ell)} = 1$), and applying (2.11), we obtain the above bound. Fix $\vec{\beta} := (\beta^1, \dots, \beta^r) \in I_{nt_1} \times \cdots \times I_{nt_r}$ so that none of the indices has a parent which is an ancestor of another index (in particular all are distinct). Call such a $\vec{\beta}$ a good value of $\vec{\beta}$. Then, in particular, $\vec{\beta}$ uniquely determines a non-degenerate shape $\mathcal{F}(\vec{\beta}) \in \Sigma_r$ where $\beta^1, \dots, \beta^r$ label the r leaves and one can define the internal vertices of the shape by locating the branch points from the root to β^1, then the new branch points while proceeding from the root to β^2, and so on up to β^r. See e.g. Fig. 7. In this way we label the internal vertices by $\beta^{r+1}, \dots, \beta^{2r-1}$ using our labelling convention in Definition 2.5 (now with β^i in place of i). For example (assuming $r > 1$), $\beta^{r+1} = \beta^1|\kappa_{r+1}$, where $\kappa_{r+1} = \max\{\kappa : \beta^1|\kappa = \beta^\ell|\kappa \text{ for all } \ell > 1\} \in \{0, \dots, \min\{|\beta^\ell|\} - 2\}$ (the upper bound

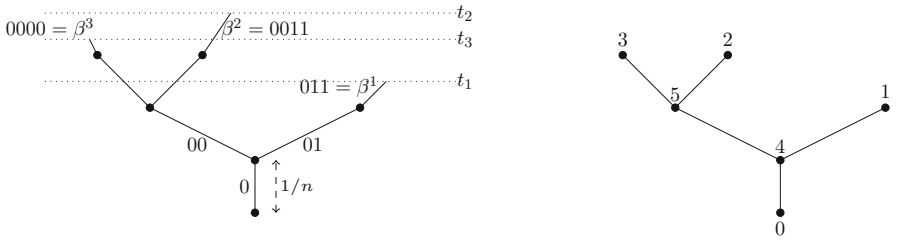

Fig. 7. On the left is (part of) a GW tree with $\beta^1, \beta^2, \beta^3$ indicated. Here $|\beta^1| = |\beta^5| = 2$, $|\beta^2| = |\beta^3| = 3$, and $|\beta^4| = 1$, and this contributes to (2.13) when $t_2, t_3 \in [3/n, 4/n)$ and $t_1 \in [2/n, 3/n)$ as depicted. On the right is the corresponding tree shape. The edge lengths associated to the latter are determined by taking differences of the u_e, where $u_4 = 1/n, u_5 = 2/n, u_1 = t_1, u_2 = t_2, u_3 = t_3$

since $\vec{\beta}$ is good), and then continue down the branch towards β^1 until there is only one leaf (β^1) along the remaining tree. Note that each β^ℓ for $\ell > r$ is of the form $\beta^i|\kappa_\ell$ for some $i = i(\ell) \le r$ and some $\kappa_\ell < |\underline{\beta}^i|$, i.e. is an ancestor of some β^i.

We introduce tree distances $\vec{u}(\vec{\beta}) = (u_1, \dots, u_{2r-1})$ for the above shape, with $u_i \in \mathbb{Z}_+/n \setminus \{0\}$ for $i > r$, by setting

$$u_\ell = \begin{cases} t_\ell & \text{if } \ell \le r, \\ (|\beta^\ell| + 1)/n & \text{if } \ell \in \{r+1, \dots, 2r-1\}. \end{cases}$$

Recall that u_ℓ is the distance from vertex β^ℓ to the root and so edge distances can be found by differencing. Denote this tree shape with edge lengths by $\mathbb{T}(\vec{\beta})$. Note that the fact that $\vec{\beta}$ is good ensures that $u_\ell < |\beta^i|/n \le u_i$, whenever β^ℓ is an ancestor of β^i for $\ell > r$ and $i \le r$. In fact, the possible values of $\vec{u}$ are now given by the discrete analogue of $\mathcal{M}(\vec{t}, F)$,

$$\vec{u} \in \mathcal{M}_n(\vec{t}, F) := \{\vec{u} \in (0,\infty)^{2r-1} : u_i = t_i \text{ for } i \le r, u_i \in \mathbb{Z}_+/n \setminus \{0\} \text{ for } i > r \\ \text{and } u_k < u_j \text{ whenever } \beta^k \text{ is an ancestor of } \beta^j\}. \tag{2.14}$$

In the above notation we use the fact that the ordering of the leaves given by $\vec{t}$, the shape F, and our convention on numbering internal vertices, determines the ancestral relationship between the β^k, not the particular choice of $\vec{\beta}$. The definition of u_ℓ for the internal branch points $\ell > r$ ensures that

$$(W^1, \dots, W^r) := (\hat{B}^{\beta^1}, \dots, \hat{B}^{\beta^r}) \text{ is a tree-indexed Brownian motion} \\ \text{with variance parameter } \sigma^2 \text{ on } \mathbb{T}(\vec{\beta}). \tag{2.15}$$

To see this, note that at a branch point $\beta^\ell = \beta^i \wedge \beta^j$ for leaves i, j and $\ell > r$, the Brownian paths $\hat{B}^{\beta^i}$ and $\hat{B}^{\beta^j}$ do not split apart and evolve independently until time $(|\beta^\ell| + 1)/n = u_\ell$.

We now decompose the sum over good $\vec{\beta}$ in (2.13) according to its shape, F, and edge lengths $\vec{u}$. Abbreviating $(\beta_1, \dots, \beta_r) \in I_{nt_1} \times \cdots \times I_{nt_r}$ as $\vec{\beta} \in I_{n\vec{t}}$, and writing $\vec{\beta} \subset GW$ for $\{\beta^1, \dots, \beta^r\} \subset GW$, the right hand side of (2.13) becomes

$$\sum_{F\in\Sigma_r}\frac{1}{n^{r-1}}\sum_{\vec{u}\in\mathcal{M}_n(\vec{t},F)}\sum_{\substack{\vec{\beta}\in I_{n\vec{t}}:\\ \vec{\beta}\text{ good}}}\mathbb{1}_{\{F(\vec{\beta})=F\}}\mathbb{1}_{\{\vec{u}(\vec{\beta})=\vec{u}\}}\mathbb{P}(\vec{\beta}\subset GW)\mathbb{E}\Big[\prod_{\ell=1}^{r}\phi^{(\ell)}(\hat{B}^{\beta^\ell}_{\cdot\wedge t_\ell})\Big]+\mathcal{O}\Big(\frac{1}{n}\Big). \tag{2.16}$$

Recall the notation (2.10). Choose $F\in\Sigma_r$, $\vec{u}\in\mathcal{M}_n(\vec{t},F)$, and $\vec{\beta}\in I_{n\vec{t}}$ such that $F(\vec{\beta})=F$ and $\vec{u}(\vec{\beta})=\vec{u}$. Let $N=N(F,\vec{u})\in\mathbb{Z}_+$ be the number of ancestors of $\beta^1,\dots,\beta^r$ in the index set I. Note that N is equal to n times the sum of (truncated) edge lengths in $\mathbb{T}(\vec{\beta})$ determined by $\vec{u}'$ where $u'_\ell=u_\ell$ if $\ell>r$ and $u'_\ell=[u_\ell]_n=[t_\ell]_n$ if $\ell\le r$ (see e.g. the left hand side of Fig. 7). (Here we identify each edge of rescaled length $1/n$ with the index of its entry vertex in I.) Therefore N is a function of $(F,\vec{u})$ as the notation suggests. It follows immediately that $\mathbb{P}(\vec{\beta}\subset GW)=2^{-N}$ since $\vec{\beta}\subset GW$ if and only if each of these ancestors has two offspring.

It follows from this, (2.15), and Proposition 2.9, that (2.16) equals

$$\begin{aligned}&\sum_{F\in\Sigma_r}\frac{1}{n^{r-1}}\sum_{\vec{u}\in\mathcal{M}_n(\vec{t},F)}\sum_{\vec{\beta}\in I_{n\vec{t}},\,\vec{\beta}\text{ good}}\mathbb{1}_{\{F(\vec{\beta})=F\}}\mathbb{1}_{\{\vec{u}(\vec{\beta})=\vec{u}\}}\mathbb{P}(\vec{\beta}\subset GW)\Phi(F,\vec{u},\boldsymbol{s},\boldsymbol{k})+\mathcal{O}\Big(\frac{1}{n}\Big)\\ &=\sum_{F\in\Sigma_r}\frac{1}{n^{r-1}}\sum_{\vec{u}\in\mathcal{M}_n(\vec{t},F)}2^{-N}\Phi(F,\vec{u},\boldsymbol{s},\boldsymbol{k})\sum_{\vec{\beta}\in I_{n\vec{t}}}\mathbb{1}_{\{F(\vec{\beta})=F\}}\mathbb{1}_{\{\vec{u}(\vec{\beta})=\vec{u}\}}+\mathcal{O}\Big(\frac{1}{n}\Big).\end{aligned}$$

Here dropping the "good" requirement on $\vec{\beta}$, at the cost of $\mathcal{O}\Big(\frac{1}{n}\Big)$, is again an easy calculation along the lines of that done earlier.

For fixed $F\in\Sigma_r$ and $\vec{u}\in\mathcal{M}_n(\vec{t},F)$, the number of choices for $\vec{\beta}\subset I$ with this shape and edge lengths in the above is 2^N. This is because there are two choices for the offspring labels for each of the N "ancestors" above. Therefore combining the above equalities leads to

$$E_{\mu_n^{\mathrm{BBM}}}\Big[\prod_{\ell=1}^{r}H^{(n)}_{t_\ell}(\phi^{(\ell)})\Big]=\sum_{F\in\Sigma_r}\frac{1}{n^{r-1}}\sum_{\vec{u}\in\mathcal{M}_n(\vec{t},F)}\Phi(F,\vec{u},\boldsymbol{s},\boldsymbol{k})+\mathcal{O}\Big(\frac{1}{n}\Big).$$

As $n\to\infty$ in the above, the $(r-1)$-fold Riemann sum converges to the $(r-1)$-dimensional integral in the right-hand side of the proposition, and so the result now follows from (2.12). For the Riemann sum convergence, we note that the $\vec{u}$ dependence of the integrand admits finitely many jump discontinuities. □

2.3. Lattice tree f.d.d. We now turn to the LT setting. Fix $m\in\mathbb{N}$, $t>0$, $\vec{k}=(k_1,\dots,k_m)\in\mathbb{R}^{dm}$, and $\vec{s}=(0=s_0,\dots,s_m=t)$, where $s_i<s_{i+1}$. Then

$$H^{(n)}_t(\phi_{\vec{s},\vec{k}})=\int_{\mathcal{D}}\phi\,dH^{(n)}_t=\frac{1}{C_0n}\sum_{\sqrt{n}x\in\mathcal{T}_{nt}}\prod_{j=1}^{m}\mathrm{e}^{\mathrm{i}k_j\left(w^{(n)}_{s_j}(t,x)-w^{(n)}_{s_{j-1}}(t,x)\right)}. \tag{2.17}$$

Letting $\vec{x}_m=(x_1,\dots,x_m)$ and setting $x_0=0\in\mathbb{Z}^d$ we have

$$E_{\mu_n^{\mathrm{LT}}}[H^{(n)}_t(\phi_{\vec{s},\vec{k}})]=\frac{C_1}{C_0}\sum_{\vec{x}_m\in(\mathbb{Z}^d)^m}\prod_{j=1}^{m}\mathrm{e}^{\mathrm{i}\frac{k_j}{\sqrt{n}}(x_j-x_{j-1})}\mathbb{P}\Big(x_m\in\mathcal{T}_{nt},\cap_{j=1}^{m}\{w_{ns_j}(nt,x_m)=x_j\}\Big). \tag{2.18}$$

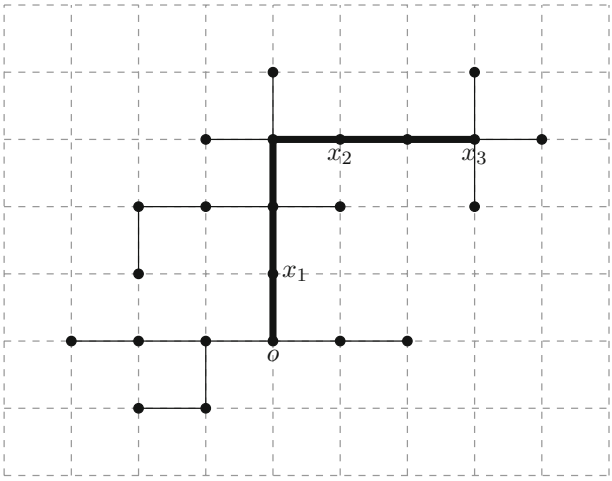

Fig. 8. A depiction of the event in the detailed 1-particle function with $n = 1$, $t = 6$, $s_1 = 1$ and $s_2 = 4$, with the path $s \mapsto w_s(6, x_3)$ in bold (recall the notation from (1.1))

We call the quantity $\mathbb{P}\big(x_m \in \mathcal{T}_{nt}, \cap_{j=1}^m \{w_{ns_j}(nt, x_m) = x_j\}\big)$ a *detailed 1-particle function*, (see e.g. Fig. 8), and the Fourier transform of the increments is called a *detailed* 1-particle transform, i.e.

$$\sum_{\vec{x}_m \in (\mathbb{Z}^d)^m} \prod_{j=1}^m \mathrm{e}^{\mathrm{i}k_j(x_j - x_{j-1})} \mathbb{P}\Big(x_m \in \mathcal{T}_{nt}, \cap_{j=1}^m \{w_{ns_j}(nt, x_m) = x_j\}\Big).$$

Related quantities arising from expectations of the form

$$E_{\mu_n^{\mathrm{LT}}}\left[\prod_{\ell=1}^r H_{t_\ell}^{(n)}(\phi^{(\ell)})\right],$$

with $\phi^{(\ell)} = \phi_{\vec{s}^{(\ell)}, \vec{k}^{(\ell)}}$ ($\vec{s}^{(\ell)}, \vec{k}^{(\ell)}$ as in Theorem 2.7) are called *detailed r-particle transforms*. Therefore Theorem 2.7 amounts to verifying the appropriate asymptotics for the detailed r-particle transforms.

When $m = 1$, the detailed 1-particle function is simply $\mathbb{P}(x_1 \in \mathcal{T}_{nt})$, and its Fourier transform becomes $\sum_{x \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i}k_1 x} \mathbb{P}(x \in \mathcal{T}_{nt})$. These quantities are called the 1-particle functions (traditionally in the literature these have been called the 2-point functions, with the two points being the origin o and x_1). For $\vec{n} \in \mathbb{Z}_+^r$ and $\vec{x} = (x_1, \ldots, x_r) \in \mathbb{Z}^{dr}$ we can define the r-particle functions (see e.g. Fig. 9):

$$\rho_{\vec{n}}(\vec{x}) = \mathbb{P}(\cap_{i=1}^r \{x_i \in \mathcal{T}_{n_i}\}),$$

and (their Fourier transforms) the r-particle transforms for $\vec{k} \in (\mathbb{R}^d)^r$:

$$\hat{\rho}_{\vec{n}}(\vec{k}) = \sum_{\vec{x} \in (\mathbb{Z}^d)^r} \mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{x}} \rho_{\vec{n}}(\vec{x}).$$

We write $\mathcal{O}(x)$ to denote a quantity whose absolute value is bounded by a constant times x. Using the inductive method of [12,14] the following was shown in [17, Theorem 3.7]:

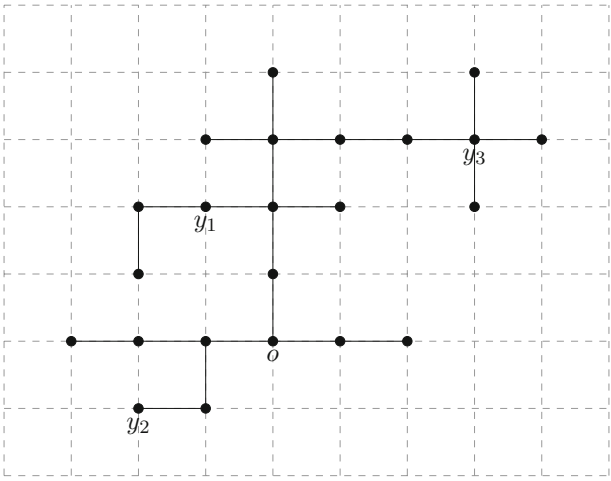

Fig. 9. A depiction of the event in the 3-particle function $\rho_{(3,3,6)}(y_1, y_2, y_3)$

Theorem 2.10 [17]. *Fix $d > 8$. There exists $L_0 = L_0(d) \gg 1$ such that for every $L \geq L_0$:*

There exist $K, C_A > 0$ such that, for every $\delta \in (0, 1 \wedge \frac{d-8}{2})$,

$$\sup_{n\in\mathbb{Z}_+} \sup_{k\in\mathbb{R}^d} |\hat{\rho}_n(k)| = \sup_{n\in\mathbb{Z}_+} \hat{\rho}_n(0) \leq K, \tag{2.19}$$

and

$$\hat{\rho}_n\left(\frac{k}{\sqrt{n}}\right) = C_A e^{-\frac{\sigma_0^2|k|^2}{2}} \left[1 + \mathcal{O}\left(\frac{|k|^2}{n^\delta}\right) + \mathcal{O}\left(n^{-\frac{d-8}{2}}\right)\right]. \tag{2.20}$$

Recall that the constant C_A is equal to A' in the paper [17], while σ_0^2 is equal to $v\sigma^2/d$ in [17]. The error terms (see [17, Theorem 3.7, Lemma 3.8]) in (2.20) depend on d, L but are uniform in $\{k \in \mathbb{R}^d : |k|^2 \leq C \log n\}$ (where C depends on δ). Taking $k = 0$ above we see that, as claimed in Sect. 1.1, $C_A = \lim_{n\to\infty} \mathbb{E}[|\mathcal{T}_n|]$. Asymptotics for the r-particle transforms are provided in [17, Theorem 1.14]. In particular there exists $C_V > 0$ depending on D, d such that

$$n^{-1}\mathbb{E}[|\mathcal{T}_n|^2] = n^{-1}\hat{\rho}_{(n,n)}(0,0) \to C_V C_A^3. \tag{2.21}$$

Recall that the constant C_V in our paper is equal to $V\rho^2$ in [17]. Our task is to "upgrade" these kinds of results from [17] to get asymptotics for the "detailed" r-particle transforms. This is the focus of the next section.

2.4. The LT detailed r-particle transforms and proof of Theorem 2.7. Recall the labelling convention for internal vertices (branch points) and edges in $\digamma$ from Definition 2.5.

A lattice tree $T \ni o$ having $r + 1$ leaves ($o = x_0$ and $x_1, \ldots, x_r$), $r - 1$ vertices $x_{r+1}, \ldots, x_{2r-1}$ of degree 3, and all other vertices degree 2, has an associated abstract tree Γ as follows: $x_i \mapsto i$, and any two vertices i, i' in Γ are connected via a single edge if the shortest path from x_i to $x_{i'}$ in T passes through no other x_j. All vertices in Γ are degree 1 or 3. Relabelling the vertices of degree 3 according to the labelling convention

in Definition 2.5 gives an abstract shape Γ', which is the *shape of* T and the points $x_1, \dots, x_r$ (and o), and we write $v_g \in \{x_{r+1}, \dots, x_{2r-1}\}$ for the vertex in T that mapped to branch point $g \in \Gamma'$.

Given F, $\check{\boldsymbol{y}} = (\check{y}_{e,i})_{i\in[j_e],e\in[2r-1]}$, and $\check{\boldsymbol{n}} = (\check{n}_{e,i})_{i\in[j_e],e\in[2r-1]}$ with each $\check{y}_{e,i} \in \mathbb{Z}^d$ and each $\check{n}_{e,i} \in \mathbb{N}$, let $\check{\boldsymbol{T}}(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}})$ denote the set of lattice trees $T \ni o$ such that:

(*) for each $\ell \in [r]$ the tree T contains $x_\ell = \sum_{e\in\mathcal{E}_\ell(F)} \sum_{i=1}^{j_e} \check{y}_{e,i}$, and the shape of the minimal subtree T' of T containing o and $x_1, \dots, x_r$ is F, and for each branch point $g \in F$, the corresponding vertex v_g is tree distance $\sum_{f\prec g} \sum_{i=1}^{j_f} \check{n}_{f,i} + \sum_{i=1}^{j_g} \check{n}_{g,i}$ from the root in T', and

(**) for each $\ell \in [r]$, each $e \in \mathcal{E}_\ell(F)$, and each $i_e \in \{1, \dots, j_e\}$, the path from o to x_ℓ in T passes through the point $\sum_{f\prec e} \sum_{i=1}^{j_f} \check{y}_{f,i} + \sum_{i=1}^{i_e} \check{y}_{e,i} \in \mathbb{Z}^d$ at time (tree distance from the root) $\sum_{f\prec e} \sum_{i=1}^{j_f} \check{n}_{f,i} + \sum_{i=1}^{i_e} \check{n}_{e,i} \in \mathbb{Z}^d$.

Let

$$t_{\check{\boldsymbol{n}}}^{(F)}(\check{\boldsymbol{y}}) = \rho\mathbb{P}\Big(\mathcal{T} \in \check{\boldsymbol{T}}(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}})\Big). \tag{2.22}$$

Given n, F, and $\check{\boldsymbol{n}}$ as above, and $\check{\boldsymbol{k}} = (\check{k}_{e,i})_{i\in[j_e],e\in[2r-1]}$ with each $\check{k}_{e,i} \in \mathbb{R}^d$, define

$$\hat{t}_{\check{\boldsymbol{n}}}^{(F)}(\check{\boldsymbol{k}}) = \sum_{\check{\boldsymbol{y}}} \prod_{e=1}^{2r-1} \prod_{i=1}^{j_e} \mathrm{e}^{\mathrm{i}\check{k}_{e,i}\cdot\check{y}_{e,i}} t_{\check{\boldsymbol{n}}}^{(F)}(\check{\boldsymbol{y}}).$$

The following proposition will be proved in Sect. 4.5 via modifications of [17, Theorem 4.8] (where each $j_e = 1$) as indicated in [18]:

Proposition 2.11. *Fix $d > 8$. There exists $L_0(d)$ such that for every $L \geq L_0$: for every $\delta \in (0, 1 \wedge \frac{d-8}{2})$, $\varepsilon > 0$, $r \in \mathbb{N}$, $F \in \Sigma_r$, $(j_e)_{e\in[2r-1]} \in \mathbb{N}^{2r-1}$, $\check{\boldsymbol{n}} = (\check{n}_{e,i})_{i\in[j_e],e\in[2d-1]}$ (with each $\check{n}_{e,i} \in \mathbb{N}$ and each $\check{n}_{e,i}/n \in (\varepsilon, 1/\varepsilon)$), $R > 0$, $\check{\boldsymbol{k}} = (\check{k}_{e,i})_{i\in[j_e],e\in[2d-1]}$ (with each $\check{k}_{e,i} \in [-R, R]^d$),*

$$\hat{t}_{\check{\boldsymbol{n}}}^{(F)}\Big(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\Big) = \rho C_V^{r-1} C_A^{2r-1} \prod_{e=1}^{2r-1} \prod_{i=1}^{j_e} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \frac{\check{n}_{e,i}}{n}} + \mathcal{O}\Big(\sum_{e=1}^{2r-1} \sum_{i=1}^{j_e} \frac{1}{\check{n}_{e,i}^{\frac{d-8}{2}}}\Big) + \mathcal{O}\Big(\sum_{e=1}^{2r-1} \sum_{i=1}^{j_e} \frac{\big|\check{\boldsymbol{k}}\big|^2 \check{n}_{e,i}^{1-\delta}}{n}\Big),$$

where the constants in the error terms depend on L, δ, r, R, $(j_e)_{e\in[2r-1]}$ and $\varepsilon > 0$.

The purpose of this section is to prove Theorem 2.7 using Proposition 2.11.

We begin with generalisations of (2.17) and (2.18) (where $r = 1$). Fix $r \geq 1$ and $t_1, \dots, t_r > 0$. Let $\boldsymbol{s} = (\vec{s}^{(1)}, \dots, \vec{s}^{(r)})$, where $\vec{s}^{(\ell)} = (s_0^{(\ell)}, \dots, s_{m^{(\ell)}}^{(\ell)})$, and $0 = s_0^{(\ell)} < s_1^{(\ell)} < \cdots < s_{m^{(\ell)}}^{(\ell)} = t_\ell$ for each ℓ (so each $m^{(\ell)} \in \mathbb{N}$). Then for $\phi^{(1)}, \dots, \phi^{(r)} \in \mathcal{G}$ (with $\phi^{(\ell)} = \phi_{\vec{s}^{(\ell)},\vec{k}^{(\ell)}}$ and $\vec{k}^{(\ell)} = (k_1^{(\ell)}, \dots, k_{m^{(\ell)}}^{(\ell)}) \in (\mathbb{R}^d)^{m^{(\ell)}}$),

$$\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})$$
$$= \Big(\frac{1}{C_0 n}\Big)^r \sum_{\sqrt{n}x_1 \in \mathcal{T}_{nt_1}} \cdots \sum_{\sqrt{n}x_r \in \mathcal{T}_{nt_r}} \prod_{\ell=1}^{r} \prod_{j_\ell=1}^{m^{(\ell)}} \exp\Big\{ \mathrm{i} k_{j_\ell}^{(\ell)} \big(w^{(n)}_{s^{(\ell)}_{j_\ell}}(t_\ell, x_\ell) - w^{(n)}_{s^{(\ell)}_{j_\ell - 1}}(t_\ell, x_\ell) \big) \Big\}.$$

Take expectations and work with the un-normalised functions $w(t,x) = w(t,x)\big(T\big)$ (a slight abuse of notation, as before $w(t,x)$ was defined as a function of the random tree $\mathcal{T}$) to see that

$$E_{\mu_n^{\mathrm{LT}}}\Big[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)})\Big] = \frac{C_1}{C_0^r n^{r-1}} \sum_{\vec{x} \in (\mathbb{Z}^d)^r} \sum_{\substack{T \ni o: \\ \vec{x} \in T_{n\vec{t}}}} \mathbb{P}(\mathcal{T} = T)$$
$$\times \prod_{\ell=1}^{r} \prod_{j_\ell=1}^{m^{(\ell)}} \exp\Big\{ \mathrm{i} \frac{k_{j_\ell}^{(\ell)}}{\sqrt{n}} \big(w_{ns^{(\ell)}_{j_\ell}}(nt_\ell, x_\ell) - w_{ns^{(\ell)}_{j_\ell - 1}}(nt_\ell, x_\ell) \big) \Big\}, \tag{2.23}$$

where $\vec{x} \in T_{n\vec{t}}$ means $x_i \in T_{nt_i}$ for each $i \in [r]$.

Given $\vec{x} = (x_1, \ldots, x_r) \in (\mathbb{Z}^d)^r$ and $T \ni o$ a lattice tree with $x_1, \ldots, x_r \in T$, one can consider the minimal subtree containing the origin and these points. Typically this subtree has $r-1$ branch points that are connected to the root and the points x_i according to an abstract (rooted) shape F consisting of $2r-1$ edges $e \in \mathcal{E}(F)$ and $2r$ vertices. Call this the shape associated to $(T, \vec{x})$. Contributions from subtrees containing fewer than $r-1$ branch points (arising if (i) the number of distinct elements in $\{x_1, \ldots, x_r\}$ is smaller than r, or (ii) paths in T to one or more x_i contain paths to one or more other x_j, or (iii) the most recent common ancestor of two x_j's is the origin, or (iv) some branch point in the subtree has degree more than 3) will constitute error terms (see e.g. (2.26) below) and they will be said to have a *degenerate* shape. For a given (non-degenerate) shape $F \in \Sigma_r$, and $\vec{t} = (t_1, \ldots, t_r) \in (\mathbb{R}_{>0})^r$, recall the definition of $\mathcal{M}_n(\vec{t}, F)$ from (2.14) (but now with ℓ in place of β^ℓ). For $\vec{x} \in (\mathbb{Z}^d)^r$, $\vec{y} = (y_{r+1}, \ldots, y_{2r-1}) \in (\mathbb{Z}^d)^{r-1}$, and $\vec{u} \in \mathcal{M}_n(\vec{t}, F)$, let $\boldsymbol{T}_n(F, \vec{t}, \vec{u}, \vec{x}, \vec{y})$ denote the set of lattice trees T containing the origin and the points $x_i \in T_{\lfloor nt_i \rfloor}$ for $i \in [r]$ for which the shape associated to $(T, \vec{x})$ is F, such that for each branch point $j = r+1, \ldots, 2r-1$ in F, the spatial and temporal location of the corresponding branch point in T is (y_j, nu_j). The main contribution to (2.23) is therefore

$$\frac{C_1 \rho^{-1}}{C_0^r n^{r-1}} \sum_{F \in \Sigma_r} \sum_{\substack{\vec{u} \in \\ \mathcal{M}_n(\vec{t}, F)}} \sum_{\substack{(\vec{x}, \vec{y}) \in \\ (\mathbb{Z}^d)^{2r-1}}} \sum_{\substack{T \in \\ \boldsymbol{T}_n(F, \vec{t}, \vec{u}, \vec{x}, \vec{y})}} W(T)$$
$$\times \prod_{\ell=1}^{r} \prod_{j_\ell=1}^{m^{(\ell)}} \exp\Big\{ \mathrm{i} \frac{k_{j_\ell}^{(\ell)}}{\sqrt{n}} \big(w_{ns^{(\ell)}_{j_\ell}}(nt_\ell, x_\ell) - w_{ns^{(\ell)}_{j_\ell - 1}}(nt_\ell, x_\ell) \big) \Big\}. \tag{2.24}$$

The modulus of each exponential is bounded by 1. Next, using (2.19), and neglecting interaction between parts of the tree corresponding to the $2r-1$ different edges in the shape we get that for any shape $F \in \Sigma_r$,

$$\sum_{(\vec{x}, \vec{y})} \sum_{\substack{T \in \\ \boldsymbol{T}_n(F, \vec{t}, \vec{u}, \vec{x}, \vec{y})}} W(T) \le K_0^{2r-1}, \tag{2.25}$$

for some $K_0 > 0$. If $\bar{t} = \max_{i\in[r]} t_i$, we can sum over $\vec{u}$ to conclude that

$$\sum_{\substack{\vec{u}\in\\ \mathcal{M}_n(\vec{t},F)}} \sum_{(\vec{x},\vec{y})} \sum_{\substack{T\in\\ \boldsymbol{T}_n(F,\vec{t},\vec{u},\vec{x},\vec{y})}} W(T) \le K_0^{2r-1}(n\bar{t}+1)^{r-1}.$$

Remark 2.12. Bounds similar to (2.25) hold in great generality. For any abstract rooted tree graph (call it a generalised shape) F^* with edge set E^*, and any set of temporal lengths $(n_e)_{e\in E^*}$ (with each $n_e \in \mathbb{N}$) associated with those edges: the total weight of all lattice trees containing the origin having vertices with spatial and temporal displacements $(\vec{\Delta}_e)_{e\in E^*}$ and $(n_e)_{e\in E^*}$ with the generalised shape of the connections to these points being F^*, summed over $(\vec{\Delta}_e)_{e\in E^*}$ gives at most $K_0^{\#E^*}$. This is also obtained by ignoring interactions between different parts of the trees corresponding to different edges in E^*. ★

For degenerate shapes, one also has (2.25) (in fact the exponent $2r-1$ can be reduced). However, in comparison with (2.24), degenerate shapes give rise to sums over fewer (at most $r-2$ in fact) u_j's, each of which takes at most $n\bar{t}+1$ possible values. After summing over finitely many degenerate shapes and summing over $\vec{u}$ we may bound the version of (2.24) for degenerate shapes by

$$\frac{C}{n^{r-1}}(n\bar{t}+1)^{r-2} \le C\frac{(\bar{t}+1)^{r-2}}{n}. \tag{2.26}$$

We conclude that contributions to (2.23) from degenerate shapes are bounded in absolute value by $Cn^{-1}(\bar{t}+1)^{r-2}$ and the main contribution from non-degenerate shapes is at most $C(\bar{t}+1)^{r-1}$. If we set $m^{(\ell)} = 1, \vec{k}^{(\ell)} = 0$, we conclude the following as a special case:

Lemma 2.13. *For each $r \in \mathbb{N}$ there exists a constant $C_r > 0$ such that for all $t_1, \dots, t_r \ge 0$,*

$$\sup_{n\in\mathbb{N}} n\mathbb{E}\Big[\prod_{i=1}^{r} H_{t_i}^{(n)}(1)\Big] \le C_r(\bar{t}+1)^{r-1}.$$

Given $\varepsilon > 0$, $\vec{t}$, $\boldsymbol{s}$, and a (non-degenerate) shape $F \in \Sigma_r$, let $\mathcal{M}_{n,\varepsilon}(\vec{t}, F, \boldsymbol{s})$ denote the set of $\vec{u} \in \mathcal{M}_n(\vec{t}, F)$ for which (with $u_0 := 0$) either:

- there exist a leaf $\ell \in \{1, \dots, r\}$, a branch point $j \in \{r+1, \dots, 2r-1\}$ in the path from o to ℓ, and an $i \in \{1, \dots, m^{(\ell)}\}$, such that
$$|u_j - s_i^{(\ell)}| \le \varepsilon,$$
- there exist $i, j \in \{0, \dots, 2r-1\}$ vertices of F, such that i is an ancestor of j in F and
$$|u_i - u_j| \le \varepsilon.$$

Roughly speaking these correspond to situations where there is branching on a path close to one of the observation times along the path, or where one of the edge-lengths is short.

Let $\mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s}) = \mathcal{M}_n(\vec{t}, F) \setminus \mathcal{M}_{n,\varepsilon}(\vec{t}, F, \boldsymbol{s})$. Then the sum over $\vec{u}$ in (2.24) can be split into a sum over $\vec{u} \in \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})$ and a sum over $\vec{u} \in \mathcal{M}_{n,\varepsilon}(\vec{t}, F, \boldsymbol{s})$.

Using the same argument as for (2.26), we get that the absolute value of the sum over $\vec{u} \in \mathcal{M}_{n,\varepsilon}(\vec{t}, F, \boldsymbol{s})$ is at most

$$n^{r-1} C\varepsilon(\bar{t}+1)^{r-1} n^{-(r-1)} = C(\bar{t}+1)^{r-1}\varepsilon. \tag{2.27}$$

We therefore turn our attention to the quantity

$$\frac{C_1\rho^{-1}}{C_0^r n^{r-1}} \sum_{F\in\Sigma_r} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \sum_{\substack{(\vec{x},\vec{y})\in \\ (\mathbb{Z}^d)^{2r-1}}} \sum_{\substack{T\in \\ \boldsymbol{T}_n(F,\vec{t},\vec{u},\vec{x},\vec{y})}} W(T) \prod_{\ell=1}^{r}\prod_{j_\ell=1}^{m^{(\ell)}} \mathrm{e}^{\mathrm{i}\frac{k^{(\ell)}_{j_\ell}}{\sqrt{n}}\left(w_{ns^{(\ell)}_{j_\ell}}(nt_\ell,x_\ell) - w_{ns^{(\ell)}_{j_\ell-1}}(nt_\ell,x_\ell)\right)}. \tag{2.28}$$

We now define discrete analogues of the sets $\mathcal{I}$ following Definition 2.5. Recall the notation (2.10). Let $F \in \Sigma_r$, $\vec{t} \in (0,\infty)^r$, $\vec{u} \in \mathcal{M}_n(\vec{t}, F)$, and $\boldsymbol{s} = (\vec{s}^{(1)}, \dots, \vec{s}^{(r)})$, where $\vec{s}^{(\ell)} = (s_0^{(\ell)}, \dots, s_{m^{(\ell)}}^{(\ell)})$, and $0 = s_0^{(\ell)} < s_1^{(\ell)} < \cdots < s_{m^{(\ell)}}^{(\ell)} = t_\ell$ for each $\ell \in [r]$ be given. If $e \notin \mathcal{E}_\ell(F)$ then set $\mathcal{I}_n(e, \vec{s}^{(\ell)}) = \varnothing$. If $e \in \mathcal{E}_\ell(F)$, then let $\mathcal{I}_n(e, \vec{s}^{(\ell)})$ denote those elements of $[\vec{s}^{(\ell)}]_n := ([s_1^{(\ell)}]_n, \dots, [s_{m^{(\ell)}}^{(\ell)}]_n)$ that fall in the interval $(u_-(e), u_+(e) \wedge [t_\ell]_n)$, where $u_-(e)$, $u_+(e)$ are the elements of $\vec{u}$ corresponding to the end-vertices of e (and $u_-(e) = 0$ if e is adjacent to the root). Let $\mathcal{I}_n(e, \boldsymbol{s}) = \cup_{\ell=1}^r \mathcal{I}_n(e, \vec{s}^{(\ell)})$. The $j(e) := |\mathcal{I}_n(e, \boldsymbol{s})|$ elements of $\mathcal{I}_n(e, \boldsymbol{s})$ divide the interval $[u_-(e), u_+(e) \wedge [t_\ell]_n]$ into $j(e)+1$ subintervals - denote their lengths by $(\check{n}_{e,i}/n)_{i=1,\dots,j(e)+1}$, and set $\check{\boldsymbol{n}} = (\check{n}_{e,i})_{e\in\mathcal{E}(F);i=1,\dots,j(e)+1}$. If $j(e) = 0$ then $\check{n}_{e,1}/n = u_+(e) - u_-(e)$. Note that $j(e)$ and $\check{\boldsymbol{n}}$ depend on $F, \vec{u}, \boldsymbol{s}$ (and n), and that $\sum_{e\in\mathcal{E}_\ell(F)} \sum_{j=1}^{j(e)+1} \check{n}_{e,j} = \lfloor nt_\ell \rfloor$.

For $\ell \in [r]$, $e \in \mathcal{E}_\ell(F)$ and $a \in \{1, \dots, j(e)+1\}$ let

$$\zeta_n^{[e]}(a,\ell) = \min\left\{ i \le m^{(\ell)} : s_i^{(\ell)} \ge u_-(e) + \sum_{i_e=1}^{a} \frac{\check{n}_{e,i_e}}{n} \right\}.$$

(Note that $s_i^{(\ell)}$ is interchangeable with $[s_i^{(\ell)}]_n$ in the definition of $\zeta_n^{[e]}$.) Given $\vec{k}^{(\ell)} = (k_1^{(\ell)}, \dots, k_{m^{(\ell)}}^{(\ell)}) \in (\mathbb{R}^d)^{m^{(\ell)}}$, for each $\ell \in [r]$, and for $e \in \{1, \dots, 2r-1\}$ and $a \le j(e)+1$, let

$$\check{k}_{e,a}(n) = \sum_{\ell : e\in\mathcal{E}_\ell(F)} k^{(\ell)}_{\zeta_n^{[e]}(a,\ell)}.$$

Let $\check{\boldsymbol{k}}(n) = (\check{k}_{e,i}(n))_{e\in[2r-1], i\le j(e)+1}$ which depends on $F, \boldsymbol{s}, \vec{u}, n$ and of course $\boldsymbol{k}$.

If $n \in \mathbb{N}$, $F \in \Sigma_r$, $\boldsymbol{s}$, and $\vec{u} \in \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})$ are given, this determines $\check{\boldsymbol{n}} = \check{\boldsymbol{n}}(F, \boldsymbol{s}, \vec{u})$ as above. If we are given $\boldsymbol{k}$ as well then this also determines $\check{\boldsymbol{k}}(n)$. By expressing locations of paths in terms of their spatial increments $\check{\boldsymbol{y}} = (\check{y}_{e,i})_{i\in[j_e], e\in[2r-1]}$ (and recalling the definition of $\check{\boldsymbol{T}}(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}})$ given prior to (2.22)) we see that (2.28) is equal to

$$\frac{C_1\rho^{-1}}{C_0^r n^{r-1}} \cdot \sum_{F\in\Sigma_r} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \sum_{\check{\boldsymbol{y}}} \sum_{\substack{T\in \\ \check{\boldsymbol{T}}(F,\check{\boldsymbol{y}},\check{\boldsymbol{n}}(F,\boldsymbol{s},\vec{u}))}} W(T) \prod_{e=1}^{2r-1}\prod_{i=1}^{j(e)+1} \mathrm{e}^{\mathrm{i}\frac{\check{k}_{e,i}(n)}{\sqrt{n}}\cdot\check{y}_{e,i}}$$

$$= \frac{C_1}{C_0^r n^{r-1}} \cdot \sum_{F\in\Sigma_r} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \sum_{\check{\boldsymbol{y}}} \prod_{e=1}^{2r-1}\prod_{i=1}^{j(e)+1} \mathrm{e}^{\mathrm{i}\frac{\check{k}_{e,i}(n)}{\sqrt{n}}\cdot\check{y}_{e,i}} \mathbb{P}\Big(\mathcal{T} \in \check{\boldsymbol{T}}\big(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}}(F,\boldsymbol{s},\vec{u})\big)\Big), \tag{2.29}$$

Recall $\mathcal{M}(\vec{t}, F)$ from Definition 2.5. Given $\varepsilon > 0$ we define $\mathcal{M}_\varepsilon(\vec{t}, F, \boldsymbol{s})$ to be the set of $\vec{u} \in \mathcal{M}(\vec{t}, F)$ for which either:

- there exist a leaf $\ell \in \{1, \dots, r\}$, a branch point j in the path from o to ℓ in F, and $i \in \{1, \dots, m^{(\ell)}\}$ such that

$$|u_j - s_i^{(\ell)}| \le \varepsilon.$$

- there exists vertices $i \prec j$ of F, such that

$$|u_i - u_j| \le \varepsilon.$$

Let $\mathcal{M}_*(\vec{t}, F, \boldsymbol{s}) = \mathcal{M}(\vec{t}, F) \setminus \mathcal{M}_\varepsilon(\vec{t}, F, \boldsymbol{s})$. Then, as for (2.27), we have that

$$\int_{\vec{u} \in \mathcal{M}_\varepsilon(\vec{t}, F, \boldsymbol{s})} 1 d\vec{u} < C_r \varepsilon \bar{t}$$

Recall the definition of Φ (and its arguments) from (2.6). Below we will show that as $n \to \infty$ (2.29) converges to

$$\sum_{F \in \Sigma_r} \int_{\vec{u} \in \mathcal{M}_*(\vec{t}, F, \boldsymbol{s})} \Phi(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) d\vec{u}. \tag{2.30}$$

Fix $F \in \Sigma_r$ and consider the quantity in (2.29) with fixed F which can be written as

$$\frac{C_1}{C_0^r n^{r-1}} \sum_{\substack{\vec{u} \in \\ \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})}} \sum_{\check{\boldsymbol{y}}} \prod_{e=1}^{2r-1} \prod_{i=1}^{j(e)+1} \mathrm{e}^{\mathrm{i} \frac{\check{k}_{e,i}(n)}{\sqrt{n}} \cdot \check{y}_{e,i}} \mathbb{P}\Big(\mathcal{T} \in \check{\boldsymbol{T}}(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}}(F, \vec{u}, \boldsymbol{s}))\Big). \tag{2.31}$$

Then (2.31) is equal to

$$\frac{C_1 \rho^{-1}}{C_0^r n^{r-1}} \sum_{\substack{\vec{u} \in \\ \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})}} \hat{t}^{(F)}_{\check{\boldsymbol{n}}(F, \boldsymbol{s}, \vec{u})}\Big(\frac{\check{\boldsymbol{k}}(n)}{\sqrt{n}}\Big), \tag{2.32}$$

where we recall that $\check{\boldsymbol{k}}(n)$ depends on $F, \boldsymbol{s}, \vec{u}, n, \boldsymbol{k}$.

Proof of Theorem 2.7. Fix $r, \vec{t}$ and the $\phi^{(\ell)}$ (hence $\boldsymbol{k}$ and $\boldsymbol{s}$).

Let $\delta(\boldsymbol{s}) > 0$ denote the minimum difference between distinct values in $\boldsymbol{s}$ (recall that this includes 0 and each t_ℓ). Let $\varepsilon \in (0, (\delta(\boldsymbol{s})/2) \wedge 1)$. Above (see in particular (2.26), (2.27) and (2.32)), we have shown that the left hand side of (2.7) is equal to

$$\frac{C_1 \rho^{-1}}{C_0^r n^{r-1}} \sum_{F \in \Sigma_r} \sum_{\substack{\vec{u} \in \\ \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})}} \hat{t}^{(F)}_{\check{\boldsymbol{n}}(F, \boldsymbol{s}, \vec{u})}\Big(\frac{\check{\boldsymbol{k}}(n)}{\sqrt{n}}\Big) + \mathcal{O}(\varepsilon) + \mathcal{O}(n^{-1}),$$

where the constants in the $\mathcal{O}$ notation here only depend on $\bar{t}, r, L, d$. By definition of $\check{\boldsymbol{n}}$, each $\check{n}_{e,i}$ is equal to $\lfloor ns \rfloor - \lfloor ns' \rfloor$ for some distinct $s > s' \in \boldsymbol{s}$ (or is equal to $|\lfloor ns_i^{(\ell)} \rfloor - nu_j|$ for some branch point j in the path from o to ℓ in F, or $|nu_i - nu_j|$ for some $i \prec j$ in F). It follows from the definition of $\delta(\boldsymbol{s})$ and the fact that $\vec{u} \in \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})$ that we have

that $\check{n}_{e,i} > n\varepsilon/2$ for all e, i for n sufficiently large depending on ε (which we assume in what follows). By Proposition 2.11 (recalling that $C_0 = C_A^2 C_V$ and $C_1 = C_A C_V$, and $\delta \in (0,1)$ is as in Proposition 2.11) we see that this is equal to

$$\sum_{F\in\Sigma_r} \frac{1}{n^{r-1}} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \Bigg[\prod_{e=1}^{2r-1} \prod_{i=1}^{j(e)+1} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \frac{\check{n}_{e,i}}{n}} + \mathcal{O}\Big(\sum_{e,i} \frac{1}{\check{n}_{e,i}^{\frac{d-8}{2}}} \Big) + \mathcal{O}\Big(\sum_{e,i} \frac{|\check{\boldsymbol{k}}|^2 \check{n}_{e,i}^{1-\delta}}{n} \Big) \Bigg]$$
$$+ \mathcal{O}(\varepsilon) + \mathcal{O}(n^{-1}), \tag{2.33}$$

where in the above, $\check{\boldsymbol{n}}$ is determined by $F, \vec{u}, \boldsymbol{s}$ (and n), and $\check{\boldsymbol{k}}$ is determined by these and $\boldsymbol{k}$. In addition the constants in the error terms in square brackets depends on ε (among other things, as in Proposition 2.11). Also $\delta, \varepsilon \in (0,1)$ and $n\varepsilon/2 \le \check{n}_{e,i} \le n\bar{t}$ imply the error terms in square brackets are $\mathcal{O}((\varepsilon n)^{-(\delta\wedge(d-8)/2)})$ uniformly in $\vec{u} \in \mathcal{M}_{n,*}(\vec{t}, F, \boldsymbol{s})$ (where again the constant in the $\mathcal{O}$ notation here depends on $\varepsilon, \boldsymbol{k}$). Since the sum over $\vec{u}$ gives at most $(n\bar{t})^{r-1}$ we see that (2.33) is equal to

$$\sum_{F\in\Sigma_r} \frac{1}{n^{r-1}} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \prod_{e=1}^{2r-1} \prod_{i=1}^{j(e)+1} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \frac{\check{n}_{e,i}}{n}} + \mathcal{O}((\varepsilon n)^{-(\delta\wedge(d-8)/2)}) + \mathcal{O}(\varepsilon). \tag{2.34}$$

Recall the definition of $\check{\boldsymbol{s}} = \check{\boldsymbol{s}}(F, \vec{u}, \boldsymbol{s})$ from below (2.4). Together with the definition of $\check{\boldsymbol{n}}$ we see that $|\check{s}_{e,i} - \check{n}_{e,i}/n| \le 2/n$ for every e, i. Thus (for n large enough depending on ε) (2.34) is equal to

$$\sum_{F\in\Sigma_r} \frac{1}{n^{r-1}} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \prod_{e=1}^{2r-1} \prod_{i=1}^{j(e)+1} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \check{s}_{e,i} + \mathcal{O}(n^{-1})} + \mathcal{O}(\varepsilon), \tag{2.35}$$

where the error term in the exponent depends on $\boldsymbol{k}$ but is uniform in $\vec{u}$. Recalling (2.6), it follows that (2.35) is equal to

$$\sum_{F\in\Sigma_r} \frac{1}{n^{r-1}} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \prod_{e=1}^{2r-1} \prod_{i=1}^{j(e)+1} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \check{s}_{e,i}} + \mathcal{O}(\varepsilon)$$
$$= \sum_{F\in\Sigma_r} \frac{1}{n^{r-1}} \sum_{\substack{\vec{u}\in \\ \mathcal{M}_{n,*}(\vec{t},F,\boldsymbol{s})}} \Phi(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) + \mathcal{O}(\varepsilon).$$

As $n \to \infty$ in the above, the $(r-1)$-fold Riemann sum converges to the $(r-1)$-dimensional integral in (2.30). We have therefore shown that there exists a constant C (depending on $\boldsymbol{k}, \boldsymbol{s}, \vec{t}$) such that for any $\varepsilon > 0$, for n sufficiently large we have that

$$\Bigg| E_{\mu_n^{\mathrm{LT}}} \Bigg[\prod_{\ell=1}^{r} H_{t_\ell}^{(n)}(\phi^{(\ell)}) \Bigg] - \sum_{F\in\Sigma_r} \int_{\vec{u}\in\mathcal{M}(\vec{t},\alpha)} \Phi_{\sigma^2}(F, \vec{u}, \boldsymbol{s}, \boldsymbol{k}) d\vec{u} \Bigg| \le C\varepsilon,$$

which completes the proof. □

3. Tightness

In this section we work in an abstract setting for historical processes motivated by the historical paths $\{w(m,x) : m \in \mathbb{Z}_+, x \in \mathcal{T}_m\}$ of lattice trees and those for branching Brownian motion, $\{B^\alpha : |\alpha| \in \mathbb{Z}_+, \alpha \in GW\}$ (with $n = 1$), both introduced in Sect. 1.

As before, add Δ to $\mathbb{R}^d$ as a cemetery point. Assume on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we have

$$\forall k \in \mathbb{Z}_+,\ \mathcal{S}_k \text{ is an a.s. finite random subset of a countable set } \mathcal{S}. \tag{3.1}$$

$$\begin{aligned}&\forall k \in \mathbb{Z}_+, \beta \in \mathcal{S},\ (w_j(k,\beta))_{j\in\mathbb{Z}_+} \text{ are } \mathbb{R}^d \cup \{\Delta\}\text{-valued random variables such that}\\ &\quad \text{for } \beta \in \mathcal{S}_k,\ w_j(k,\beta) \text{ are } \mathbb{R}^d\text{-valued, } w_0(k,\beta) = 0,\ w_j(k,\beta) = w_k(k,\beta)\ \forall j \geq k,\\ &\quad \text{and for } \beta \in \mathcal{S} \setminus \mathcal{S}_k,\ w_j(k,\beta) = \Delta.\end{aligned} \tag{3.2}$$

So for each $k \in \mathbb{Z}_+$ and β in the random finite set $\mathcal{S}_k$ we have a discrete-time $\mathbb{R}^d$-valued stochastic process starting at 0 and freezing at time k.

For

$$w \in \mathcal{W} := \{w(k,\beta) : \beta \in \mathcal{S}_k, k \in \mathbb{Z}_+\} \text{ (the set of historical paths)}, \tag{3.3}$$

we define the rescaled paths by

$$w_s^{(n)}(t,\beta) = \frac{w_{\lfloor ns\rfloor}(\lfloor nt\rfloor,\beta)}{\sqrt{n}}, \quad s,t \geq 0, \tag{3.4}$$

so that for $t \geq 0$ and $\beta \in \mathcal{S}_{\lfloor nt\rfloor}$, $w^{(n)}(t,\beta) \in \mathcal{D}(\mathbb{R}^d)$. Define a càdlàg $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$-valued process by

$$H_t^{(n)} = \frac{1}{C_g n} \sum_{\beta\in\mathcal{S}_{\lfloor nt\rfloor}} \delta_{w^{(n)}(t,\beta)}, \tag{3.5}$$

where $C_g > 0$ is a model-dependent constant. We call this class of measure valued processes, the historical processes associated with $\mathcal{W}$.

Example 3.1 **(Lattice trees).** Here $\mathcal{S} = \mathbb{Z}^d$, $\mathcal{S}_m = \mathcal{T}_m$ for $m \in \mathbb{Z}_+$ and for $x \in \mathcal{S}_m$, $w(m,x)$ is the tree history from the root to (m,x) in (1.1). If $C_g = C_0$ then one can easily check that $H^{(n)}$ as defined in (3.5) agrees with the historical process for lattice trees in (1.4). Note here that the index set for $w^{(n)}$ has been changed from that in (1.2) (and so we have abused the notation) but the actual empirical measures are unchanged. Properties (3.1) and (3.2) are clear if we extend the definition of $w(m,x)$ to Δ for $x \notin \mathcal{S}_m$.

Example 3.2 **(Branching random walk).** We discretize (in time) the branching Brownian motions introduced in Sect. 1 and use the notation from that construction. We denote dependence on $n \in \mathbb{N}$ now in our notation for $\hat{B}^{\beta,(n)}$ for $\beta \in I$. Let $\mathcal{S} = I$, $\mathcal{S}_m = \{\beta \in I : \beta \in GW, |\beta| = m\}$, and for $\beta \in \mathcal{S}_m$ set

$$w_j(m,\beta) = \hat{B}^{\beta,(1)}_{j\wedge m}.$$

Then one can check that for $\alpha \in \mathcal{S}_{\lfloor nt\rfloor}$,

$$w_s^{(n)}(t,\beta) = \frac{\hat{B}^{\beta,(1)}_{\lfloor n(t\wedge s)\rfloor}}{n^{1/2}}.$$

Set $C_g = 1$ in (3.5), and for $|\beta| = \lfloor nt \rfloor$, let $Z_s^{\beta,(n)} = \hat{B}^{\beta,(n)}_{\lfloor n(s\wedge t)\rfloor/n}$ be a time discretization of the stopped Brownian paths $\hat{B}^{\beta,(n)}$. Brownian scaling shows that

$$\text{if } \tilde{H}_t^{(n)} = \frac{1}{n} \sum_{\substack{\beta \in GW: \\ |\beta| = \lfloor nt \rfloor}} \delta_{Z^{\beta,(n)}}, \text{ then } \tilde{H}^{(n)} \text{ is equal in law to}$$
$$\text{the } n\text{th historical process given by (3.5) for each } n \in \mathbb{N}. \tag{3.6}$$

Clearly $\tilde{H}^{(n)}$ is a rescaled branching random walk with Gaussian mean 0, variance σ^2 increments. Properties (3.1) and (3.2) are again clear if we extend the definition of $w(m, x)$ to Δ for $x \notin \mathcal{S}_m$.

In order to prove historical tightness, we will assume that the collection $\mathcal{W}$ (as in (3.3)) of historical paths satisfies the following condition. Recall that $w^{(n)}$ is the scaled version of w, as in (3.4).

Condition 3.3 (*Modulus of continuity*). For some $q \in (0, 1/2)$, $\theta \in (0, 1]$, and constant $C_2 > 0$, there exist random variables $(\delta_n)_{n\in\mathbb{N}}$ so that for all historical paths $w \in \mathcal{W}$ and $n \in \mathbb{N}$,

$$\forall s_i \in \mathbb{Z}_+/n, \quad |s_2 - s_1| \le \delta_n \Rightarrow |w^{(n)}_{s_2} - w^{(n)}_{s_1}| \le |s_2 - s_1|^q,$$
$$\text{where } n\mathbb{P}(\delta_n \le \rho) \le C_2\rho^\theta \quad \forall \rho \in [0, 1). \tag{3.7}$$

This condition is verified for any $q \in (0, 1/2)$ and $\theta = 1$ in [20, Theorem 6] for sufficiently spread-out lattice trees in more than 8 dimensions in Example 3.1 above (as well as a number of other models)—see Lemma 3.12 below. For the Branching Random Walks with Gaussian increments in Example 3.2 it is easy to derive it from [6, Theorem 8.1] for the same parameter values (in fact θ can be taken to be any value in $(0, \infty)$) . Here one takes the underlying diffusion to be Brownian motion, restricts the time steps to be in $\mathbb{Z}_+/n$, and then uses (3.6).

In our abstract setting, the extinction times become

$$S^{(1)} = \min\{k \in \mathbb{Z}_+ : \mathcal{S}_k = \emptyset\} \in \mathbb{Z}_+ \cup \infty,$$

so that

$$S^{(n)} := S^{(1)}/n = \inf\{t \ge 0 : H_t^{(n)}(1) = 0\},$$

agreeing with our earlier definition for lattice trees. We assume $S^{(1)}$ satisfies the following:

Condition 3.4 (*Survival bounds*). There exist $\underline{c}, \overline{c} > 0$ such that

$$\underline{c} \le \inf_{t\ge 0} \mathbb{P}(S^{(1)} > t)(t \vee 1) \le \sup_{t\ge 0} \mathbb{P}(S^{(1)} > t)(t \vee 1) \le \overline{c}. \tag{3.8}$$

This condition holds for the branching random walks in Example 3.2 by Kolmogorov's classical result for survival of critical branching processes (e.g. see [24, Theorem II.1.1(a)]) and for the lattice tree historical paths in Example 3.1 by (1.6) (or see [20, (1.22) and (1.27)]).

Definition 3.5. For a metric space, E, a collection $\{Q_n : n \in \mathbb{N}\}$ of probabilities on $\mathcal{D}(\mathbb{R}_+, E) = \mathcal{D}(E)$, is $\mathcal{C}$*-relatively compact* iff every sequence $n_k \to \infty$ has a subsequence $\{n'_k\}$ s.t. $Q_{n'_k}$ converges weakly in $\mathcal{D}(E)$ to a law, Q, supported on $\mathcal{C}(E)$, the set of continuous E-valued paths. If $\{X_n\}$ is a sequence of càdlàg E-valued processes on our underlying probability space, we say $\{X_n : n \in \mathbb{N}\}$ is $\mathcal{C}^{\text{cond}}$*-relatively compact* iff for every $s_0 > 0$, the set of conditional laws $\{\mathbb{P}(X_n \in \cdot | S^{(n)} > s_0) : n \in \mathbb{N}\}$ is $\mathcal{C}$-relatively compact in $\mathcal{D}(E)$. ◀

We start with a general tightness result for historical processes in this abstract setting:

Theorem 3.6. *Assume $H^{(n)}$ is given by (3.5), where $\mathcal{W}$ satisfies Condition 3.3. Suppose also that Condition 3.4 holds and $\{H^{(n)}_\cdot(\phi) : n \in \mathbb{N}\}$ is $\mathcal{C}^{cond}$-relatively compact in $\mathcal{D}(\mathbb{C})$ for each ϕ in a determining class D_0 (for $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$) containing 1. Then $\{H^{(n)}_\cdot : n \in \mathbb{N}\}$ is $\mathcal{C}^{cond}$-relatively compact, and for every $s_0 > 0$, every limit point, H, of $\{\mathbb{P}(H^{(n)} \in \cdot | S^{(n)} > s_0) : n \in \mathbb{N}\}$ satisfies $H_t(\mathcal{C}(\mathbb{R}^d)^c) = 0$ for all $t \geq 0$ a.s.*

In practice it is the relative compactness of $\{H^{(n)}_\cdot(\phi) : n \in \mathbb{N}\}$ for a rich class of test functions ϕ that will require most of the effort. For LT's this is done in Proposition 3.11, which is in turn proved in Sect. 3.2 below. Applying Theorem 3.6 to the case of lattice trees (conditional on survival), we will then deduce the following below:

Theorem 3.7. *Let $H^{(n)}$ be the sequence of rescaled historical processes associated with sufficiently spread-out lattice trees in $d > 8$ dimensions, defined in (1.4). Then $\{H^{(n)}_\cdot : n \in \mathbb{N}\}$ is $\mathcal{C}^{cond}$-relatively compact.*

3.1. Proofs of Theorems 3.6 and 3.7. Our starting point for proving Theorem 3.6 is a version of the Jakubowski-Kurtz Theorem for $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$-valued processes. It is a simple extension of that for $\mathcal{M}_F(\mathbb{R}^d)$-valued processes in [11, Theorem 5.2].

Theorem 3.8. *Let $D_0 \subset \mathcal{C}(\mathcal{D}(\mathbb{R}^d), \mathbb{C})$ be a determining class for $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$ containing 1.*
A sequence of probabilities $\{P_k, k \in \mathbb{N}\}$ on $\mathcal{D}(\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d)))$ is $\mathcal{C}$-relatively compact iff

$$\forall \eta > 0, \ \forall T \in \mathbb{N}, \ \text{there is a compact set } K_{\eta,T} \subset \mathcal{D}(\mathbb{R}^d) \ \text{such that}$$
$$\sup_k P_k\Big(\sup_{t \leq T} H_t(K^c_{\eta,T}) > \eta\Big) < \eta, \tag{3.9}$$

and

$$\text{for all } \phi \in D_0 \text{ the sequence of probabilities, } \{P_k(H_\cdot(\phi) \in \cdot)\},$$
$$\text{is } \mathcal{C}\text{-relatively compact in } \mathcal{D}(\mathbb{C}). \tag{3.10}$$

For $\delta, T > 0$ and $w \in \mathcal{D}(\mathbb{R}^d)$, we define

$$W'(w, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t \in [t_{i-1}, t_i)} |w_s - w_t|,$$

where the infimum is over all partitions $\{t_i\}$ such that $0 = t_0 < t_1 < \ldots t_{N-1} < T \leq t_N$ such that $t_i - t_{i-1} > \delta$ for all i. Note that W' is decreasing in δ and increasing in T. We restate [8, Ch. 3, Theorem 6.3 and Remark 6.4] with their general metric space E replaced by $\mathbb{R}^d$ and use the above monotonicity to take sequential limits and restrict $T \in \mathbb{N}$.

Proposition 3.9. *Let $\delta'_m \downarrow 0$. The closure of a set $A \subset \mathcal{D}(\mathbb{R}^d)$ is compact iff*

$$\sup_{w\in A, t\le T} |w_t| < \infty \quad \textit{and} \quad \lim_{m\to\infty} \sup_{w\in A} W'(w, \delta'_m, T) = 0, \quad \forall\, T \in \mathbb{N}.$$

For $0<q<1/2$ let $B_m=B_m(q)=\{w \in \mathcal{D}(\mathbb{R}^d) : W'(w, 2^{-m}, T)\le 2^{-(m-2)q} \forall\, T\in\mathbb{N}\}$, and for $M \in \mathbb{N}$ define

$$A_M = A_M(q) = \{w \in \mathcal{D}(\mathbb{R}^d) : |w_t| \le (t+1)2^{M+1} \ \forall t \ge 0\} \cap \left(\cap_{m=M}^{\infty} B_m\right).$$

An easy application of Proposition 3.9 shows that A_M has compact closure in $\mathcal{D}(\mathbb{R}^d)$.

Lemma 3.10. *Assume Condition 3.3, and let q, δ_n be as in (3.7). For any $n, M \in \mathbb{N}$, if $\delta_n > \max(2^{2-M}, n^{-1})$, then $H_t^{(n)}(A_M^c) = 0$ for all $t \ge 0$.*

Proof. Assume $\delta_n \ge \max(2^{2-M}, n^{-1})$, and let $m \in \mathbb{N}^{\ge M}$, $T \in \mathbb{N}$ and $w \in \mathcal{W}$. If we divide $[0, \lceil t\rceil]$ into $\lceil t\rceil 2^M$ intervals of length $2^{-M} < \delta_n$, then the triangle inequality, (3.7) and $\delta_n \ge n^{-1}$ imply

$$|w_t^{(n)}| = |w_{[t]_n}^{(n)} - w_0^{(n)}| \le \lceil t\rceil 2^M [2^{-Mq} + n^{-q}] \le (t+1)2^{M+1}, \tag{3.11}$$

where in the first inequality we have moved an interval endpoint to an appropriate neighbouring point in $\mathbb{Z}_+/n$ resulting in an error of at most n^{-q}. Consider next $W'(w^{(n)}, 2^{-m}, T)$ for $w \in \mathcal{W}$. If $2^{-m} < \frac{1}{n}$, then $W'(w^{(n)}, 2^{-m}, T) = 0$, as one can see by taking $t_i = \frac{i}{n}$, $i \in \mathbb{Z}_+$ in the definition of W', and using the fact that $w^{(n)}$ is constant on $[i/n, (i+1)/n)$ for $i \in \mathbb{Z}_+$. Assume therefore that $2^{-m} \ge \frac{1}{n}$. Now set $t_i = i2^{-m+1}$, for $i \in \mathbb{Z}_+$, which gives $t_i - t_{i-1} > 2^{-m}$ for all i. We also have

$$[t_i]_n - [t_{i-1}]_n \le 2^{1-m} + \frac{1}{n} < 2^{2-m} \le 2^{2-M} \le \delta_n. \tag{3.12}$$

By (3.7) this implies that for $s, t \in [t_{i-1}, t_i)$

$$|w_t^{(n)} - w_s^{(n)}| = |w_{[t]_n}^{(n)} - w_{[s]_n}^{(n)}| \le |[t]_n - [s]_n|^q \le 2^{-(m-2)q},$$

where in the last line we have used the middle expression in (3.12). This proves that $W'(w^{(n)}, 2^{-m}, T) \le 2^{-(m-2)q}$, which together with (3.11), shows that $w^{(n)} \in A_M$, and so completes the proof. □

Proof of Theorem 3.6. Let $n_k \to \infty$, fix $s_0 > 0$, and define probabilities on $\mathcal{D}(\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d)))$ by

$$P_{n_k}(\cdot) = \mathbb{P}(H^{(n_k)} \in \cdot | S^{(n_k)} > s_0).$$

For the first assertion we need to show this sequence of probability laws are $\mathcal{C}$-relatively compact on $\mathcal{D}(\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d)))$. For this we will use Theorem 3.8, and so need to verify the hypotheses of that result. For (3.9), for all $T \in \mathbb{N}$ we set $K_{\eta,T} = \overline{A_M}$, where M is chosen below. The compactness of this set follows from Proposition 3.9, as has already been noted above. By Lemma 3.10,

$$\begin{aligned} P_{n_k}\big(\sup_t H_t(K_{\eta,T}^c) > 0\big) &\le \mathbb{P}\big(H_t^{(n_k)}(A_M^c) > 0 \ \text{ for some } t \ge 0 \big| \, S^{(n_k)} > s_0\big) \\ &\le \mathbb{P}\big(\delta_{n_k} \le \max(2^{2-M}, n_k^{-1})\big)/\mathbb{P}(S^{(n_k)} > s_0) \\ &\le \underline{c}^{-1}(n_k s_0 + 1) n_k^{-1} C_2 (2^{(2-M)\theta} + n_k^{-\theta}), \end{aligned}$$

where in the last inequality we have used (3.8) and (3.7). The above bound is at most $\underline{c}^{-1}(s_0+1)C_2(2^{(2-M)\theta}+n_k^{-\theta})$ which will be smaller than η if we set $M = M(\eta)$ large enough and assume $n_k > N(\eta)$. This proves (3.9) for large enough k. It is easy to enlarge $K_{\eta,T}$ to obtain a compact set which satisfies (3.9) for all k. For example, for fixed $n = n_k \le N(\eta)$ and all $t \ge 0$, $H_t^{(n)}$ is supported on the space of càdlàg paths which are constant on $[i/n, (i+1)/n)$ and on $[S^{(n)}, \infty)$, and whose jumps are uniformly bounded in absolute value by

$$\max_{m/n \le S^{(n)}, \beta \in \mathcal{S}_m, 1 \le j \le m} \frac{|w_j(m,\beta) - w_{j-1}(m,\beta)|}{n^{1/2}} < \infty \quad P_n - a.s. \tag{3.13}$$

Now use (3.8) to bound $S^{(n)}$ and bound the upper bound on the jumps in (3.13), with high P_n-probability, and so obtain a compact set of paths which supports $H_t^{(n)}$ for all $t \ge 0$ with P_n probability at least $1-\eta$, for the finite many values of $n = n_k \le N(\eta)$.

The other condition (3.10) of Theorem 3.8 holds by assumption and so the $\mathcal{C}$-relative compactness is established.

For the last statement we note first that if $\Delta w_t = w_t - w_{t-}$ for $w \in \mathcal{D}(\mathbb{R}^d)$ and $t > 0$, then a simple Skorokhod topology exercise (e.g. use [8, Chapter 3, Proposition 5.3]) shows that for any $\delta > 0$,

$$\Big\{ w \in \mathcal{D}(\mathbb{R}^d) : \sup_{s>0} |\Delta w_s| \le \delta \Big\} \text{ is a closed set in } \mathcal{D}(\mathbb{R}^d).$$

Consider a weak limit H of $\{P_{n_k}\}$. By Skorokhod's representation theorem and the continuity of the limit point, $H_\cdot$, we may realize all our processes on a space with underlying law $\mathbb{P}'$ and assume $H_t^{(n_k)} \to H_t$ in $\mathcal{M}_F(\mathcal{D}(\mathbb{R}^d))$ for all $t \ge 0$, $\mathbb{P}'$-a.s. So the Portmanteau Theorem for the weak topology gives for all $t \ge 0$ and $M \in \mathbb{N}$,

$$H_t\Big(\big\{\sup_{s>0}|\Delta w_s| \le 1/M\big\}^c\Big) \le \liminf_{k\to\infty} H_t^{(n_k)}\Big(\big\{\sup_{s>0}|\Delta w_s| \le 1/M\big\}^c\Big).$$

Now fix $t > 0$ and use Fatou's Lemma to see that for $\delta > 0$,

$$\begin{aligned}
\mathbb{P}'\Big(H_t\big(\{\sup_{s>0}|\Delta w_s| > 1/M\}\big) > \delta\Big) &\le \mathbb{P}'\Big(\liminf_{k\to\infty} H_t^{(n_k)}(\{\sup_{s>0}|\Delta w_s| > 1/M\}) > \delta\Big) \\
&\le \liminf_{k\to\infty} \mathbb{P}'\Big(H_t^{(n_k)}(\{\sup_{s>0}|\Delta w_s| > 1/M\}) > \delta\Big) \\
&\le \liminf_{k\to\infty} P_{n_k}(\delta_{n_k} < 1/n_k) = 0.
\end{aligned}$$

In the last inequality we use the fact that for k large enough $\delta_{n_k} \ge 1/n_k$ implies that for all ancestral paths, and all $s > 0$, $|\Delta w_s^{(n_k)}| \le (1/n_k)^q < 1/M$, and in the final equality we use Conditions 3.3 and 3.4. Now let $M \uparrow \infty$ to see that H_t is supported by $\mathcal{C} = \mathcal{C}(\mathbb{R}^d)$ a.s. for each $t > 0$. Therefore $H_t(\mathcal{C}^c) = 0\ \forall t \in \mathbb{Q}^{>0}$. So using the openness of $\mathcal{C}^c$ and the Portmanteau theorem again, we get from the continuity of $t \to H_t$ that $H_t(\mathcal{C}^c) = 0$ for all $t \ge 0$ a.s. □

Let Lip_K denote the set of functions $\phi : \mathcal{D}(\mathbb{R}^d) \to \mathbb{R}$ such that for each $w, w' \in \mathcal{D}(\mathbb{R}^d)$, $|\phi(w)| \le K$ and $|\phi(w) - \phi(w')| \le K\|w - w'\|$, where $\|w\| = \sup_{t\in\mathbb{R}_+}|w_t|$.

Proposition 3.11. *For critical sufficiently spread-out lattice trees in dimensions $d > 8$: For each $\phi \in \mathrm{Lip}_1$, $\{H_\cdot^{(n)}(\phi) : n \in \mathbb{N}\}$ is $\mathcal{C}^{cond}$-relatively compact in $\mathcal{D}(\mathbb{R})$.*

The proof of this key result is more complicated and so is deferred until Sect. 3.2. Assuming this, Theorem 3.7 now follows:

Proof of Theorem 3.7. We have already noted that the historical process for lattice trees is a special case of the general framework in this Section, that Condition 3.3 was verified in [20] with $q = 1/4$ and $\theta = 1$ (see Lemma 3.12 below), and Condition 3.4 holds by (1.6). Proposition 3.11 shows the last hypothesis of Theorem 3.6 holds with $D_0 = \mathrm{Lip}_1$. D_0 is a determining class because it includes appropriate multiples of all finite-dimensional Lipschitz continuous functions. The result now follows from Theorem 3.6. □

One can also prove the analogue of Theorem 3.7 for the branching random walks in Example 3.2, where the analogue of Proposition 3.11 yields easily to martingale methods, but the convergence results here can be readily proved as in [24, Chapter II].

3.2. Tightness for lattice trees. The goal of this section is to prove Proposition 3.11. For lattice trees, we will use the modulus of continuity in the following form:

Lemma 3.12. *For each $n \in \mathbb{N}$ there exists a random $\delta_n \geq \frac{1}{n}$ and a constant $c > 0$ satisfying $n\mathbb{P}(\delta_n \leq \rho) \leq c\rho$ for every $\rho \in [0, 1)$ and every $w \in \mathcal{W}$ (the system of ancestral paths to points in the tree)*

$$|s_2 - s_1| \leq \delta_n \Rightarrow |w^{(n)}_{s_2} - w^{(n)}_{s_1}| \leq c(|s_2 - s_1|^{1/4} + n^{-1/4}).$$

Proof. Apply [20, Theorem 6] with $\alpha = 1/4$. The fact that we can take $\delta_n \geq \frac{1}{n}$ follows from the finite-range assumption on the lattice trees, which gives $|w^{(n)}_{i/n} - w^{(n)}_{(i-1)/n}| \leq Ln^{-1/2} \leq Ln^{-1/4}$, and so allows us to replace δ_n with $\delta_n \vee (1/n)$. □

The other main ingredient we use is a bound on the fourth moments of the increments of the total mass:

Proposition 3.13. *There is a $\gamma > 1$ and for any $T > 0$, there is a c_T such that for all $n \in \mathbb{N}$ and all $s_1, s_2 \in (\mathbb{Z}_+/n) \cap [0, T]$,*

$$n\mathbb{E}\left[(H^{(n)}_{s_2}(1) - H^{(n)}_{s_1}(1))^4\right] \leq c_T|s_2 - s_1|^{\gamma}.$$

The above is condition (ii) of [11, Theorem 2.2] with $k = 0$ and is verified in that reference (see [11, Theorem 3.3, Lemma 3.5, and Section 7]).

For $w \in \mathcal{D}(\mathbb{R}^d)$ and $t \geq 0$ let $w^t \in \mathcal{D}(\mathbb{R}^d)$ be defined by $w^t_s = w_{s\wedge t}$ and for $\phi \in \mathrm{Lip}_1$ let $\phi^t \in \mathrm{Lip}_1$ be defined by $\phi^t(w) = \phi(w^t)$. Define $\mathcal{T}^{(n)}_t = n^{-1/2}\mathcal{T}_{nt}$. We will use $\mathcal{T}^{(n)}_t$ as our index set for $w^{(n)}$, as in (1.2), and so depart from the notation in (3.4).

Lemma 3.14. *Let δ_n be as in Lemma 3.12, and assume that $0 \leq v \leq t_1 < t_2$ satisfy*

$$t_2 - v \leq \delta_n(\omega).$$

Then for $\phi \in \mathrm{Lip}_1$ and $i = 1, 2$,

$$|H^{(n)}_{t_i}(\phi) - H^{(n)}_{t_i}(\phi^v)| \leq c((t_2 - v)^{1/4} + n^{-1/4})(X^{(n)}_{t_1}(1) \vee X^{(n)}_{t_2}(1)).$$

Proof. Note that $H_t^{(n)}(\phi^t) = H_t^{(n)}(\phi)$ (recall (1.1), (1.2), and (1.4)), and therefore for t_i and v as above,

$$\begin{aligned}
|H_{t_i}^{(n)}(\phi) - H_{t_i}^{(n)}(\phi^v)| &\le \frac{c}{n} \sum_{x \in \mathcal{T}_{t_i}^{(n)}} \left|\phi\big(w^{(n)}(t_i, x)\big) - \phi\big((w^{(n)}(t_i, x))^v\big)\right| \\
&\le \frac{c}{n} \sum_{x \in \mathcal{T}_{t_i}^{(n)}} \|w^{(n)}(t_i, x) - (w^{(n)}(t_i, x))^v\| \\
&= \frac{c}{n} \sum_{x \in \mathcal{T}_{t_i}^{(n)}} \sup_{s \in [v, t_i]} |w_s^{(n)}(t_i, x) - w_v^{(n)}(t_i, x)| \\
&\le c(|t_i - v|^{1/4} + n^{-1/4}) X_{t_i}^{(n)}(1),
\end{aligned}$$

where we have used $t_2 - v \le \delta_n$ and Lemma 3.12 in the last line. The result follows. □

For a lattice tree T containing x (and o), let $T_{\not> x}$ denote the tree consisting of all vertices that are not descendants of x. If $x \notin T$ then let $T_{\not> x} = \varnothing$. Let $\mathcal{F}_{\not> x} = \sigma(\mathcal{T}_{\not> x})$. Let $\mathbb{T}_x$ denote the set of lattice trees containing the vertex x. If $x \in T$, let $R_x(T) \in \mathbb{T}_x$ denote the descendants of x in T together with x and all the edges joining them (if $x \notin T$, let $R_x(T) = \varnothing$), and let $R_x(T) - x \in \mathbb{T}_o$ denote the translation of $R_x(T)$ by $-x$.

Lemma 3.15. *For $x \in \mathbb{Z}^d$, for every Borel measurable $\varphi^* : \mathbb{T}_o \to \mathbb{R}_+$, and $\varphi : \{(x, R) : x \in \mathbb{Z}^d, R \in \mathbb{T}_x\} \to \mathbb{R}_+$ defined by $\varphi(x, R) = \varphi^*(R - x)$ a.s.*

$$\mathbb{E}\left[\varphi(x, R_x(\mathcal{T}))\middle|\mathcal{F}_{\not> x}\right] \mathbb{1}_{\{x \in \mathcal{T}\}} \le \rho \mathbb{E}[\varphi(o, \mathcal{T})] \mathbb{1}_{\{x \in \mathcal{T}\}}.$$

Proof. Let $\varphi^*(R) := \mathbb{1}_{\{R \in F\}}$, where $F \subset \mathbb{T}_o$. For S such that $\mathbb{P}(x \in \mathcal{T}, \mathcal{T}_{\not> x} = S) > 0$,

$$\begin{aligned}
&\mathbb{E}\left[\varphi(x, R_x(\mathcal{T}))\middle|\mathcal{T}_{\not> x} = S,\ x \in \mathcal{T}\right] \mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}\}} \\
&\quad = \sum_{n \in \mathbb{Z}_+} \frac{\mathbb{E}[\mathbb{1}_{\{R_x(\mathcal{T}) - x \in F\}} \mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}_n\}}]}{\mathbb{P}(\mathcal{T}_{\not> x} = S, x \in \mathcal{T})} \mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}\}}
\end{aligned}$$

By [20, Lemma 9.4] this is at most

$$\rho \mathbb{P}(\mathcal{T} \in F) \sum_n \frac{\mathbb{E}[\mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}_n\}}]}{\mathbb{P}(\mathcal{T}_{\not> x} = S, x \in \mathcal{T})} \mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}\}} = \rho \mathbb{P}(\mathcal{T} \in F) \mathbb{1}_{\{\mathcal{T}_{\not> x} = S\}} \mathbb{1}_{\{x \in \mathcal{T}\}}.$$

Summing over S gives

$$\mathbb{E}\left[\varphi(x, R_x(\mathcal{T}))\middle|\mathcal{F}_{\not> x}\right] \mathbb{1}_{\{x \in \mathcal{T}\}} \le \rho \mathbb{P}(\mathcal{T} \in F) \mathbb{1}_{\{x \in \mathcal{T}\}}.$$

The right-hand side is equal to $\rho \mathbb{E}[\varphi(o, \mathcal{T})] \mathbb{1}_{\{x \in \mathcal{T}\}}$ as claimed. Use linearity to get the result for simple non-negative functions, and monotone convergence to complete the proof. □

Assume $0 \le v \le t_1 < t_2$ and $\phi \in \mathrm{Lip}_1$. We want to bound

$$\begin{aligned}
|H_{t_2}^{(n)}(\phi) - H_{t_1}^{(n)}(\phi)| \le{}& |H_{t_2}^{(n)}(\phi) - H_{t_2}^{(n)}(\phi^v)| + |H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)| \\
&+ |H_{t_1}^{(n)}(\phi) - H_{t_1}^{(n)}(\phi^v)|. \qquad (3.14)
\end{aligned}$$

Lemma 3.14 will allow us to handle the first and last terms; the majority of the work will be in bounding the expected 4th power of the middle term. For fixed n, $T \in \mathbb{T}_o$ and $x \in n^{-1/2}\mathbb{Z}^d$, let $R_x^{(n)}(T^{(n)}) = n^{-1/2}R_{\sqrt{n}x}(T) \subset T^{(n)}$ denote the subtree consisting of x and its descendants. Write $\mathcal{R}_x^{(n)} = R_x^{(n)}(\mathcal{T}^{(n)})$.

$$H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v) = \frac{1}{C_0 n}\left[\sum_{z_2 \in \mathcal{T}_{t_2}^{(n)}} \phi((w^{(n)}(t_2, z_2))^v) - \sum_{z_1 \in \mathcal{T}_{t_1}^{(n)}} \phi((w^{(n)}(t_1, z_1))^v)\right].$$

Using the tree structure and $v \le t_1 < t_2$, this is equal to

$$\begin{aligned}
&\frac{1}{C_0 n}\left[\sum_{x \in \mathcal{T}_v^{(n)}} \sum_{z_2 \in \mathcal{T}_{t_2}^{(n)} \cap \mathcal{R}_x^{(n)}} \phi((w^{(n)}(t_2, z_2))^v) - \sum_{x \in \mathcal{T}_v^{(n)}} \sum_{z_1 \in \mathcal{T}_{t_1}^{(n)} \cap \mathcal{R}_x^{(n)}} \phi((w^{(n)}(t_1, z_1))^v)\right] \\
&= \frac{1}{C_0 n}\left[\sum_{x \in \mathcal{T}_v^{(n)}} \phi(w^{(n)}(v, x))\Big(\sum_{z_2 \in \mathcal{T}_{t_2}^{(n)} \cap \mathcal{R}_x^{(n)}} 1 - \sum_{z_1 \in \mathcal{T}_{t_1}^{(n)} \cap R_x(\mathcal{T})} 1\Big)\right] \\
&= \frac{1}{C_0 n}\left[\sum_{x \in \mathcal{T}_v^{(n)}} \phi(w^{(n)}(v, x))\left[|\mathcal{T}_{t_2}^{(n)} \cap \mathcal{R}_x^{(n)}| - |\mathcal{T}_{t_1}^{(n)} \cap \mathcal{R}_x^{(n)}|\right].\right]
\end{aligned}$$

If $x \in \mathbb{Z}^d/\sqrt{n}$, let $\Delta_{x,v}^{(n)} = \mathbb{1}_{\{x \in \mathcal{T}_v^{(n)}\}}(|\mathcal{T}_{t_2}^{(n)} \cap \mathcal{R}_x^{(n)}| - |\mathcal{T}_{t_1}^{(n)} \cap \mathcal{R}_x^{(n)}|)$. If $\vec{x}_m = (x_1, \dots, x_m)$, then

$$(H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v))^m = \frac{1}{C_0^m n^m} \sum_{\vec{x}_m \in (\mathcal{T}_v^{(n)})^m} \prod_{j=1}^m \phi(w^{(n)}(v, x_j)) \prod_{j'=1}^m \Delta_{x_{j'},v}^{(n)}. \quad (3.15)$$

Let $\mathbb{1}_{\{x,v\}}^{(n)} := \mathbb{1}_{\{x \in \mathcal{T}_v^{(n)}\}}$ and recall $\gamma > 1$ is as in Proposition 3.13.

Lemma 3.16. *Let $\varepsilon \in (0, 1]$, $K > 0$ and $T \in \mathbb{N}$. There is a $C_{K,T} > 0$ so that for $n \in \mathbb{N}$, $0 < p \le 4$, all $x \in \mathbb{Z}^d/\sqrt{n}$, all $t_i \in \mathbb{Z}_+/n$ such that $0 \le t_1 \le t_2 \le t_1 + 1$, and all $0 \le v \le t_1 - K(t_2 - t_1)^\varepsilon$,*

$$n\mathbb{E}\Big[|\Delta_{x,v}^{(n)}/n|^p \Big| \mathcal{F}_{\not\ni x}\Big] \le C_{K,T}|t_2 - t_1|^{\gamma p/4 - \varepsilon}\mathbb{1}_{\{x,v\}}^{(n)} \quad a.s. \quad (3.16)$$

Proof. By Lemma 3.15 with the function $\varphi^*(R) = \big||R_{n(t_2-v)}| - |R_{n(t_1-v)}|\big|^p$, the left hand side of (3.16) is at most

$$\begin{aligned}
&Cn\rho\mathbb{E}\Big[|H_{t_2-v}^{(n)}(1) - H_{t_1-v}^{(n)}(1)|^p\Big]\mathbb{1}_{\{x,v\}}^{(n)} \\
&= Cn\rho\mathbb{E}\Big[|H_{t_2-v}^{(n)}(1) - H_{t_1-v}^{(n)}(1)|^p \Big| H_{t_1-v}^{(n)}(1) > 0\Big]\mathbb{P}(H_{t_1-v}^{(n)}(1) > 0)\mathbb{1}_{\{x,v\}}^{(n)},
\end{aligned}$$

where we have used the fact that the integrand is 0 on $\{H_{t_1-v}^{(n)}(1) = 0\}$. By Jensen's inequality, this is at most

$$Cn\mathbb{P}(H_{t_1-v}^{(n)}(1) > 0)\mathbb{E}\Big[|H_{t_2-v}^{(n)}(1) - H_{t_1-v}^{(n)}(1)|^4 \Big| H_{t_1-v}^{(n)}(1) > 0\Big]^{p/4}\mathbb{1}_{\{x,v\}}^{(n)}$$

$$\le Cn\mathbb{P}(H^{(n)}_{t_1-v}(1)>0)\left(n^{-1}\mathbb{P}(H^{(n)}_{t_1-v}(1)>0)^{-1}n\mathbb{E}\Big[|H^{(n)}_{t_2-v}(1)-H^{(n)}_{t_1-v}(1)|^4\Big]\right)^{p/4}\mathbb{1}^{(n)}_{\{x,v\}}$$
$$\le C_T(n\mathbb{P}(H^{(n)}_{t_1-v}(1)>0))^{1-p/4}(|t_2-t_1|^\gamma)^{p/4}\mathbb{1}^{(n)}_{\{x,v\}},$$

where we have used Proposition 3.13, and that, without loss of generality, $t_2-t_1\ge n^{-1}$, so $[t_2-v]_n-[t_1-v]_n\le t_2-t_1+n^{-1}\le 2(t_2-t_1)$. Now use the uniform bound on the survival probability from Condition 3.4 for lattice trees, to bound the above by

$$C(t_1-v)^{p/4-1}|t_2-t_1|^{\gamma p/4}\mathbb{1}^{(n)}_{\{x,v\}},$$

Since $t_1-v\ge K(t_2-t_1)^\varepsilon$ and $|t_2-t_1|\le 1$, this is at most

$$C_{K,T}|t_2-t_1|^{\frac{p}{4}(\gamma+\varepsilon)-\varepsilon}\mathbb{1}_{\{x,v\}}\le C_{K,T}|t_2-t_1|^{\frac{p}{4}\gamma-\varepsilon}\mathbb{1}^{(n)}_{\{x,v\}},$$

as required. □

In proving our next result, we will make use of Lemma 2.13 with each $t_i=t$.

Proposition 3.17. *There are $\eta,\varepsilon\in(0,1]$, and for any $T\in\mathbb{N}$ a constant C_T, such that for all $\phi\in \mathrm{Lip}_1$, all $t_1,t_2\in[0,T]$ satisfying $(2n)^{-1}\le t_2-t_1\le 1/2$ and $v\le t_1-5(t_2-t_1)^\varepsilon$ (v may be negative), and all $n\in\mathbb{N}$,*

$$n\mathbb{E}\Big[(H^{(n)}_{t_2}(\phi^{v^+})-H^{(n)}_{t_1}(\phi^{v^+}))^4\Big]\le C_T|t_2-t_1|^{1+\eta}.$$

Proof. We first show that it suffices to prove the above for $t_i\in\mathbb{Z}_+/n$ satisfying

$$t_i\le T,\ t_1\le t_2\le t_1+1,\ \text{and any } v\le t_1-(t_2-t_1)^\varepsilon. \tag{3.17}$$

Assume this result and let n, t_i and v be as in the theorem. Using $t_2-t_1\le 1/2$, we have

$$[t_2]_n-[t_1]_n\le(t_2-t_1)+\frac{1}{n}\le\frac{1}{2}+\frac{1}{n}<1.$$

In addition, using $t_2-t_1\ge 1/(2n)$ we have

$$[t_2]_n-[t_1]_n\le(t_2-t_1)+\frac{1}{n}\le 3(t_2-t_1), \tag{3.18}$$

which implies

$$\begin{aligned}[t_1]_n-([t_2]_n-[t_1]_n)^\varepsilon&\ge t_1-\frac{1}{n}-(3(t_2-t_1))^\varepsilon\\&\ge t_1-2(t_2-t_1)-3^\varepsilon(t_2-t_1)^\varepsilon\\&\ge t_1-5(t_2-t_1)^\varepsilon\ge v.\end{aligned}$$

The above inequalities show that our hypotheses (3.17) hold for $[t_i]_n$ and the given v. Using the fact that $H^{(n)}_{t_i}=H^{(n)}_{[t_i]_n}$ we have from our assumed result, that

$$n\mathbb{E}\Big[(H^{(n)}_{t_2}(\phi^{v^+})-H^{(n)}_{t_1}(\phi^{v^+}))^4\Big]\le C_T|[t_2]_n-[t_1]_n|^{1+\eta}\le C_T3^{1+\eta}(t_2-t_1)^{1+\eta}$$

(the last by (3.18)), as required.

So consider now only $t_i \in \mathbb{Z}_+/n$ satisfying (3.17) and $t_2 > t_1$ (without loss of generality). We first assume $v \le 0$. In this case for all $x \in \mathcal{T}_{t_i}^{(n)}$, $w^{(n)}(t_i, x)^0$ is the zero path, $\bar{0}$, and so

$$H_{t_i}^{(n)}(\phi^{v^+}) = H_{t_i}^{(n)}(\phi^0) = \phi(\bar{0}) H_{t_i}^{(n)}(1) = \phi(\bar{0}) X_{t_i}^{(n)}(1).$$

The required inequality now follows (recall $t_i \in \mathbb{Z}_+/n$) from Proposition 3.13 and $|\phi(\bar{0})| \le 1$.

So assume henceforth that $0 \le v \le t_1 - (t_2 - t_1)^{\varepsilon}$. For $x \in \mathcal{T}_v^{(n)}$, and $\phi \in \mathrm{Lip}_1$, write $\phi_{x,v}^{(n)} := \phi(w^{(n)}(v, x))$. Note that from (3.15) we have

$$\begin{aligned} D_4 &:= n\mathbb{E}\left[(H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v))^4 \right] \\ &= \frac{1}{C_0^4 n^3} \mathbb{E}\Bigg[\sum_{\substack{\vec{x}=(x_1,x_2,x_3,x_4) \\ \in (\mathbb{Z}^d/\sqrt{n})^4}} \prod_{j=1}^{4} \left(\phi_{x_j,v}^{(n)} \Delta_{x_j,v}^{(n)} \right) \Bigg] \\ &= \frac{1}{C_0^4 n^3} \mathbb{E}\Bigg[\sum_{\vec{x} \in (\mathbb{Z}^d/\sqrt{n})^4} \phi_{\vec{x},v}^{(n)} \Delta_{\vec{x},v}^{(n)} \Bigg], \end{aligned} \tag{3.19}$$

where $\Delta_{\vec{x},v}^{(n)}$ denotes the product of the indicators $\Delta_{x_i,v}^{(n)}$ over the elements x_i of the vector $\vec{x}$ and $\phi_{\vec{x},v}^{(n)}$ is the product (running over the elements of the vector $\vec{x}$) of the $\phi_{x_i,v}^{(n)}$.

We'd like to condition $\Delta_{x_4,v}^{(n)}$ on $\mathcal{F}_{\not\succ x_4}$ in order to extract a positive power of $t_2 - t_1$ using Lemma 3.16. This is complicated by the fact that there are terms in the sums where other $x_i = x_4$. If we specify for which i this is true for then we will also have a constraint that the remaining x_j are not equal to x_4. After conditioning we wish to restore the possibility that these $x_j = x_4$ in order to recover a term of the form $(H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v))$ raised to some power smaller than 4 and so derive a recursive inequality which will bound the mean of fourth power of this increment. This results in an inclusion–exclusion argument below. To shorten the notation we will drop the dependence on v and n in our notation and also suppress the summation range of $\vec{x}$.

In what follows, $A_1 \subset [4]$ denotes the set of indices i for which $x_i = x_4$ (so in particular $4 \in A_1$). Then letting $A_1^c = [4] \setminus A_1$, and writing $x(A) := \{x_i : i \in A\}$ and $\vec{x}_A$ for the vector $\vec{x}$ with coordinates restricted to A, we have

$$D_4 = n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[\sum_{\vec{x}_{A_1^c}} \phi_{\vec{x}_{A_1^c}} \Delta_{\vec{x}_{A_1^c}} \sum_{x_4 \notin x(A_1^c)} \phi_{x_4}^{|A_1|} \Delta_{x_4}^{|A_1|} \Bigg],$$

where in the case $A_1 = [4]$ we interpret the term in the expectation as $\sum_{x_4} \phi_{x_4}^4 \Delta_{x_4}^4$.

Taking conditional expectation with respect to $\mathcal{F}_{\not\succ x_4}$ and using the fact that (for $x_i \ne x_4$), $\mathbb{1}_{\{x_4,v\}} \phi_{x_i} \Delta_{x_i}$ is $\mathcal{F}_{\not\succ x_4}$-measurable (as is $\mathbb{1}_{\{x_4,v\}} \phi_{x_4}$) we have that D_4 is equal to

$$n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[\sum_{\vec{x}_{A_1^c}} \phi_{\vec{x}_{A_1^c}} \Delta_{\vec{x}_{A_1^c}} \sum_{x_4 \notin x(A_1^c)} \phi_{x_4}^{|A_1|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\succ x_4}] \Bigg].$$

Interpreting the empty sum $\sum_{x_4 \in x(A_1^c)}$ as zero when $A_1^c = \varnothing$, we can write the above as

$$n^{-3}C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[\sum_{\vec{x}_{A_1^c}} \phi_{\vec{x}_{A_1^c}} \Delta_{\vec{x}_{A_1^c}} \sum_{x_4} \phi_{x_4}^{|A_1|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\ni x_4}] \Bigg] \tag{3.20}$$

$$- n^{-3}C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[\sum_{\vec{x}_{A_1^c}} \phi_{\vec{x}_{A_1^c}} \Delta_{\vec{x}_{A_1^c}} \sum_{x_4 \in x(A_1^c)} \phi_{x_4}^{|A_1|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\ni x_4}] \Bigg]. \tag{3.21}$$

Note that $|A_1| + |A_1^c| = 4$ and reason as in (3.19) to see that (3.20) equals

$$C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[(H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v))^{|A_1^c|} \sum_{x_4} \phi_{x_4}^{|A_1|} n \mathbb{E}[(\Delta_{x_4}/n)^{|A_1|} | \mathcal{F}_{\not\ni x_4}] \Bigg],$$

which, by Lemma 3.16 and $|\phi| \le 1$, is bounded in absolute value by

$$C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[|H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)|^{|A_1^c|} \sum_{x_4} \mathbb{1}_{\{x_4, v\}} \Bigg] |t_2 - t_1|^{(\gamma |A_1|/4) - \varepsilon}.$$

Expressing the sum over x_4 in terms of $H_v^{(n)}(1)$ this is equal to

$$C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} n \mathbb{E}\Bigg[|H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)|^{|A_1^c|} H_v^{(n)}(1) \Bigg] |t_2 - t_1|^{(\gamma |A_1|/4) - \varepsilon}.$$

By Hölder's inequality this is at most

$$C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} n \mathbb{E}\Big[|H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)|^4 \Big]^{|A_1^c|/4} \mathbb{E}\Big[H_v^{(n)}(1)^{4/(4-|A_1^c|)} \Big]^{(4-|A_1^c|)/4} |t_2 - t_1|^{(\gamma |A_1|/4) - \varepsilon}$$

$$\le C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} D_4^{|A_1^c|/4} |t_2 - t_1|^{\gamma |A_1|/4 - \varepsilon} \Big(n \mathbb{E}\Big[H_v^{(n)}(1)^{4/(4-|A_1^c|)} \Big] \Big)^{(4-|A_1^c|)/4}.$$

Note that for $b \le 3$ we have that for $H \ge 0$, $H^{4/(4-b)} \le H + H^4$. Since $n\mathbb{E}[(H_v^{(n)}(1))^r] < C_{r,T}$ for each $r \in \mathbb{N}$ (by Lemma 2.13), this shows that this quantity is at most (C may depend on T throughout)

$$C \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} D_4^{|A_1^c|/4} |t_2 - t_1|^{(\gamma |A_1|/4) - \varepsilon}.$$

We turn now to the quantity (3.21), and it is convenient to introduce further notation. For sets $A_i \subset [4]$, let $B_i = \cup_{j=1}^i A_j$. In particular $B_1 = A_1$. Thus (3.21) is equal to the negative of

$$n^{-3}C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \mathbb{E}\Bigg[\sum_{\vec{x}_{B_1^c}} \phi_{\vec{x}_{B_1^c}} \Delta_{\vec{x}_{B_1^c}} \sum_{x_4} \mathbb{1}_{\{x_4 \in x(B_1^c)\}} \phi_{x_4}^{|A_1|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\ni x_4}] \Bigg]. \tag{3.22}$$

Abusing notation by writing $x(A) = x$ to mean that $x_i = x$ for each $i \in A$ we can write

$$\mathbb{1}_{\{x_4 \in x(B_1^c)\}} = \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \mathbb{1}_{\{x(A_2)=x_4\}} \mathbb{1}_{\{x_4 \notin x(B_2^c)\}},$$

which is simply the statement that $x_4 \in x(B_1^c)$ if and only if the set $A_2 := \{i \in [4] \setminus B_1 : x_4 = x_i\}$ is non-empty. Thus, since $x_i = x_4$ for $i \in A_2$ in this expression, (3.22) is equal to

$$n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \mathbb{E}\Big[\sum_{\vec{x}_{B_2^c}} \phi_{\vec{x}_{B_2^c}} \Delta_{\vec{x}_{B_2^c}} \sum_{x_4} \mathbb{1}_{\{x_4 \notin x(B_2^c)\}} \phi_{x_4}^{|B_2|} \Delta_{x_4}^{|A_2|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\ni x_4}] \Big], \tag{3.23}$$

where we have also used the fact that $\phi_{\vec{x}_{B_1^c}} = \phi_{\vec{x}_{B_2^c}} \phi_{\vec{x}_{A_2}} = \phi_{\vec{x}_{B_2^c}} \phi_{x_4}^{|A_2|}$, and $|A_2| + |A_1| = |B_2|$. In the case $B_2^c = \varnothing$ the term in the expectation in (3.23) should be interpreted as $\sum_{x_4} \phi_{x_4}^4 \Delta_{x_4}^{|A_2|} \mathbb{E}[\Delta_{x_4}^{|A_1|} | \mathcal{F}_{\not\ni x_4}]$.

We can again condition on $\mathcal{F}_{\not\ni x_4}$ to see that (3.23) is equal to

$$n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \mathbb{E}\Big[\sum_{\vec{x}_{B_2^c}} \phi_{\vec{x}_{B_2^c}} \Delta_{\vec{x}_{B_2^c}} \sum_{x_4 \notin x(B_2^c)} \phi_{x_4}^{|B_2|} \prod_{i=1}^{2} \mathbb{E}[\Delta_{x_4}^{|A_i|} | \mathcal{F}_{\not\ni x_4}] \Big].$$

Using inclusion–exclusion in the sum over x_4 this can be written as

$$n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \mathbb{E}\Big[\sum_{\vec{x}_{B_2^c}} \phi_{\vec{x}_{B_2^c}} \Delta_{\vec{x}_{B_2^c}} \sum_{x_4} \phi_{x_4}^{|B_2|} \prod_{i=1}^{2} \mathbb{E}[\Delta_{x_4}^{|A_i|} | \mathcal{F}_{\not\ni x_4}] \Big] \tag{3.24}$$

$$- n^{-3} C_0^{-4} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \mathbb{E}\Big[\sum_{\vec{x}_{B_2^c}} \phi_{\vec{x}_{B_2^c}} \Delta_{\vec{x}_{B_2^c}} \sum_{x_4 \in x(B_2^c)} \phi_{x_4}^{|B_2|} \prod_{i=1}^{2} \mathbb{E}[\Delta_{x_4}^{|A_i|} | \mathcal{F}_{\not\ni x_4}] \Big], \tag{3.25}$$

where the sum over x_4 in (3.25) is interpreted as 0 when $B_2^c = \varnothing$. The quantity (3.24) is equal to [reasoning as in (3.19)]

$$\sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \frac{C}{n} \mathbb{E}\Big[(H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v))^{|B_2^c|} \sum_{x_4} \phi_{x_4}^{|B_2|} \prod_{i=1}^{2} n \mathbb{E}[(\Delta_{x_4}/n)^{|A_i|} | \mathcal{F}_{\not\ni x_4}] \Big].$$

We have also used $|B_2^c| + |A_1| + |A_2| = 4$ to get the correct powers of n. Using Lemma 3.16 again as before, we may bound the summand (in absolute value) by

$$C \mathbb{E}\Big[|H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)|^{|B_2^c|} H_v^{(n)}(1) \Big] \prod_{i=1}^{2} |t_2 - t_1|^{(\gamma |A_i|/4) - \varepsilon} \tag{3.26}$$

$$\leq n^{-1} |t_2 - t_1|^{(\gamma |B_2|/4) - 2\varepsilon} D_4^{|B_2^c|/4}, \tag{3.27}$$

where we have again used Hölder's inequality, Lemma 2.13, and $\sum_{i=1}^{2}|A_i| = |B_2|$ since A_1 and A_2 are disjoint. As in (3.23), the negative of (3.25) is equal to

$$\frac{C_0^{-4}}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\mathbb{E}\Big[\sum_{\vec{x}_{B_3^c}}\phi_{\vec{x}_{B_3^c}}\Delta_{\vec{x}_{B_3^c}}\sum_{x_4\notin x(B_3^c)}\phi_{x_4}^{|B_3|}\Delta_{x_4}^{|A_3|}\prod_{i=1}^{2}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]\Big],$$

where if $B_3^c = \varnothing$ the term in the expectation is interpreted as $\sum_{x_4}\phi_{x_4}^4\Delta_{x_4}^{|A_3|}\prod_{i=1}^{2}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]$. Conditioning again, this is equal to

$$\begin{aligned}
&\frac{C_0^{-4}}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\mathbb{E}\Big[\sum_{\vec{x}_{B_3^c}}\phi_{\vec{x}_{B_3^c}}\Delta_{\vec{x}_{B_3^c}}\sum_{x_4\notin x(B_3^c)}\phi_{x_4}^{|B_3|}\prod_{i=1}^{3}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]\Big]\\
&\quad=\frac{C_0^{-4}}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\mathbb{E}\Big[\sum_{\vec{x}_{B_3^c}}\phi_{\vec{x}_{B_3^c}}\Delta_{\vec{x}_{B_3^c}}\sum_{x_4}\phi_{x_4}^{|B_3|}\prod_{i=1}^{3}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]\Big] \qquad (3.28)\\
&\quad\quad-\frac{C_0^{-4}}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\mathbb{E}\Big[\sum_{\vec{x}_{B_3^c}}\phi_{\vec{x}_{B_3^c}}\Delta_{\vec{x}_{B_3^c}}\sum_{x_4\in x(B_3^c)}\phi_{x_4}^{|B_3|}\prod_{i=1}^{3}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]\Big]. \qquad (3.29)
\end{aligned}$$

As in (3.26) and (3.27), the term (3.28) is bounded in absolute value by

$$\begin{aligned}
&\frac{C}{n^2}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}n\mathbb{E}\Big[|H_{t_2}^{(n)}(\phi^v)-H_{t_1}^{(n)}(\phi^v)|^{|B_3^c|}H_v^{(n)}(1)\Big]\prod_{i=1}^{3}|t_2-t_1|^{\gamma|A_i|/4-\varepsilon}\\
&\le\frac{C}{n^2}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}D_4^{|B_3^c|/4}|t_2-t_1|^{\gamma|B_3|/4-3\varepsilon}.
\end{aligned}$$

Since in (3.29) B_3^c can contain at most one element, the sums over $\vec{x}_{B_3^c}$ and $x_4\in x(B_3^c)$ therein reduce to a sum over x_4 (with $\vec{x}_{B_3^c}=x_4$). After conditioning again we get that the negative of (3.29) is equal to

$$\frac{C_0^{-4}}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\sum_{\substack{A_4\subset B_3^c:\\ A_4\neq\varnothing}}\mathbb{E}\Big[\sum_{x_4}\phi_{x_4}^{|B_4|}\prod_{i=1}^{4}\mathbb{E}[\Delta_{x_4}^{|A_i|}|\mathcal{F}_{\not\succ x_4}]\Big], \qquad (3.30)$$

where we note that if this term is to be non-zero then each $|A_i|=1$, and in particular $B_4=[4]$. By Lemma 3.16 and then Lemma 2.13, (3.30) is bounded in absolute value by

$$\frac{C}{n^3}\sum_{\substack{A_1\subset[4]:\\ 4\in A_1}}\sum_{\substack{A_2\subset B_1^c:\\ A_2\neq\varnothing}}\sum_{\substack{A_3\subset B_2^c:\\ A_3\neq\varnothing}}\sum_{\substack{A_4\subset B_3^c:\\ A_4\neq\varnothing}}\mathbb{E}\Big[\sum_{x_4}\mathbb{1}_{\{x_4,v\}}\Big]\prod_{i=1}^{4}|t_2-t_1|^{\gamma|A_i|/4-\varepsilon}$$

$$\leq \frac{C}{n^3} \sum_{\substack{A_1 \subset [4]: \\ 4 \in A_1}} \sum_{\substack{A_2 \subset B_1^c: \\ A_2 \neq \varnothing}} \sum_{\substack{A_3 \subset B_2^c: \\ A_3 \neq \varnothing}} \sum_{\substack{A_4 \subset B_3^c: \\ A_4 \neq \varnothing}} |t_2 - t_1|^{\gamma - 4\varepsilon}.$$

After dropping some negative powers of n, we have shown above that

$$D_4 \leq C' \sum_{\ell=0}^{3} D_4^{\ell/4} |t_2 - t_1|^{\frac{\gamma}{4}(4-\ell) - 4\varepsilon},$$

Thus, letting $d = D_4 |t_2 - t_1|^{16\varepsilon - \gamma}$ and recalling that $|t_2 - t_1| \leq 1$, we have

$$d \leq C' \sum_{\ell=0}^{3} d^{\ell/4}.$$

Recall that D_4 is finite by Lemma 2.13, and so from the above, $d \leq C = C(C')$, and therefore $D_4 \leq C|t_2 - t_1|^{\gamma - 16\varepsilon}$. Choosing $\varepsilon < (\gamma - 1)/16$ completes the proof. □

For $v < 0$ define $\phi^v = \phi^0$ so that $\phi^v = \phi^{v^+}$.

Proof of Proposition 3.11. Let $\phi \in \mathrm{Lip}_1$ and $n_k \to \infty$. For a fixed $s_0 > 0$ we must show that $\{n_k\}$ has a subsequence $\{n_k'\}$ along which $\mathbb{P}\big(H_\cdot^{(n_k')}(\phi) \in \cdot | S^{(n_k')} > s_0\big)$ converges weakly to a continuous limit. The argument remains unchanged if we assume $n_k = k$, and to ease the notation we will assume this. So our goal is to show that

$$\{\mathbb{P}(H_\cdot^{(n)}(\phi) \in \cdot) | S^{(n)} > s_0) : n \in \mathbb{N}\} \text{ has a weakly convergent subsequence in } \mathcal{D}(\mathbb{R}_+, \mathbb{R}) \text{ to a continuous limit.} \tag{3.31}$$

For $T \in \mathbb{N}$, define

$$X_T^{(n)*}(1) = \sup_{t \leq T} X_t^{(n)}(1).$$

Now fix $T \in \mathbb{N}$ and assume

$$t_1, t_2 \in [0, T],\ 0 \leq t_2 - t_1 \leq 1 \qquad \text{and} \qquad t_1 - 5(t_2 - t_1)^\varepsilon \leq v \leq t_1, \tag{3.32}$$

where ε is as in Proposition 3.17; note that v may be negative. Recall from (3.14) that

$$|H_{t_2}^{(n)}(\phi) - H_{t_1}^{(n)}(\phi)| \leq \Big[\sum_{i=1}^{2} |H_{t_i}^{(n)}(\phi) - H_{t_i}^{(n)}(\phi^v)|\Big] + |H_{t_2}^{(n)}(\phi^v) - H_{t_1}^{(n)}(\phi^v)|. \tag{3.33}$$

Note that (3.32) implies $t_2 - v^+ \leq t_2 - v \leq (t_2 - t_1) + 5(t_2 - t_1)^\varepsilon$, and so if $\delta_n(\omega)$ is as in Lemma 3.12, then Lemma 3.14 (applied to $v^+ \geq 0$) together with the facts that $\phi^v = \phi^{v^+}$ and $t_2 - t_1 \leq 1$ show that

$$\begin{aligned} &(t_2 - t_1) + 5(t_2 - t_1)^\varepsilon \leq \delta_n \text{ implies} \\ &\sum_{i=1}^{2} |H_{t_i}^{(n)}(\phi) - H_{t_i}^{(n)}(\phi^v)| \leq C X_T^{(n)*}(1)[(t_2 - t_1)^{\varepsilon/4} + n^{-1/4}]. \end{aligned} \tag{3.34}$$

If $\eta > 0$ is as in Proposition 3.17, let $\eta_0 = \eta/8$. Proposition 3.17 shows that for $m, n \in \mathbb{N}$ satisfying

$$m \le (\log_2 n) + 1, \quad \text{that is,} \quad 2^{-m} \ge \frac{1}{2n},$$

then, by taking a union bound over $k \in \mathbb{Z}_+ : 0 \le k2^{-m} \le T+1$,

$$n\mathbb{P}\Big(\max_{0\le k\le 2^m(T+1)} \big| H^{(n)}_{(k+1)2^{-m}}(\phi^{(k2^{-m}-5\cdot 2^{-m\varepsilon})^+}) - H^{(n)}_{k2^{-m}}(\phi^{(k2^{-m}-5\cdot 2^{-m\varepsilon})^+}) \big| > 2^{-m\eta_0} \Big)$$
$$\le 2^{4m\eta_0}(T+2)2^m C_{T+2} 2^{-m(1+\eta)} = C'_T 2^{-m(\eta/2)}.$$

By a union bound there is an $M^{(n)}_0(\omega) \in \mathbb{N}^{\ge 2}$ so that

$$\text{for all } M \ge 2,\ n\mathbb{P}(M^{(n)}_0 \ge M) \le C_{T,\eta} 2^{-M\eta/2}, \tag{3.35}$$

and for all $m \in \mathbb{N}$ satisfying $M^{(n)}_0 \le m \le (\log_2 n) + 1$, we have

$$\max_{0\le k\le 2^m(T+1)} \big| H^{(n)}_{(k+1)2^{-m}}(\phi^{(k2^{-m}-5\cdot 2^{-m\varepsilon})^+}) - H^{(n)}_{k2^{-m}}(\phi^{(k2^{-m}-5\cdot 2^{-m\varepsilon})^+}) \big| \le 2^{-m\eta_0}.$$

Set $\eta_1 = (\varepsilon/4) \wedge \eta_0 > 0$. Combine the above bound with (3.34) and use it in (3.33) (with $T+1$ in place of T in the latter two) to see that for all natural numbers m satisfying

$$\frac{1}{2n} \le 2^{-m} \le 2^{-M^{(n)}_0} \text{ and } 6\cdot 2^{-m\varepsilon} \le \delta_n,$$

we have

$$\max_{0\le k\le 2^m(T+1)} |H^{(n)}_{(k+1)2^{-m}}(\phi) - H^{(n)}_{k2^{-m}}(\phi)| \le 2C {X^{(n)}_{T+2}}^*(1)(2^{-m\varepsilon/4} + n^{-1/4}) + 2^{-m\eta_0}$$
$$\le (6C {X^{(n)}_{T+2}}^*(1) + 1)2^{-m\eta_1}. \tag{3.36}$$

Set $m_n = \lfloor (\log_2 n) + 1 \rfloor$ and $T_n = \{j2^{-m_n} : j \in \mathbb{Z}_+ \cap [0, (T+1)2^{m_n}]\}$. Lévy's binary expansion argument and (3.36) shows that if

$$t_1, t_2 \in T_n, \text{ and } 0 \le t_2 - t_1 < 2^{-M^{(n)}_0} \wedge (\delta_n/6)^{1/\varepsilon},$$

then

$$|H^{(n)}_{t_2}(\phi) - H^{(n)}_{t_1}(\phi)| \le C({X^{(n)}_{T+2}}^*(1) + 1)|t_2 - t_1|^{\eta_1}. \tag{3.37}$$

Since $\frac{1}{2^{m_n}} < \frac{1}{n}$, for any $t \in [0, T]$ we may choose $\{t\}_n \in \big[[t]_n, [t]_n + \frac{1}{n}\big) \cap T_n$. Let

$$\delta'_n = \frac{1}{3}(2^{-M^{(n)}_0} \wedge (\delta_n/6)^{1/\varepsilon}).$$

Let $t_1, t_2 \in [0, T] \cap \mathbb{Z}_+/n$ be such that $0 < t_2 - t_1 \le \delta'_n$. Then $|t_i - \{t_i\}_n| < 1/n \le |t_2 - t_1|$, which implies that

$$|\{t_2\}_n - \{t_1\}_n| \le |t_2 - t_1| + \frac{2}{n} < 3|t_2 - t_1| \le 2^{-M^{(n)}_0} \wedge (\delta_n/6)^{1/\varepsilon}.$$

Thus (3.37) holds for $\{t_2\}_n, \{t_1\}_n$, that is,

$$|H^{(n)}_{\{t_2\}_n}(\phi) - H^{(n)}_{\{t_1\}_n}(\phi)| \le C({X^{(n)}_{T+2}}^*(1) + 1)|\{t_2\}_n - \{t_1\}_n|^{\eta_1}$$
$$\le C3^{\eta_1}({X^{(n)}_{T+2}}^*(1) + 1)|t_2 - t_1|^{\eta_1}.$$

Now use the fact that $H^{(n)}_{t_i}(\phi) = H^{(n)}_{\{t_i\}_n}(\phi)$ for $i = 1, 2$ to conclude that:

$$\text{for all } t_1 < t_2 \in [0, T] \cap (\mathbb{Z}_+/n) \text{ such that } t_2 - t_1 \le \delta'_n \text{ we have} \tag{3.38}$$

$$|H^{(n)}_{t_2}(\phi) - H^{(n)}_{t_1}(\phi)| \le C_T(X^{(n)}_{T+2}{}^*(1) + 1)|t_2 - t_1|^{\eta_1}. \tag{3.39}$$

Next use Lemma 3.12 and (3.35) to see that for $r \in (0, \frac{1}{12})$, then

$$\begin{aligned} n\mathbb{P}(\delta'_n \le r) &\le n\mathbb{P}(M^{(n)}_0 \ge \log_2(1/3r)) + n\mathbb{P}(\delta_n \le 6(3r)^{\varepsilon}) \\ &\le C_{T,\eta}(3r)^{\eta/2} + c6(3r)^{\varepsilon} \\ &\le C'_{T,\eta} r^{\varepsilon \wedge (\eta/2)} \le C'_{T,\eta} r^{\eta_1}. \end{aligned} \tag{3.40}$$

Our objective now follows easily from (3.39) and (3.40). Let $\{\tilde{H}^{(n)}_t,\ t \ge 0\}$ be the continuous process obtained by linearly interpolating $\{H^{(n)}_{j/n}(\phi) : j \in \mathbb{Z}_+\}$. It follows from (3.38) and (3.39), with $T + 1$ in place of T, that for some C'_T,

$$\begin{aligned} &\text{if } t_1 \le t_2 \in [0, T] \text{ and } (t_2 - t_1) \vee \frac{1}{n} \le \delta'_n, \text{ then} \\ &|\tilde{H}^{(n)}_{t_2} - \tilde{H}^{(n)}_{t_1}| \le C'_T(X^{(n)}_{T+3}{}^*(1) + 1)|t_2 - t_1|^{\eta_1}. \end{aligned} \tag{3.41}$$

For $t_2 - t_1 \ge \frac{1}{n}$ this is an easy consequence of the triangle inequality and the fact that $\delta'_n \ge \frac{1}{n}$. For $0 < t_2 - t_1 < \frac{1}{n}$, either $[t_1]_n = [t_2]_n$ and the linear interpolation and $\delta'_n \ge 1/n$ easily give the desired bound, or $[t_2]_n = [t_1]_n + 1/n$, and the triangle inequality gives

$$|\tilde{H}^{(n)}_{t_2} - \tilde{H}^{(n)}_{t_1}| \le |\tilde{H}^{(n)}_{t_2} - \tilde{H}^{(n)}_{[t_2]_n}| + |\tilde{H}^{(n)}_{[t_2]_n} - \tilde{H}^{(n)}_{t_1}|$$

which leads to the required bound using the linear interpolation and $\delta'_n \ge 1/n$ again.

Recall that $|\phi| \le 1$ implies $|\tilde{H}^{(n)}_0| \le \frac{1}{C_0 n}$. We now fix $T \in \mathbb{N}$, and for $\delta, M > 0$, define a compact set of paths in $\mathcal{C} = \mathcal{C}([0, T], \mathbb{R})$ by

$$\begin{aligned} K_{\delta,M} = \big\{ w \in \mathcal{C} : |w_0| \le C_0^{-1} \text{ and } \forall t_1, t_2 \in [0, T], \ \text{ if } |t_2 - t_1| \le \delta \text{ then} \\ |w_{t_2} - w_{t_1}| \le C'_T(M + 1)|t_2 - t_1|^{\eta_1} \big\}. \end{aligned}$$

Compactness is clear by the Arzela–Ascoli Theorem. Recall that $s_0 > 0$. It follows from (3.41) and (3.40) that for small enough $\delta_k > 0$ and large enough $M_k, n_k \in \mathbb{N}$,

$$\begin{aligned} n\mathbb{P}(\tilde{H}^{(n)}_{\cdot}|_{[0,T]} \notin K_{\delta,M}, S^{(n)} > s_0) &\le n\mathbb{P}\big(X^{(n)}_{T+3}{}^*(1) > M, S^{(n)} > s_0\big) + n\mathbb{P}\big(\delta'_n \le \frac{1}{n} \vee \delta\big) \\ &\le 2^{-k}, \quad \text{if } \delta \le \delta_k,\ M \ge M_k, \text{ and } n \ge n_k. \end{aligned} \tag{3.42}$$

Here we are using the tightness of the maximum total mass processes from [11, Theorem 1.2 and Corollary 1.3]. By further decreasing δ_k and increasing M_k we can realize the bound in (3.42) for all $n \in \mathbb{N}$. It now follows that for the compact sets $\hat{K}_m = \cap_{k=m}^{\infty} K_{\delta_k, M_k}$ we have

$$\text{for all } m, n \in \mathbb{N}, \quad n\mathbb{P}\big(\tilde{H}^{(n)}|_{[0,T]} \notin \hat{K}_m, S^{(n)} > s_0\big) \le 2^{1-m}. \tag{3.43}$$

We use the lower bound on the survival probability from (3.8):

$$\mathbb{P}(S^{(n)} > s_0) \ge \underline{c}((ns_0) \vee 1)^{-1}.$$

Combine the above with (3.43) to conclude that for all $m, n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}\big(\tilde{H}^{(n)}|_{[0,T]} \notin \hat{K}_m | S^{(n)} > s_0\big) &\le \frac{\mathbb{P}(\tilde{H}^{(n)}|_{[0,T]} \notin \hat{K}_m,\, S^{(n)} > s_0)}{\underline{c}/((ns_0) \vee 1)} \\ &\le \underline{c}^{-1}(ns_0+1)\mathbb{P}(\tilde{H}^{(n)}|_{[0,T]} \notin \hat{K}_m,\, S^{(n)} > s_0) \\ &\le \underline{c}^{-1}(s_0+1)2^{1-m}.\end{aligned}$$

This shows that $\{\mathbb{P}(\tilde{H}^{(n)} \in \cdot | S^{(n)} > s_0) : n \in \mathbb{N}\}$ is tight in $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ and so by Prohorov's theorem is relatively compact in $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$. This implies (see, e.g., [8, Proposition 10.4 in Chapter 3]) that $\{\mathbb{P}(H^{(n)}(\phi) \in \cdot | S^{(n)} > s_0) : n \in \mathbb{N}\}$ is $\mathcal{C}$-relatively compact in $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$, proving (3.31), as required. □

4. Proof of Proposition 2.11

The goal of this section is to prove Proposition 2.11. The proof is a modification of that of [17, Theorem 4.8], so we will not give all of the details here. Instead we will indicate the main ideas of the proof, and refer the reader to [17] for various details.

For $F \in \Sigma_r$, [17, Theorem 4.8] proves Proposition 2.11 in the simplified setting where $j_e = 1$ for every $e \in \mathcal{E}(F)$. In that reference (and with $j_e = 1$ for each e) the quantity $\hat{t}^{(F)}_{\check{\boldsymbol{n}}}(\cdot)$ is written as $\hat{t}_{\mathcal{N}(F,\check{\boldsymbol{n}})}(\cdot)$, where $\mathcal{N}(F, \check{\boldsymbol{n}})$ denotes a skeleton network consisting of inserting $\check{n}_{e,1} - 1$ vertices into edge e, for each $e \in \mathcal{E}(F)$. The quantity $\rho^{-1}\hat{t}_{\mathcal{N}(F,\check{\boldsymbol{n}})}$ then encodes (in Fourier space) the probability of our random tree $\mathcal{T}$ connecting the origin to r specified space–time points with the spatial and temporal locations of the branch points, as well as the "shape" of the connections also specified (consider the set $\check{\boldsymbol{T}}(F, \check{\boldsymbol{y}}, \check{\boldsymbol{n}})$ in the case where each $j_e = 1$). In our paper j_e need not be equal to 1. In this more general setting, $\hat{t}^{(F)}_{\check{\boldsymbol{n}}}(\cdot)$ encodes (in Fourier space) the probability of a subset of the above event, where now the spatial locations at various other fixed times are also specified. The appropriate skeleton network is now a *marked skeleton network* $\mathcal{N}^+$ (see below), where certain vertices on the skeleton network $\mathcal{N}$ at fixed times (graph distance from the root) are marked.

The approach in [17, proof of Theorem 4.8] relies on the so-called *lace expansion* and involves an inductive argument (on r). To be more precise [17] uses the lace expansion on a tree network (introduced in [15] for networks of self-avoiding walks) in the context of lattice trees, with the expansion applied at the closest branch point to the root in the network $\mathcal{N}$. The expansion gives rise to certain diagrams that involve lattice trees connecting or intersecting in various ways. Some of these connections are of fixed temporal length, and others are of unrestricted length. A crucial part of the analysis involves bounding these diagrams. The bounds depend on the complexity of the diagram, as well as the total temporal length in the diagram. Diagrams where either the complexity or the length is large give small contributions (recall that we are in high dimensions), as they are asking for either lots of intersections, or for intersections to occur over a large distance.

The point of this discussion is that, in our setting, when j_e need not be 1, one can perform exactly the same expansion. It turns out that there are essentially no new diagrams to deal with in our setting. Below we introduce the definition of a marked skeleton network (see also Fig. 10) and then proceed in the following subsections to expand the above outline of the proof of Proposition 2.11.

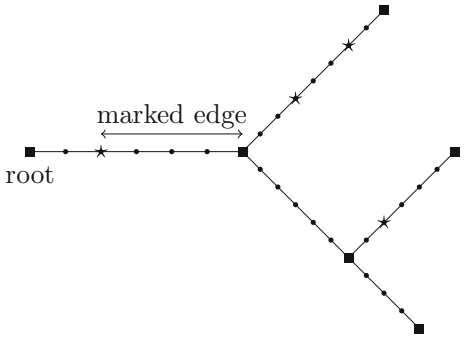

Fig. 10. An example of a marked skeleton network from a shape $F \in \Sigma_4$. Branch points and leaves are ■, marked points are ⋆

Definition 4.1. Given $F \in \Sigma_r$ and $\check{\boldsymbol{n}} = (\check{n}_{e,i})_{i\in[j_e],e\in[2r-1]}$ where $j_e \in \mathbb{N}$ for each $e \in \mathcal{E}(F)$, define $\mathcal{N}^+(F, \check{\boldsymbol{n}})$ to be the *marked skeleton network* which is obtained from F by

- inserting $j_e - 1$ *marked points* into edge e of F for each $e \in [2r - 1]$, thus each edge e in F becomes a path of j_e edges, called *marked edges*, which are labelled as (e, i) for $i \le j_e$; and
- inserting $\check{n}_{e,i} - 1$ vertices into every marked edge (e, i), so $\check{n}_{e,i} \in \mathbb{N}$ denotes the length of the marked edge (e, i).

Write $\boldsymbol{E}(\mathcal{N}^+)$ for the set of marked edges of $\mathcal{N}^+$. Marked edges are *adjacent* if they share a vertex in common.

The *branch* $\mathcal{N}^+_e$ of $\mathcal{N}^+$ associated to an edge e of F is the set of vertices of $\mathcal{N}^+$ consisting of the endpoints of e together with all points (marked or not) inserted into that edge as per the definition of $\mathcal{N}^+$. The set of branches is written $B(\mathcal{N}^+) := (\mathcal{N}^+_e)_{e\in\mathcal{E}(F)}$. Two distinct branches $\mathcal{N}^+_e$ and $\mathcal{N}^+_{e'}$ are *adjacent* if and only if they have a vertex in common (equivalent to e and e' being adjacent in F).
A *special point* of $\mathcal{N}^+$ is any marked point, branch point or leaf. ◀

Remark 4.2. The sets of (all) vertices and edges of a marked skeleton network $\mathcal{N}^+$ will be denoted by $\mathcal{N}^+$ and $E(\mathcal{N}^+)$ respectively (note the abuse of notation that $\mathcal{N}^+$ denotes both the marked skeleton and its set of vertices). The cardinality of $E(\mathcal{N}^+)$ is $\#E(\mathcal{N}^+) = \sum_{e\in F}\sum_{i=1}^{j_e} \check{n}_{e,i}$ and the number of vertices is 1 larger. All special points are also vertices of $\mathcal{N}^+$, while marked edges should be considered as distinct objects from edges, even for marked edges (e, i) such that $\check{n}_{e,i} = 1$ (note that we have thus far specified a labelling scheme for marked edges, but not edges). The set of marked edges of $\mathcal{N}^+$ is $\boldsymbol{E}(\mathcal{N}^+)$. ★

4.1. Asymptotics of the detailed 1-particle transform. For the case where $r = 1$, there exists only one shape in Σ_1 which consists of a single edge e. In this case, we use the notation $[\check{n}_1, \dots, \check{n}_\ell]$, for $(\check{n}_i)_{i\le\ell} \in \mathbb{N}^\ell$ with $\ell \ge 0$ to designate the corresponding marked skeleton network (containing no branch point) with $\check{\boldsymbol{n}} = \{\check{n}_{e,1}, \dots, \check{n}_{e,\ell}\}$.

One of the main results of [18] (see Theorem 4.3(ii) of that reference) can be reformulated as the following proposition (the error terms are not stated explicitly in [18,

Theorem 4.3], but if we keep track of them we get the following result), which is the $r = 1$ case of Proposition 2.11:

Proposition 4.3. *Fix $d > 8$. There exists $L_0(d) \gg 1$ such that for all $L \geq L_0$:*
For each $\delta \in (0, 1 \wedge \frac{d-8}{2})$, $R > 0$, every $\ell \in \mathbb{N}$ and $(\check{n}_i)_{i\leq \ell} \in \mathbb{N}^\ell$ and for any $\check{\boldsymbol{k}} \in [-R, R]^\ell$ we have, for the unique shape $\digamma \in \Sigma_1$,

$$\hat{t}^{(\digamma)}_{\check{\boldsymbol{n}}}\left(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\right) = C_A \prod_{i=1}^{\ell} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_i^2}{2}(\frac{\check{n}_i}{n})} + \mathcal{O}\left(\sum_{i=1}^{\ell} \frac{1}{\check{n}_i^{\frac{d-8}{2}}}\right) + \mathcal{O}\left(|\check{\boldsymbol{k}}|^2 \sum_{i=1}^{\ell} \frac{\check{n}_i^{1-\delta}}{n}\right),$$

where the error depends on R, δ, L, d, ℓ, and any lower bound on $\min_{i\leq \ell} \check{n}_i/n$ and upper bound on $\max_{i\leq \ell} \check{n}_i/n$.

Note that in [18] each $\check{n}_i$ is of the form $\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor$, where $0 = t_0 < t_1 < \cdots < t_\ell \leq t^*$ and where the error term depends on $\min\{t_i - t_{i-1}\}$ and t^*.

4.2. Lace expansion. We will use the lace expansion (and induction on r) to reduce our required estimates on a shape in Σ_r with $r \geq 2$ to the shape in Σ_1. In the following we let $\mathcal{N}^+ = \mathcal{N}^+(\digamma, \check{\boldsymbol{n}})$ for some $\digamma \in \Sigma_r$ and some $\check{\boldsymbol{n}}$, where $r \geq 2$. Since each $\check{n}_{e,i}/n \geq \varepsilon$ in Proposition 2.11, for fixed ε we may assume that n is sufficiently large so that each $\check{n}_{e,i} \geq 2$ in what follows.

Definition 4.4. If $\mathcal{N}^+$ is a marked skeleton network, we say that $\mathcal{M}^+$ is a *marked subnetwork* of $\mathcal{N}^+$ and write $\mathcal{M}^+ \subset \mathcal{N}^+$ if

- as a graph, $\mathcal{M}^+$ is a (connected) subgraph of $\mathcal{N}^+$, and
- the marked points of $\mathcal{M}^+$ are those vertices in $\mathcal{M}^+$ that were marked points in $\mathcal{N}^+$ (i.e. marked points are inherited from $\mathcal{N}^+$).

As usual we also write $\mathcal{M}^+$ for the set of vertices of the marked subnetwork $\mathcal{M}^+$. ◀

Definition 4.5. Let $\mathcal{M}^+$ be a marked subnetwork of some marked skeleton network $\mathcal{N}^+$.

1. A *bond* vv' is a pair of distinct vertices v, v' of $\mathcal{M}^+$. The set of vertices in the unique path in $\mathcal{M}^+$ from v to v' is written $[v, v']$. We say that the bond vv' covers the vertices in $[v, v']$ (and the edges therein). We write $vv' \in \mathcal{M}^+$ to mean that vv' is a bond in $\mathcal{M}^+$.
2. A *graph* on $\mathcal{M}^+$ is a set of bonds and we denote the set of graphs on $\mathcal{M}^+$ by $\mathcal{G}_{\mathcal{M}^+}$.
3. Let $\mathcal{R}_{\mathcal{M}^+}$ denote the set of bonds in $\mathcal{M}^+$ that cover 2 or more special points. Furthermore set $\mathcal{G}^{-\mathcal{R}}_{\mathcal{M}^+} = \{\Gamma \in \mathcal{G}_{\mathcal{M}^+},\ \Gamma \cap \mathcal{R}_{\mathcal{M}^+} = \emptyset\}$, i.e. the graphs on $\mathcal{M}^+$ that do not contain any bond in $\mathcal{R}_{\mathcal{M}^+}$.
4. A graph $\Gamma \in \mathcal{G}_{\mathcal{M}^+}$ is said to be *connected* on $\mathcal{M}^+$ if every edge of $\mathcal{M}^+$ is covered by some $st \in \Gamma$. Let $\mathcal{G}^{\mathrm{con}}_{\mathcal{M}^+}$ be the set of connected graphs on $\mathcal{M}^+$, and $\mathcal{G}^{-\mathcal{R},\mathrm{con}}_{\mathcal{M}^+} = \mathcal{G}^{\mathrm{con}}_{\mathcal{M}^+} \cap \mathcal{G}^{-\mathcal{R}}_{\mathcal{M}^+}$.
5. Given $\Gamma \in \mathcal{G}_{\mathcal{M}^+}$ and $\mathcal{A} \subset \mathcal{M}^+$, we define $\Gamma|_{\mathcal{A}} = \{vv' \in \Gamma,\ v, v' \in \mathcal{A}\}$.
6. For a vertex $v \in \mathcal{M}^+$ and $\Gamma \in \mathcal{G}_{\mathcal{M}^+}$, we let $\mathcal{A}_v(\Gamma)$ be the largest connected subnetwork $\mathcal{A}$ of $\mathcal{M}^+$ containing v and such that $\Gamma|_{\mathcal{A}}$ is a connected graph on $\mathcal{A}$. In words, this is the connected component of covered (by Γ) vertices containing v. By convention we take $\mathcal{A}_v(\Gamma) = \{v\}$ if no bond in Γ covers v .

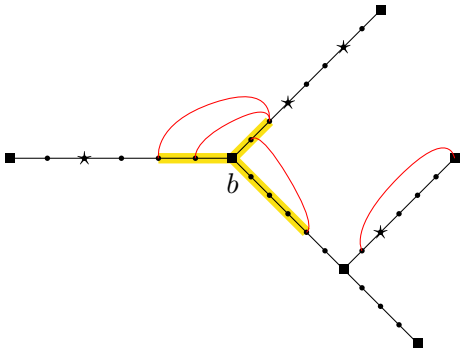

Fig. 11. A graph Γ on a marked skeleton network $\mathcal{N}^+$, with b denoting the branch point nearest to the root. The rightmost bond is in $\mathcal{R}$ since it covers two special points. Also, $\Gamma \in \mathcal{E}^b_{\mathcal{N}^+}$ since $\mathcal{A}_b(\Gamma)$ (highlighted) contains a neighbour of a marked point

7. If $v \in \mathcal{N}^+$, we let $\mathcal{E}^v_{\mathcal{N}^+}$ denote the set of graphs $\Gamma \in \mathcal{G}_{\mathcal{N}^+}$ such that $\mathcal{A}_v(\Gamma)$ contains a vertex adjacent to some special point $u \neq v$ of $\mathcal{N}^+$, and $\mathcal{E}^{-\mathcal{R},v}_{\mathcal{N}^+_+} = \mathcal{G}^{-\mathcal{R}}_{\mathcal{N}^+} \cap \mathcal{E}^v_{\mathcal{N}^+}$. See e.g. Fig. 11. ◀

In this section, for a bond $vv' \in \mathcal{N}^+$, $U_{vv'}$ will denote a quantity in $\{-1, 0\}$. Observe that (with $\mathcal{R} = \mathcal{R}_{\mathcal{N}^+}$),

$$\prod_{vv' \in \mathcal{N}^+} [1 + U_{vv'}] = \prod_{vv' \in \mathcal{N}^+ \setminus \mathcal{R}} [1 + U_{vv'}] - \Big(\prod_{vv' \in \mathcal{N}^+ \setminus \mathcal{R}} [1 + U_{vv'}] \Big) \Big(1 - \prod_{vv' \in \mathcal{R}} [1 + U_{vv'}] \Big). \tag{4.1}$$

Definition 4.6. For $\vec{m} \in \mathbb{Z}^3_+$ we write $\mathcal{S}_{\vec{m}}$ for the (unmarked) network consisting of paths of lengths $(m_j)_{j=1}^3$ respectively meeting at a common vertex. If exactly i of the m_j are strictly positive then this is a *star-shaped network* of degree i. The case $i = 0$ is a single vertex. The *central point* of $\mathcal{S}_{\vec{m}}$ is the common vertex of the 3 paths. ◀

Definition 4.7. For a marked skeleton network $\mathcal{N}^+ = \mathcal{N}^+(F, \check{\boldsymbol{n}})$ with $F \in \Sigma_r$ for some $r \geq 2$, let b denote the branch point lying on the same branch as the root. Let $\mathcal{S}^-_{\mathcal{N}^+}$ be the largest subnetwork of $\mathcal{N}^+$ containing b and which does not contain a neighbour of any other special point of $\mathcal{N}^+$ ◀

Remark 4.8. If $\Gamma \in \mathcal{G}^{-\mathcal{R}}_{\mathcal{N}^+} \setminus \mathcal{E}^b_{\mathcal{N}_+}$ then $\mathcal{A}_b(\Gamma)$ is a (connected subnetwork of a) star-shaped network of degree at most 3 (since $F \in \Sigma_r$ with $r \geq 2$). ★

Definition 4.9. If $\mathcal{N}^+$ is a marked skeleton network and $\mathcal{A} \subset \mathcal{S}^-_{\mathcal{N}^+}$ with $b \in \mathcal{A}$, then the vertex set $\mathcal{N}^+ \setminus \mathcal{A}$ (with the edge structure and marked points induced from $\mathcal{N}^+$) consists of exactly three marked skeleton networks (each of which is connected) that we write as $(\mathcal{N}^+ \setminus \mathcal{A})_i$ for $i = 1, 2, 3$. Those three subnetworks together contain all special points of $\mathcal{N}^+$ except b. ◀

For a subnetwork $\mathcal{A} \subset \mathcal{N}^+$, let $\mathcal{K}(\mathcal{A}) = \sum_{\Gamma \in \mathcal{G}^{-\mathcal{R}}_{\mathcal{A}}} \prod_{st \in \Gamma} U_{st}$. Then we can write

$$\mathcal{K}(\mathcal{N}^+) := \sum_{\Gamma \in \mathcal{G}^{-\mathcal{R}}_{\mathcal{N}^+}} \prod_{st \in \Gamma} U_{st}$$

$$= \sum_{\substack{\mathcal{A} \subset \mathcal{S}^-_{\mathcal{N}^+} \\ b \in \mathcal{A}}} \sum_{\Gamma \in \mathcal{G}^{-\mathcal{R},\mathrm{con}}_{\mathcal{A}}} \prod_{st \in \Gamma} U_{st} \prod_{i=1}^{3} \Big[\sum_{\Gamma_i \in \mathcal{G}^{-\mathcal{R}}_{(\mathcal{N}^+ \setminus \mathcal{A})_i}} \prod_{v_i v_i' \in \Gamma_i} U_{v_i v_i'} \Big] + \sum_{\Gamma \in \mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+}} \prod_{st \in \Gamma} U_{st}, \tag{4.2}$$

with the convention that $\sum_{\Gamma_i \in \mathcal{G}^{-\mathcal{R}}_{\emptyset}} \sum_{v_i v_i' \in \Gamma_i} U_{v_i v_i'} = 1$. In words this decomposition says that the set of graphs Γ on $\mathcal{N}^+$ containing no bonds that cover two or more special points consists of (i) those graphs Γ for which the induced connected subnetwork containing b also contains a neighbour of some other special point (this is the last term in (4.2)) and (ii) those graphs Γ for which this induced subnetwork does not contain the neighbour of another special point. For (ii) the induced connected subnetwork is some set $\mathcal{A}$ contained in $\mathcal{S}^-_{\mathcal{N}^+}$ so we can first sum over the possibilities for $\mathcal{A}$ and then sum over connected graphs on $\mathcal{A}$ and graphs on each $(\mathcal{N}^+ \setminus \mathcal{A})_i$. Introducing

$$\mathcal{J}(\mathcal{A}) = \sum_{\Gamma \in \mathcal{G}^{-\mathcal{R},\mathrm{con}}_{\mathcal{A}}} \prod_{st \in \Gamma} U_{st},$$

then (4.2) becomes

$$\mathcal{K}(\mathcal{N}^+) = \sum_{\substack{\mathcal{A} \subset \mathcal{S}^-_{\mathcal{N}^+}: \\ b \in \mathcal{A}}} \mathcal{J}(\mathcal{A}) \prod_{i=1}^{3} \mathcal{K}\big((\mathcal{N}^+ \setminus \mathcal{A})_i\big) + \sum_{\Gamma \in \mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+}} \prod_{st \in \Gamma} U_{st}. \tag{4.3}$$

4.3. Application of the Lace expansion. Given $\mathcal{N}^+ = \mathcal{N}^+(F, \check{\boldsymbol{n}})$ for some $F \in \Sigma_r$ ($r \geq 1$), and $\check{\boldsymbol{n}} = (\check{n}_{e,i})_{i \in [j_e], e \in [2r-1]} \in \mathbb{N}^{\boldsymbol{E}(\mathcal{N}^+)}$, and given $\check{\boldsymbol{y}} = (\check{y}_{e,i})_{i \in [j_e], e \in [2r-1]} \in (\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+)}$, define $t_{\mathcal{N}^+}(\check{\boldsymbol{y}}) = t^{(F)}_{\check{\boldsymbol{n}}}(\check{\boldsymbol{y}})$. This notation will help us deal with various subnetworks. Recalling (2.22), we have

$$t_{\mathcal{N}^+}(\check{\boldsymbol{y}}) = \sum_{T \in \check{\boldsymbol{T}}(\check{\boldsymbol{n}}, \check{\boldsymbol{y}})} W(T).$$

Definition 4.10. Given $\mathcal{N}^+$ and $\check{\boldsymbol{y}}$ as above, we define $\Omega_{\mathcal{N}^+}(\check{\boldsymbol{y}})$ to be the set of embeddings $\omega = (\omega(s))_{s \in \mathcal{N}^+}$ of $\mathcal{N}^+$ into $\mathbb{Z}^d$ such that

1. the root is mapped to 0,
2. adjacent vertices in $\mathcal{N}^+$ are mapped to points in $\mathbb{Z}^d$ at (ℓ_∞) distance at most L from each other.
3. the endpoint of the marked edge (e, j) that is farthest from the root (this endpoint is necessarily a special point) is mapped to $\sum_{f \prec e} \sum_{k \leq j_f} \check{y}_{f,k} + \sum_{k \leq j} \check{y}_{e,k}$ for all $e \in [2r-1]$ and $j \leq j_e$. ◀

For a collection of lattice trees $(R_s)_{s \in \mathcal{N}^+}$ and for a bond (pair of distinct vertices) st of $\mathcal{N}^+$ define

$$U_{st} = \begin{cases} 0, & \text{if } R_s \cap R_t = \emptyset, \\ -1, & \text{otherwise.} \end{cases} \tag{4.4}$$

Definition 4.11. Given $x \in \mathbb{Z}^d$, we write $\sum_{R \ni x}$ to denote a sum over lattice trees R containing the point $x \in \mathbb{Z}^d$. ◀

As for [17, Eq. (4.17)] we can write

$$t_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}: \\ R_s \ni \omega(s) \forall s \in \mathcal{N}^+}} \prod_{t \in \mathcal{N}^+} W(R_t) \prod_{uu' \in \mathcal{N}^+} [1 + U_{uu'}], \tag{4.5}$$

as any combination $(\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}}), (R_s)_{s\in\mathcal{N}^+})$ such that the R_s are mutually avoiding lattice trees, uniquely defines a lattice tree $T \in \check{\boldsymbol{T}}(\digamma, \check{\boldsymbol{n}}, \check{\mathbf{y}})$ and vice versa. Here, R_s is the tree hanging off the vertex $s \in \mathcal{N}^+$. Note that in the shorthand notation of [17] (4.5) would be written as

$$t_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \prod_{s \in \mathcal{N}^+} \sum_{R_s \ni \omega(s)} W(R_s) \prod_{uu' \in \mathcal{N}^+} [1 + U_{uu'}]. \tag{4.5'}$$

Recalling Definition 4.5 and (4.1), we set

$$\phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}: \\ R_s \ni \omega(s) \forall s \in \mathcal{N}^+}} \prod_{t \in \mathcal{N}^+} W(R_t) \Big(\prod_{uu' \in \mathcal{R}^c} [1 + U_{uu'}] \Big) \Big(1 - \prod_{vv' \in \mathcal{R}} [1 + U_{vv'}] \Big), \tag{4.6}$$

which is 0 unless $U_{vv'} = -1$ for some $vv' \in \mathcal{R}$, and (recalling the last term in (4.3))

$$\phi^{b}_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}: \\ R_s \ni \omega(s) \forall s \in \mathcal{N}^+}} \prod_{t \in \mathcal{N}^+} W(R_t) \sum_{\Gamma \in \mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+}} \prod_{vv' \in \Gamma} U_{vv'}.$$

By (4.1) we have

$$t_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}: \\ R_s \ni \omega(s) \forall s \in \mathcal{N}^+}} \prod_{t \in \mathcal{N}^+} W(R_t) \mathcal{K}(\mathcal{N}^+) - \phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}})$$

and by (4.3)

$$\begin{aligned} & t_{\mathcal{N}^+}(\check{\mathbf{y}}) \\ &= \sum_{\omega \in \Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}: \\ R_s \ni \omega(s) \forall s \in \mathcal{N}^+}} \prod_{t \in \mathcal{N}^+} W(R_t) \sum_{\substack{\mathcal{A} \subset \mathcal{S}^-_{\mathcal{N}^+}, \\ b \in \mathcal{A}}} \mathcal{J}(\mathcal{A}) \prod_{i=1}^{3} \mathcal{K}((\mathcal{N}^+ \setminus \mathcal{A})_i) \\ &\quad + \phi^{b}_{\mathcal{N}^+}(\check{\mathbf{y}}) - \phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}}). \end{aligned} \tag{4.7}$$

This decomposition is related to Fig. 11 where, loosely speaking, the term in $\mathcal{J}$ corresponds the interactions induced by bonds around the first branch point and the three terms in $\mathcal{K}$ correspond to three new smaller networks. Some notation associated to this decomposition is introduced in the next definition.

Definition 4.12. For a marked skeleton network $\mathcal{N}^+$, let $\check{e}_1, \check{e}_2, \check{e}_3$ be the three marked edges incident to the branch point b. Note that each for $k = 1, 2, 3$, $\check{e}_k = (e_k, i_k)$ for some $e_k \in [2r-1]$ and some $i_k \in \{1, j_{e_k}\}$ (this marked edge is necessarily the last marked edge on the branch containing the origin and the first marked edge on the other two branches containing b).

Given $\vec{m} = (m_k)_{k=1}^3$ such that $0 \le m_k \le \check{n}_{\check{e}_k} - 2$ where $\check{n}_{\check{e}_k} := \check{n}_{e_k, i_k}$, $k = 1, 2, 3$, we define $(\mathcal{N}^+_{k,\vec{m}})_{k=1,2,3}$ as the three components of $\mathcal{N}^+ \setminus \mathcal{S}_{\vec{m}}$ as in Definition 4.9 (recall Definition 4.6, and note that each $\mathcal{N}^+_{k,\vec{m}}$ is itself a marked skeleton network). Since each $m_k < \check{n}_{e_k,i_k}$, there is a bijection between marked edges of $\mathcal{N}^+$ and the marked edges of $(\mathcal{N}^+_{k,\vec{m}})_{k=1,2,3}$. The marked edge $\check{e}_k$ is split between $\mathcal{S}_{\vec{m}}$ and $\mathcal{N}^+_{k,\vec{m}}$, but we will abuse notation by retaining this label to refer to the corresponding truncated edge in both components.

Set $\boldsymbol{E}^*(\mathcal{N}^+_{k,\vec{m}}) = \boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}}) \setminus \{\check{e}_k\}$. For $k = 1, 2, 3$, write $\check{\boldsymbol{n}}^{\vec{m},k} \in \mathbb{N}^{\boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}})}$ for the vector whose components encode the lengths of marked edges in $\mathcal{N}^+_{k,\vec{m}}$, i.e.

$$\check{n}^{\vec{m},k}_{\check{e}} = \begin{cases} \check{n}_{\check{e}} & \text{if } \check{e} \in \boldsymbol{E}^*(\mathcal{N}^+_{k,\vec{m}}), \\ \check{n}_{\check{e}} - (m_k + 1) & \text{if } \check{e} = \check{e}_k. \end{cases}$$

Similarly for $\check{\boldsymbol{y}} \in (\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+)}$, $k \in \{1, 2, 3\}$ and $v_k \in \mathbb{Z}^d$ we write $\check{\boldsymbol{y}}^{v_k,k} \in (\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}})}$ for the vector whose components are

$$\check{y}^{v_k,k}_{\check{e}} = \begin{cases} \check{y}_{\check{e}} & \text{if } \check{e} \in \boldsymbol{E}^*(\mathcal{N}^+_{k,\vec{m}}), \\ \check{y}_{\check{e}} - v_k & \text{if } \check{e} = \check{e}_k. \end{cases}$$

◀

Let $\check{\boldsymbol{n}}^b := (\check{n}_{\check{e}_1}, \check{n}_{\check{e}_2}, \check{n}_{\check{e}_3})$ be the lengths of the marked edges adjacent to b in $\mathcal{N}^+$. Define

$$\mathcal{H}_{\check{\boldsymbol{n}}^b} = \Big\{\vec{m} \ : \ 0 \le m_k \le \frac{\check{n}_{\check{e}_k}}{3} \wedge (\check{n}_{\check{e}_k} - 2), \ k = 1, 2, 3\Big\}$$
$$\overline{\mathcal{H}}_{\check{\boldsymbol{n}}^b} = \{\vec{m} \ : \ 0 \le m_k \le \check{n}_{\check{e}_k} - 2, \ k = 1, 2, 3\} \setminus \mathcal{H}_{\check{\boldsymbol{n}}^b}.$$

Remark 4.13. For $\vec{m} \in \mathcal{H}_{\check{\boldsymbol{n}}^b}$, we know that for $k \in \{1, 2, 3\}$ and $\check{e} \in \boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}})$ we have $\frac{2}{3}\check{n}_{\check{e}} \le \check{n}^{\vec{m},k}_{\check{e}} \le \check{n}_{\check{e}}$ (with $\check{n}^{\vec{m},k}_{\check{e}} = \check{n}$ whenever $\check{e} \in \boldsymbol{E}^*(\mathcal{N}^+_{k,\vec{m}})$). In particular, recalling that there is a bijection between marked edges of $\mathcal{N}^+$ and the marked edges of $(\mathcal{N}^+_{k,\vec{m}})_{k=1,2,3}$, we can see that for any $a \in \mathbb{R}$ there exist $c(a), C(a) > 0$ such that for $\vec{m} \in \mathcal{H}_{\check{\boldsymbol{n}}^b}$

$$c(a) \sum_{k=1}^{3} \sum_{\check{e} \in \boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}})} (\check{n}^{\vec{m},k}_{\check{e}})^a \le \sum_{\check{e} \in \boldsymbol{E}(\mathcal{N}^+)} \check{n}^a_{\check{e}} \le C(a) \sum_{k=1}^{3} \sum_{\check{e} \in \boldsymbol{E}(\mathcal{N}^+_{k,\vec{m}})} (\check{n}^{\vec{m},k}_{\check{e}})^a.$$

★

Finally, we set

$$\phi^{\pi}_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega\in\Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}:\\ R_s\ni\omega(s)\forall s\in\mathcal{N}^+}} \prod_{t\in\mathcal{N}^+} W(R_t) \sum_{\vec{m}\in\overline{\mathcal{H}}_{\vec{n}^b}} \mathcal{J}(\mathcal{S}_{\vec{m}}) \prod_{i=1}^{3} \mathcal{K}\big(\mathcal{N}^+_{i,\vec{m}}\big),$$

and (noting the change to the sum over $\vec{m}$)

$$Q_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega\in\Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}:\\ R_s\ni\omega(s)\forall s\in\mathcal{N}^+}} \prod_{t\in\mathcal{N}^+} W(R_t) \sum_{\vec{m}\in\mathcal{H}_{\vec{n}^b}} \mathcal{J}(\mathcal{S}_{\vec{m}}) \prod_{i=1}^{3} \mathcal{K}\big(\mathcal{N}^+_{i,\vec{m}}\big). \tag{4.8}$$

From the argument above and (4.7), we can see that

$$t_{\mathcal{N}^+}(\check{\mathbf{y}}) = Q_{\mathcal{N}^+}(\check{\mathbf{y}}) + \phi^{\pi}_{\mathcal{N}^+}(\check{\mathbf{y}}) + \phi^{b}_{\mathcal{N}^+}(\check{\mathbf{y}}) - \phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}}). \tag{4.9}$$

The last three terms are error terms. The relevant estimates (bounds) are given in the following lemma, whose proof (which is very similar to the corresponding error bounds in [17]) will be presented in Sect. 4.6.

Lemma 4.14. *Fix $d > 8$. There exists $L_0(d)$ such that for all $L \geq L_0$ and for a marked skeleton network $\mathcal{N}^+$,*

$$\sum_{\check{\mathbf{y}}} \left|\phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}})\right| = \mathcal{O}\Big(\sum_{e=1}^{2r-1}\sum_{j=1}^{j_e} \frac{1}{\check{n}_{e,j}^{\frac{d-8}{2}}}\Big), \tag{e:$\mathcal{R}$}$$

$$\sum_{\check{\mathbf{y}}} \left|\phi^{\pi}_{\mathcal{N}^+}(\check{\mathbf{y}})\right| = \mathcal{O}\Big(\sum_{e=1}^{2r-1}\sum_{j=1}^{j_e} \frac{1}{\check{n}_{e,j}^{\frac{d-8}{2}}}\Big), \tag{e:π}$$

$$\sum_{\check{\mathbf{y}}} \left|\phi^{b}_{\mathcal{N}^+}(\check{\mathbf{y}})\right| = \mathcal{O}\Big(\sum_{e=1}^{2r-1}\sum_{j=1}^{j_e} \frac{1}{\check{n}_{e,j}^{\frac{d-8}{2}}}\Big), \tag{e:b}$$

where the constants in the $\mathcal{O}$ notation depend on d and the number of special points in $\mathcal{N}^+$.

We end this section by introducing an important quantity that will appear in the decomposition of $Q_{\mathcal{N}^+}$ and describes the interactions induced by the term $\mathcal{J}(\mathcal{S}_{\vec{m}})$ in (4.8).

Definition 4.15. For $\vec{m} \in \mathbb{Z}^3_+$ and $\vec{u} \in (\mathbb{Z}^d)^3$ we define $\pi_{\vec{0}}(\vec{u}) = \rho \mathbb{1}_{\{\vec{u}=0\}}$ and if some $m_i > 0$,

$$\pi_{\vec{m}}(\vec{u}) = \sum_{\omega\in\Omega_{\mathcal{S}_{\vec{m}}(\vec{u})}} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{S}_{\vec{m}}}:\\ R_s\ni\omega(s)\forall s\in\mathcal{S}_{\vec{m}}}} \prod_{t\in\mathcal{S}_{\vec{m}}} W(R_t)\mathcal{J}(\mathcal{S}_{\vec{m}}),$$

where the set of embeddings $\Omega_{\mathcal{S}_{\vec{m}}(\vec{u})}$ is defined similarly to Definition 4.10: the root of $\mathcal{S}_{\vec{m}}$ (which is the vertex along branch 1 at graph distance m_1 from the central vertex - if $m_1 = 0$ this is simply the central vertex itself) is mapped to 0; adjacent vertices in $\mathcal{S}_{\vec{m}}$ are mapped to points in $\mathbb{Z}^d$ at distance at most L; and the central point is mapped to u_1 and the leaves on branches i for $i = 2, 3$ are mapped to $u_1 + u_2$ and $u_1 + u_3$ respectively. ◀

Remark 4.16. This definition of $\pi_{\vec{m}}$ is exactly the same as the one in [17] (Definition 4.12) and as such the results on this quantity, that rely heavily on diagrammatic estimates, can be transferred directly to our context. ★

Recall from (2.21) and the discussion thereafter that $C_V = \rho^2 V$. The constant V was defined in [17, (4.30)] as

$$V = z_D^3 \sum_{\vec{m}\in\mathbb{Z}_+^3} \sum_{\vec{u}\in(\mathbb{Z}^d)^3} \pi_{\vec{m}}(\vec{u}) = z_D^3 \sum_{\vec{m}\in\mathbb{Z}_+^3} \hat{\pi}_{\vec{m}}(0). \tag{4.10}$$

For $N \in \mathbb{N}$ we define

$$\pi_{\vec{m}}^{(N)}(\vec{u}) = \sum_{L\in\mathcal{L}^{(N)}(\mathcal{S}_{\vec{m}})} \sum_{\omega\in\Omega_{\mathcal{S}_{\vec{m}}}(\vec{u})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{S}_{\vec{m}}}:\\ R_s\ni\omega(s)\forall s\in\mathcal{S}_{\vec{m}}}} \prod_{t\in\mathcal{S}_{\vec{m}}} W(R_t) \prod_{st\in L}(-U_{st}) \prod_{vv'\in\mathcal{C}(L)}(1+U_{vv'}), \tag{4.11}$$

where $L \in \mathcal{L}^{(N)}(\mathcal{S}_{\vec{m}})$ is the set of *laces on* $\mathcal{S}_{\vec{m}}$ *with* N *bonds* and $\mathcal{C}(L)$ denotes the set of bonds which are *compatible with* L. We refer to [17, Section 2] for the precise definitions, and give only a rough description here: A lace L on $\mathcal{S}_{\vec{m}}$ is either a minimal graph covering $\mathcal{S}_{\vec{m}}$ (i.e. the removal of any bond in L results in a graph that no longer covers $\mathcal{S}_{\vec{m}}$) or one that is almost minimal (in this case there is a bond covering the branch point whose removal results in a minimal graph covering $\mathcal{S}_{\vec{m}}$). There is a rule for (uniquely) defining a lace $\boldsymbol{L}(\Gamma)$ associated to a connected graph Γ on $\mathcal{S}_{\vec{m}}$. For a fixed lace L the bonds compatible with L are those for which adding them to L results in a connected graph Γ' for which $\boldsymbol{L}(\Gamma') = L$.

In our work we only need a few facts about $\pi_{\vec{m}}^{(N)}(\cdot)$, including the obvious fact that $\pi_{\vec{m}}^{(N)}(\vec{u}) \geq 0$ and that (see [17, (4.28)–(4.29)]) if some $m_i > 0$ then

$$\pi_{\vec{m}}(\vec{u}) = \sum_{N=1}^{\infty} (-1)^N \pi_{\vec{m}}^{(N)}(\vec{u}). \tag{4.12}$$

Equations (4.12) and (4.11) are the lace expansion. A key result about this expansion is the following minor correction of [17, Proposition 4.13].

Proposition 4.17. *Fix $d > 8$. There exists $L_0(d)$ such that for all $L \geq L_0$ there exists a constant $C > 0$ (independent of L) and $B_N(\vec{m}) > 0$ such that for all $N \geq 1$ and $\vec{\ell} = (\ell_1, \ell_2, \ell_3) \in \mathbb{Z}_+^3 \setminus \{(0,0,0)\}$ we have for $j \in \{1,2,3\}$,*

$$\sum_{\vec{u}\in(\mathbb{Z}^d)^3} |u_j|^{2q} \pi_{\vec{m}}^{(N)}(\vec{u}) \leq C(L^2N^2\|\vec{m}\|_\infty)^q B_N(\vec{m}), \qquad \text{for } q \in \{0,1\}, \quad \text{and} \tag{4.13}$$

$$\sum_{N=1}^{\infty} \sum_{\vec{m}:m_j\geq\ell_j} B_N(\vec{m}) \leq \frac{C}{(\ell_j \vee 1)^{\frac{d-8}{2}}}, \quad \text{and} \tag{4.14}$$

$$\sum_{N=1}^{\infty} N^2 \sum_{\vec{m}\leq\vec{\ell}} \|\vec{m}\|_\infty B_N(\vec{m}) \leq C \times \begin{cases} \|\vec{\ell}\|_\infty^{\frac{10-d}{2}\vee 0} & \text{if } d \neq 10 \\ \log(\|\vec{\ell}\|_\infty \vee 2) & \text{if } d = 10. \end{cases} \tag{4.15}$$

The correction is that the $\vee 1$ and $\vee 2$ are missing in [17, Proposition 4.13], but what we have stated above is what is actually proved therein. Here we have also not included the extra decay in L appearing in these bounds in [17, Proposition 4.13] as we do not need it.

4.4. Decomposition of $\mathcal{Q}_{\mathcal{N}^+}$. By (4.8) we can see that $Q_{\mathcal{N}^+(F,\check{\boldsymbol{n}})}(\check{\mathbf{y}})$ can be decomposed into 4 parts: the connected component $\mathcal{S}_{\vec{m}}$ of bonds stemming from the branching point (term in $\mathcal{J}$) and the three subgraphs of $\mathcal{N}^+$ remaining after the removal of this connected component (terms in $\mathcal{K}$). These four subgraphs are not connected by any bonds by definition of $\mathcal{J}$ and $\mathcal{K}$ on the respective subgraphs. Furthermore the star-shaped subgraph $\mathcal{S}_{\vec{m}}$ contains the special point b, while all other special points are contained in one of the other subgraphs. This means that our problem can be reduced to three independent similar problems for smaller lengths. This reasoning translates into the following lemma which can be proved exactly as for [17, Lemma 4.14] so we do not repeat the proof. Recall the definition of $\check{\mathbf{y}}^{v_i,i}$ and the marked skeleton networks $\mathcal{N}^+_{i,\vec{m}}$ in Definition 4.12.

Lemma 4.18. *For a marked skeleton network $\mathcal{N}^+ = \mathcal{N}^+(F, \check{\boldsymbol{n}})$ and $\check{\mathbf{y}} \in (\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+)}$,*

$$Q_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\vec{m}\in\mathcal{H}_{\check{\boldsymbol{n}}_b}} \sum_{\vec{u}\in(\mathbb{Z}^d)^3} \pi_{\vec{m}}(\vec{u}) \prod_{i=1}^{3} \Big(z_D \sum_{v_i} D(v_i - u_i) t_{\mathcal{N}^+_{i,\vec{m}}}(\check{\mathbf{y}}^{v_i,i}) \Big).$$

For any marked skeleton network $\mathcal{N}^+(F, \check{\boldsymbol{n}})$, we introduce the Fourier transform of $t_{\mathcal{N}^+}$ and $Q_{\mathcal{N}^+}$ for any $\check{\boldsymbol{k}} = (\check{k}_{e,j})_{j\in[j_e],e\in[2r-1]}$ by

$$\hat{t}_{\mathcal{N}^+}(\check{\boldsymbol{k}}) = \sum_{\check{\mathbf{y}}\in(\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+)}} \prod_{e=1}^{2r-1} \prod_{j=1}^{j_e} \mathrm{e}^{\mathrm{i}\check{k}_{e,j}\check{y}_{e,j}} t_{\mathcal{N}^+}(\check{\mathbf{y}}),$$

$$\hat{Q}_{\mathcal{N}^+}(\check{\boldsymbol{k}}) = \sum_{\check{\mathbf{y}}\in(\mathbb{Z}^d)^{\boldsymbol{E}(\mathcal{N}^+)}} \prod_{e=1}^{2r-1} \prod_{j=1}^{j_e} \mathrm{e}^{\mathrm{i}\check{k}_{e,j}\check{y}_{e,j}} Q_{\mathcal{N}^+}(\check{\mathbf{y}}).$$

Lemma 4.18 implies that

$$\hat{Q}_{\mathcal{N}^+}(\check{\boldsymbol{k}}) = z_D^3 \sum_{\vec{m}\in\mathcal{H}_{\check{\boldsymbol{n}}_b}} \hat{\pi}_{\vec{m}}(\check{\boldsymbol{k}}^b) \prod_{i=1}^{3} \hat{D}(\check{k}_{\check{e}_i}) \hat{t}_{\mathcal{N}^+_{i,\vec{m}}}(\check{\boldsymbol{k}}^i), \tag{4.16}$$

where $\check{\boldsymbol{k}}^b = (\check{k}_{\check{e}_1}, \check{k}_{\check{e}_2}, \check{k}_{\check{e}_3})$ (meaning the part of $\check{\boldsymbol{k}}$ corresponding to marked edges incident to the branch point b) and $\check{\boldsymbol{k}}^i$ denotes the vector of $\check{k}_{e,j}$ corresponding to marked edges e in $\mathcal{N}^+_{i,\vec{m}}$. Note that (4.16) is exactly the "marked" network analog of the unmarked relation [17, (4.39)].

4.5. Proof of Proposition 2.11. The proof now closely follows that of [17, Theorem 4.8] with obvious (and straightforward) modifications. We will present the main ideas, but not the details. The goal is to prove that

$$\hat{t}_{\mathcal{N}^+(F,\check{n})}\Big(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\Big) = \rho C_V^{r-1} C_A^{2r-1} \prod_{e=1}^{2r-1} \prod_{i=1}^{j_e} \mathrm{e}^{-\sigma_0^2 \frac{\check{k}_{e,i}^2}{2} \frac{\check{n}_{e,i}}{n}} + \mathcal{O}\Big(\sum_{e=1}^{2r-1} \sum_{i=1}^{j_e} \frac{1}{\check{n}_{e,i}^{\frac{d-8}{2}}} \Big) + \mathcal{O}\Big(\sum_{e=1}^{2r-1} \sum_{i=1}^{j_e} \frac{\big|\check{\boldsymbol{k}}\big|^2 \check{n}_{e,i}^{1-\delta}}{n} \Big).$$

From (4.9), our bounds on the error terms therein (Lemma 4.14), and (4.16) we have that

$$\hat{t}_{\mathcal{N}^+(F,\check{n})}\Big(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\Big) = \hat{Q}_{\mathcal{N}^+}\Big(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\Big) + \mathcal{O}\Big(\sum_{e=1}^{2r-1}\sum_{i=1}^{j_e}\frac{1}{\check{n}_{e,i}^{\frac{d-8}{2}}}\Big)$$
$$= z_D^3 \sum_{\vec{m}\in\mathcal{H}_{\check{\boldsymbol{n}}_b}} \hat{\pi}_{\vec{m}}\Big(\frac{\check{\boldsymbol{k}}^b}{\sqrt{n}}\Big)\prod_{j=1}^{3}\hat{D}\Big(\frac{\check{k}_{\check{e}_j}}{\sqrt{n}}\Big)\hat{t}_{\mathcal{N}^+_{j,\vec{m}}}\Big(\frac{\check{\boldsymbol{k}}^j}{\sqrt{n}}\Big) + \mathcal{O}\Big(\sum_{e=1}^{2r-1}\sum_{i=1}^{j_e}\frac{1}{\check{n}_{e,i}^{\frac{d-8}{2}}}\Big). \tag{4.17}$$

We proceed by induction on r for networks with shape $F \in \Sigma_r$, using Lemma 4.3 for the initializing case ($r = 1$).

Let $\delta \in (0, 1 \wedge \frac{d-8}{2})$. By the induction hypothesis applied to each $\mathcal{N}^+_{j,\vec{m}}$ (having $r_j + 1$ leaves, where $r_1 = 1$ and $r_2 + r_3 = r$) we may write

$$\prod_{j=1}^{3}\hat{t}_{\mathcal{N}^+_{j,\vec{m}}}\Big(\frac{\check{\boldsymbol{k}}^j}{\sqrt{n}}\Big) \approx \rho^3 C_V^{r-2} C_A^{2r-1}\prod_{j=1}^{3}\Big[\prod_{\check{e}\in\boldsymbol{E}(\mathcal{N}^+_{j,\vec{m}})} \mathrm{e}^{-\sigma_0^2\frac{\check{k}_{\check{e}}^2}{2}\frac{\check{n}_{\check{e}}^{\vec{m},j}}{n}}\Big], \tag{4.18}$$

where we recall that the notation $\check{n}_{\cdot}^{\vec{m},j}$ was introduced in Definition 4.12. The error terms in the above approximation are obtained from the induction hypothesis and Remark 4.13 (using the fact that $\check{n}_{\check{e}}^{\vec{m},j}$ is comparable to $\check{n}_{\check{e}}$—they are identical unless $\check{e} = \check{e}_j$ for some $j \le 3$ - since $\vec{m} \in \mathcal{H}_{\check{\boldsymbol{n}}_b}$). We then use the fact that

$$\prod_{j=1}^{3}\mathrm{e}^{-\sigma_0^2\frac{\check{k}_{\check{e}_j}^2}{2}\frac{\check{n}_{\check{e}_j}^{\vec{m},j}}{n}} - \prod_{j=1}^{3}\mathrm{e}^{-\sigma_0^2\frac{\check{k}_{\check{e}_j}^2}{2}\frac{\check{n}_{\check{e}_j}}{n}} \le \frac{C}{n}\sum_{j=1}^{3}\check{k}_{\check{e}_j}^2 m_j,$$

(4.13) with $q = 0$ and (4.15) of Proposition 4.17, and $|\hat{D}| \le 1$ to get an error term in (4.17) (when replacing the right-hand side of (4.18) with $\check{n}_{\check{e}}$ in the exponent instead of $\check{n}_{\check{e}}^{\vec{m},j}$) of at most $\mathcal{O}(\sum_{j=1}^{3}\check{k}_{\check{e}_j}^2\check{n}_{\check{e}_j}^{1-\delta}n^{-1})$. For the relevant details of this part of the argument, and in particular for the bounds on the error terms, one can look at the derivation of [17, (4.56)].

Now $\hat{D}(\check{k}_{\check{e}_j}/\sqrt{n}) = 1 + \mathcal{O}(|\check{k}_{\check{e}_j}|^2/n)$ and

$$|\hat{\pi}_{\vec{m}}(\check{\boldsymbol{k}}^b/\sqrt{n}) - \hat{\pi}_{\vec{m}}(0)| \le C\frac{|\check{\boldsymbol{k}}^b|^2}{n}\sum_{j=1}^{3}\sum_{\vec{u}}|u_j|^2|\pi_{\vec{m}}(\vec{u})|,$$

which, when summed over $m_j \le \check{n}_{\check{e}_j}$, $j = 1, 2, 3$, gives at most $CL^2n^{-1}|\check{\boldsymbol{k}}^b|^2\sum_{j=1}^{3}\check{n}_{\check{e}_j}^{1-\delta}$ (see (4.13) with $q = 1$ and (4.15)). Next, $\sum_{\vec{m}\in\mathcal{H}_{\check{\boldsymbol{n}}_b}}\hat{\pi}_{\vec{m}}(0)$ differs from the full sum $\sum_{\vec{m}}\hat{\pi}_{\vec{m}}(0)$ by at most $C\sum_{j=1}^{3}\check{n}_{\check{e}_j}^{-(d-8)/2}$ by (4.14). Combining the above and recalling (4.10) and that $C_V = \rho^2 V$ reveals that

$$\hat{t}_{\mathcal{N}^+(F,\check{n})}\Big(\frac{\check{\boldsymbol{k}}}{\sqrt{n}}\Big) \approx z_D^3\sum_{\vec{m}}\hat{\pi}_{\vec{m}}(0)\rho^3 C_V^{r-2}C_A^{2r-1}\prod_{e=1}^{2r-1}\prod_{i=1}^{j_e}\mathrm{e}^{-\sigma_0^2\frac{\check{k}_{e,i}^2}{2}\frac{\check{n}_{e,i}}{n}}$$
$$= \rho C_V^{r-1}C_A^{2r-1}\prod_{e=1}^{2r-1}\prod_{i=1}^{j_e}\mathrm{e}^{-\sigma_0^2\frac{\check{k}_{e,i}^2}{2}\frac{\check{n}_{e,i}}{n}}.$$

An analysis of the error terms involved in the various $\approx$ approximations is handled rigorously in [17, Sections 4.3–4.5], making use of Proposition 4.17. □

4.6. Proof of Lemma 4.14. The proof of Lemma 4.14 relies on diagrammatic estimates. These estimates are built from a single lemma which gives the bounds on the simplest diagrams. For $u \in \mathbb{Z}^d$ let us denote

$$h_m(u) = \begin{cases} z_D^2 (D * t_{m-2} * D)(u) & \text{if } m \geq 2 \\ z_D D(u) & \text{if } m = 1 \\ \mathbb{1}_{\{u=0\}} & \text{if } m = 0, \end{cases}$$

where $t_m(u) = \rho \mathbb{P}(u \in \mathcal{T}_m)$ (so $t_0(u) = \rho \mathbb{1}_{\{u=0\}}$), and we recall that $*$ denotes the convolution of functions on $\mathbb{Z}^d$. Note that in [17] there is a ζ in the definition, but this $\zeta = 1$ because of Lemma 3.9 of [17]. Note that for $m \geq 2$,

$$\begin{aligned} t_m(u) &= \sum_{\omega: 0 \xrightarrow{m} u} W(\omega) \sum_{\substack{(R_s)_{0 \leq s \leq m}: \\ R_s \ni \omega(s) \forall s}} \prod_{t=0}^{m} W(R_t) \prod_{uv \in [0,m]} [1 + U_{uv}] \\ &\leq \sum_{\omega: 0 \xrightarrow{m} u} W(\omega) \sum_{R_m \ni \omega(m)} W(R_m) \sum_{R_0 \ni \omega(0)} W(R_0) \sum_{\substack{(R_s)_{1 \leq s \leq m-1}: \\ R_s \ni \omega(s) \forall s}} \prod_{t=1}^{m-1} W(R_t) \prod_{uv \in [1,m-1]} [1 + U_{uv}] \\ &\leq \sum_{\omega: 0 \xrightarrow{m} u} W(\omega) \rho^2 \sum_{\substack{(R_s)_{1 \leq s \leq m-1}: \\ R_s \ni \omega(s) \forall s}} \prod_{t=1}^{m-1} W(R_t) \prod_{uv \in [1,m-1]} [1 + U_{uv}] = \rho^2 h_m(u). \quad (4.19) \end{aligned}$$

Let us recall partially from [17, Lemma 5.4],[4] in which the function $\varrho : \mathbb{Z}^d \to \mathbb{R}_+$ is defined by $\varrho(x) = \rho \mathbb{P}(x \in \mathcal{T})$.

Lemma 4.19. *Fix $d > 8$. There exists $L_0(d)$ such that for all $L \geq L_0$: For any $l \geq 1$ there exists $C_l > 0$ such that for all $k \in \{0, 1, 2, 3, 4\}$ and $\vec{m}^{(l)} = (m_1, \ldots, m_l) \in \mathbb{Z}_+^l$ and $m = \sum_{i=1}^l m_i$, then*

$$\left\| *_{i=1}^l h_{m_i} * \varrho^{(k)} \right\|_\infty \leq \frac{C_l}{m^{\frac{d-2k}{2}}}, \text{ and } \left\| *_{i=1}^l h_{m_i} \right\|_1 \leq C_l.$$

For a given skeleton network $\mathcal{N}^+$, let $r_+ = \#\boldsymbol{E}(\mathcal{N}^+)$. If there is a bond uu' covering two special points then either we can find two non-neighbouring marked edges $\check{e} \ni u$ and $\check{e}' \ni u'$, or (at least) one of u, u' is a leaf of $\mathcal{N}^+$. In order to accommodate the latter cases, for the proof of Lemma 4.14(e:$\mathcal{R}$) it is notationally convenient to adjoin to each leaf in $\mathcal{N}^+$ a "phantom" marked edge of length 0, and write $\boldsymbol{E}(\mathcal{N}^{++})$ for this enlarged set of marked edges. For marked edges $\check{e}, \check{e}' \in \boldsymbol{E}(\mathcal{N}^{++})$ write $\check{e} \sim \check{e}'$ if they are adjacent, and $\check{e} \not\sim \check{e}'$ otherwise. Recall from (4.4) that in the notation U_{st}, st is a pair of vertices in $\mathcal{N}^+$. For non-adjacent marked edges $\check{e}, \check{e}' \in \boldsymbol{E}(\mathcal{N}^{++})$ and $m \leq \check{n}_{\check{e}}$, and $m' \leq \check{n}'_{\check{e}'}$, write $st(\check{e}, \check{e}', m, m')$ to denote the pair of vertices in $\mathcal{N}^+$ corresponding to the m-th vertex

[4] [17, Lemma 5.4] is stated under some additional hypotheses (display (3.24) in that paper). Nevertheless [17, Theorem 3.7] shows that the assumptions (3.24) are met.

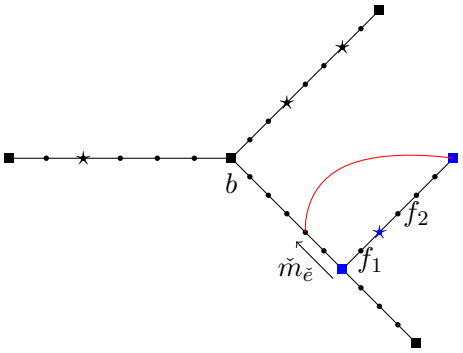

Fig. 12. A skeleton network $\mathcal{N}^+$ with a bond in $\mathcal{R}$. This bond has endpoints in the marked edge $\check{e} \in \mathcal{N}^+$ and the "phantom" marked edge $\check{e}' \in \mathcal{N}^{++}$ of lengths $\check{n}_{\check{e}} = 6$ and $\check{n}'_{\check{e}'} = 0$ respectively. We write $st(\check{e}, \check{e}', \check{m}_{\check{e}}, \check{m}'_{\check{e}'})$ for this bond. Here, $\check{m}_{\check{e}} = 2$ is indicated, while $\check{m}'_{\check{e}'} = 0$. The set of marked edges $\boldsymbol{E}^+_{\check{e},\check{e}'}$ on the path from $\check{e}$ to $\check{e}'$ is $\{f_1, f_2\}$ from ■ to ⋆ and ⋆ to ■ as indicated

along marked edge $\check{e}$ in the direction away from $\check{e}'$ and the m'-th vertex along marked edge $\check{e}'$ in the direction away from $\check{e}'$. If e.g. $\check{e}$ was one of the phantom marked edges then $\check{n}_{\check{e}} = 0$ and the relevant vertex is actually the leaf that $\check{e}$ was adjoined to. See e.g. Fig. 12.

For $0 \le a \le b \le \check{n}_{\check{e}}$, write $\check{e}[a, b]$ to denote that part of the marked edge $\check{e}$ consisting of the a-th to the b-th vertices (with ordering directed away from $\check{e}'$ as above) and similarly define $\check{e}'[a', b']$ for $0 \le a' \le b' \le \check{n}'_{\check{e}'}$.

Proof of Lemma 4.14(e:$\mathcal{R}$). In the definition of $\phi^{\mathcal{R}}_{\mathcal{N}^+}$ (see (4.6)), we can see that

$$1 - \prod_{vv' \in \mathcal{R}} [1 + U_{vv'}] \le \sum_{\check{e},\check{e}' \in \boldsymbol{E}(\mathcal{N}^{++}),\ \check{e} \not\sim \check{e}'} \ \sum_{\substack{\check{m}_{\check{e}} \le \check{n}_{\check{e}}, \\ \check{m}_{\check{e}'} \le \check{n}_{\check{e}'}}} -U_{st(\check{e},\check{e}',\check{m}_{\check{e}},\check{m}'_{\check{e}'})},$$

since if there is a bond uu' covering two special points then we can find two non-neighbouring marked edges in $\boldsymbol{E}(\mathcal{N}^{++})$ containing u and u' respectively (for more details see [17, Section 6.5.1]).

For a marked edge $\check{f} \in \boldsymbol{E}(\mathcal{N}^+)$, write $\overline{\prod}_{uv \in \check{f}}$ for a product over pairs of distinct vertices u, v in the interior of $\check{f}$ (i.e. u, v are vertices in $\check{f}$ that are not the endvertices of $\check{f}$). For non-adjacent marked edges $\check{e}, \check{e}'$ and $0 \le a \le b \le \check{n}_{\check{e}}$ as above, write $\overline{\prod}_{uv \in \check{e}[a,b]}$ for a product over pairs of distinct u, v in $\check{e}[a, b]$ that are neither endvertex of this set, and similarly define $\overline{\prod}_{uv \in \check{e}'[a',b']}$ for $0 \le a' \le b' \le \check{n}'_{\check{e}'}$.

Fix distinct $\check{e} \not\sim \check{e}'$ in $\boldsymbol{E}(\mathcal{N}^{++})$ and $\check{m}_{\check{e}} \le \check{n}_{\check{e}}, \check{m}_{\check{e}'} \le \check{n}_{\check{e}'}$. By ignoring the constraints of non-intersection between various R_i (bounding some $1 + U_{uu'}$ by 1), we obtain

$$\begin{aligned}\prod_{uu' \in \mathcal{R}^c} [1 + U_{uu'}] \le \prod_{\check{f} \in \boldsymbol{E}(\mathcal{N}^+) \setminus \{\check{e},\check{e}'\}} \Big[\overline{\prod_{uu' \in \check{f}}} [1 + U_{uu'}] \Big] \\ \times \Big[\overline{\prod_{u_0 u'_0 \in \check{e}[0,\check{m}_{\check{e}}]}} [1 + U_{u_0 u'_0}] \Big] \Big[\overline{\prod_{u_1 u'_1 \in \check{e}[\check{m}_{\check{e}},\check{n}_{\check{e}}]}} [1 + U_{u_1 u'_1}] \Big] \\ \times \Big[\overline{\prod_{u_2 u'_2 \in \check{e}'[0,\check{m}'_{\check{e}'}]}} [1 + U_{u_2 u'_2}] \Big] \Big[\overline{\prod_{u_3 u'_3 \in \check{e}'[\check{m}'_{\check{e}'},\check{n}'_{\check{e}'}]}} [1 + U_{u_3 u'_3}] \Big].\end{aligned}$$

where e.g. if $\check{m}_{\check{e}} \in \{0, \check{n}_{\check{e}}\}$ then the corresponding empty product is 1. (Note that this kind of approach is used to prove (2.25) as well as the more general statement appearing in Remark 2.12.) Using the above inequalities we can see that

$$\begin{aligned}
0 &\le \Big(\prod_{uu' \in \mathcal{R}^c} [1 + U_{uu'}] \Big) \Big(1 - \prod_{vv' \in \mathcal{R}} [1 + U_{vv'}] \Big) \\
&\le \sum_{\substack{\check{e},\check{e}' \in \boldsymbol{E}(\mathcal{N}^{++}), \\ \check{e} \not\sim \check{e}'}} \sum_{\substack{\check{m}_{\check{e}} \le \check{n}_{\check{e}}, \\ \check{m}_{\check{e}'} \le \check{n}_{\check{e}'}}} [-U_{st(\check{e},\check{e}',\check{m}_{\check{e}},\check{m}'_{\check{e}'})}] \prod_{\check{f} \in \boldsymbol{E}(\mathcal{N}^+) \setminus \{\check{e},\check{e}'\}} \Big[\overline{\prod_{uu' \in \check{f}}} [1 + U_{uu'}] \Big] \\
&\quad \times \Big[\overline{\prod_{u_0 u'_0 \in \check{e}[0,\check{m}_{\check{e}}]}} [1 + U_{u_0 u'_0}] \Big] \Big[\overline{\prod_{u_1 u'_1 \in \check{e}[\check{m}_{\check{e}},\check{n}_{\check{e}}]}} [1 + U_{u_1 u'_1}] \Big] \\
&\quad \times \Big[\overline{\prod_{u_2 u'_2 \in \check{e}'[0,\check{m}'_{\check{e}'}]}} [1 + U_{u_2 u'_2}] \Big] \Big[\overline{\prod_{u_3 u'_3 \in \check{e}'[\check{m}'_{\check{e}'},\check{n}'_{\check{e}'}]}} [1 + U_{u_3 u'_3}] \Big],
\end{aligned} \tag{4.20}$$

Note that if e.g. $\check{e}$ is a phantom marked edge then the corresponding sum over $\check{m}_{\check{e}}$ contains only the value $0 = \check{n}_{\check{e}}$.

Now, the ω in (4.6) can be broken up at every special point and at the two vertices corresponding to $st(\check{e}, \check{e}', m, m')$. The graph then becomes broken up into (at most) $r_+ + 2$ segments. Let us now introduce the set $\boldsymbol{E}^+_{\check{e},\check{e}'}$ of marked edges which connect (but do not include) $\check{e}$ to $\check{e}'$ which is non-empty since $\check{e}$ and $\check{e}'$ are not neighbours, and $\bar{\boldsymbol{E}}^+_{\check{e},\check{e}'} = \boldsymbol{E}^+_{\check{e},\check{e}'} \cup \{\check{e}, \check{e}'\}$. Letting $\check{\mathbf{y}}_{\check{e},\check{e}'} = (\check{y}_{\check{f}})_{\check{f} \in \boldsymbol{E}^+_{\check{e},\check{e}'}}$ with each $\check{y}_{\check{f}} \in \mathbb{Z}^d$ we have from (4.20) and Remark 1.1 that (cf. [17, (6.18)])

$$\begin{aligned}
&\sum_{\check{\mathbf{y}}} \Big| \phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}}) \Big| \\
&\le \rho^{2(r_+ + 2)} \sum_{\substack{\check{e},\check{e}' \in \boldsymbol{E}(\mathcal{N}^{++}), \\ \check{e} \not\sim \check{e}'}} \sum_{\substack{\check{m}_{\check{e}} \le \check{n}_{\check{e}}, \\ \check{m}_{\check{e}'} \le \check{n}_{\check{e}'}}} \Big(\prod_{\check{f}' \in \boldsymbol{E}(\mathcal{N}^+) \setminus \bar{\boldsymbol{E}}^+_{\check{e},\check{e}'}} \sum_{\check{y}_{\check{f}'}} h_{\check{n}_{\check{f}'}}(\check{y}_{\check{f}'}) \Big) \sum_{\check{\mathbf{y}}_{\check{e},\check{e}'}} \sum_{u_{\check{e}}, u_{\check{e}'}} \Big[\prod_{\check{f} \in \boldsymbol{E}^+_{\check{e},\check{e}'}} h_{\check{n}_{\check{f}}}(\check{y}_{\check{f}}) \\
&\times h_{\check{m}_{\check{e}}}(u_{\check{e}}) h_{\check{m}_{\check{e}'}}(u_{\check{e}'}) \varrho^{(2)} \Big(u_{\check{e}} + u_{\check{e}'} + \sum_{\check{f} \in \boldsymbol{E}^+_{\check{e},\check{e}'}} \check{y}_{\check{f}} \Big) \sum_{\check{y}_{\check{e}}} h_{\check{n}_{\check{e}} - \check{m}_{\check{e}}}(\check{y}_{\check{e}} - u_{\check{e}}) \sum_{\check{y}_{\check{e}'}} h_{\check{n}_{\check{e}'} - \check{m}_{\check{e}'}}(\check{y}_{\check{e}'} - u_{\check{e}'}) \Big].
\end{aligned}$$

This arises because e.g. if $\check{f}$ and $\check{f}'$ are two distinct marked edges for which there is no U_{st} term appearing anywhere in (4.20) with s and t vertices of $\check{f}$ and $\check{f}'$ respectively, then the corresponding segments of ω (and the sets of lattice trees $R_\cdot$ hanging off them) have been decoupled. Segments of ω and corresponding elements of $\check{\mathbf{y}}$ can then be summed over "independently", with factors of ρ arising at endvertices, similarly to (4.19). Similarly, the presence of the term $[-U_{st(\cdot,\cdot,\cdot,\cdot)}]$ in (4.20) forces two corresponding trees $R_\cdot$ to intersect, which yields the $\varrho^{(2)}$ term above. Recalling that $\sum_{\check{y}} h_n(\check{y}) \le C_1$ for any n by Lemma 4.19 we see that

$$\sum_{\check{\mathbf{y}}} \Big| \phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\mathbf{y}}) \Big| \le \rho^{2(r_+ + 2)} C_1^{r_+} \sum_{\substack{\check{e},\check{e}' \in \boldsymbol{E}(\mathcal{N}^{++}), \\ \check{e} \not\sim \check{e}'}} \sum_{\substack{\check{m}_{\check{e}} \le \check{n}_{\check{e}}, \\ \check{m}_{\check{e}'} \le \check{n}_{\check{e}'}}} (h_{\check{m}_{\check{e}}} * h_{\check{m}_{\check{e}'}} * \varrho^{(2)} \underset{\check{f} \in \boldsymbol{E}^+_{\check{e},\check{e}'}}{*} h_{\check{n}_{\check{f}}})(0).$$

(The power of C_1 is $r_+ - \#\bar{\boldsymbol{E}}^+_{\check{e},\check{e}'} + 2 \le r_+$, and so assuming $C_1 \ge 1$ without loss of generality, the above follows.) The notation in the last convolution above means that

there is one term in the convolution for each $\check{f} \in \boldsymbol{E}^+_{\check{e},\check{e}'}$. By Lemma 4.19 with $k = 2$ and $l = l_+ := 2 + \#\boldsymbol{E}^+_{\check{e},\check{e}'}$ we have that for $n_{\check{e},\check{e}'} = \sum_{\check{f}\in\boldsymbol{E}^+_{\check{e},\check{e}'}} \check{n}_{\check{f}}$,

$$\begin{aligned}\sum_{\check{\boldsymbol{y}}} \left|\phi^{\mathcal{R}}_{\mathcal{N}^+}(\check{\boldsymbol{y}})\right| &\le \rho^{2(r_++2)} C_1^{r_+} \sum_{\substack{\check{e},\check{e}'\in\boldsymbol{E}(\mathcal{N}^{++}),\\ \check{e}\not\sim\check{e}'}} \sum_{\substack{\check{m}_{\check{e}}\le\check{n}_{\check{e}},\\ \check{m}_{\check{e}'}\le\check{n}_{\check{e}'}}} \frac{C_{l_+}}{(\check{m}_{\check{e}}+\check{m}_{\check{e}'}+n_{\check{e},\check{e}'})^{\frac{d-4}{2}}}\\ &\le C(r_+) \sum_{\substack{\check{e},\check{e}'\in\boldsymbol{E}(\mathcal{N}^{++}),\\ \check{e}\not\sim\check{e}'}} \frac{1}{n_{\check{e},\check{e}'}^{\frac{d-8}{2}}}\\ &\le C'(r_+) \sum_{e=1}^{2r-1}\sum_{j=1}^{j_e} \frac{1}{\check{n}_{e,j}^{\frac{d-8}{2}}}.\end{aligned}$$

□

Proof of Lemma 4.14(e:π). Similarly to Lemma 4.18 (but note the change in the first summation) we have that

$$\phi^{\pi}_{\mathcal{N}^+}(\check{\boldsymbol{y}}) = \sum_{\vec{m}\in\overline{\mathcal{H}}_{\check{\boldsymbol{n}}_b}} \sum_{\vec{u}\in(\mathbb{Z}^d)^3} \pi_{\vec{m}}(\vec{u}) \prod_{j=1}^{3} \Big(z_D \sum_{v_j} D(v_j - u_j) t_{\mathcal{N}^+_{j,\vec{m}}}(\check{\boldsymbol{y}}^{v_j,j}) \Big).$$

Therefore, for any $\check{\boldsymbol{y}} \in (\mathbb{Z}^d)^{r_+}$,

$$\begin{aligned}\left|\phi^{\pi}_{\mathcal{N}^+}(\check{\boldsymbol{y}})\right| &\le C \left| \sum_{\vec{m}\in\overline{\mathcal{H}}_{\check{\boldsymbol{n}}_b}} \sum_{\vec{u}} \pi_{\vec{m}}(\vec{u}) \sum_{\check{\boldsymbol{y}}} \prod_{j=1}^{3} \sum_{v_j} D(v_j - u_j) t_{\mathcal{N}^+_{j,\vec{m}}}(\check{\boldsymbol{y}}^{v_j,j}) \right|\\ &\le C \sum_{N=1}^{\infty} \sum_{\vec{m}\in\overline{\mathcal{H}}_{\check{\boldsymbol{n}}_b}} \sum_{\vec{u}} \pi^{(N)}_{\vec{m}}(\vec{u}) \sum_{\check{\boldsymbol{y}}} \prod_{j=1}^{3} \sum_{v_j} D(v_j - u_j) t_{\mathcal{N}^+_{j,\vec{m}}}(\check{\boldsymbol{y}}^{v_j,j}).\end{aligned}$$

Using a generalisation of (2.25) as in Remark 2.12, and then (4.13) and (4.14), we have

$$\sum_{\check{\boldsymbol{y}}} \left|\phi^{\pi}_{\mathcal{N}^+}(\check{\boldsymbol{y}})\right| \le C \sum_{N=1}^{\infty} \sum_{\vec{m}\in\overline{\mathcal{H}}_{\check{\boldsymbol{n}}_b}} \sum_{\vec{u}} \pi^{(N)}_{\vec{m}}(\vec{u}) K_0^{r_+} \le C \sum_{N=1}^{\infty} \sum_{k=1}^{3} \sum_{\vec{m}:m_k\ge\check{n}_{\check{e}_k}/3} B_N(\vec{m}) \le \sum_{k=1}^{3} \frac{C'}{\check{n}_{\check{e}_k}^{\frac{d-8}{2}}}.$$

The result follows. □

The proof of Lemma 4.14(e:*b*) is again an adaptation of the proof in [17] (specifically in [17, Section 6.5.3]). Here we will indicate the changes to the argument required for the present setting of a marked skeleton network. We start by adapting [17, Definition 2.2]. Given a graph $\Gamma \in \mathcal{E}^b_{\mathcal{N}^+}$ on $\mathcal{N}^+$, a special point v of $\mathcal{N}^+$ and a marked edge e of which v is an endpoint, we define the *bond associated to e at v* as follows: If there is no bond in Γ covering v that has an endpoint strictly on e then there is no bond associated to e at v. Otherwise from the set of such bonds we choose the one whose endpoint in e is farthest from v. If this is not unique then we choose from this set one according to a fixed but arbitrary rule (e.g. choose from those whose other endpoint is strictly on some edge e' of smallest label the one whose endpoint on e' is farthest from v in this direction).

Proof of Lemma 4.14(e:*b*). Recall that

$$\phi^b_{\mathcal{N}^+}(\check{\mathbf{y}}) = \sum_{\omega\in\Omega_{\mathcal{N}^+}(\check{\mathbf{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}:\\ R_s\ni\omega(s)\forall s\in\mathcal{N}^+}} \prod_{t\in\mathcal{N}^+} W(R_t) \sum_{\Gamma\in\mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+}} \prod_{vv'\in\Gamma} U_{vv'}. \tag{4.21}$$

Recall also that $(\check{e}_i)_{i=1}^3$ are the marked edges adjacent to b and denote their end vertices (other than b) as $(\check{v}_i)_{i=1}^3$, which are special points.

For $F\subset\{1,2,3\}$ let

$$\mathcal{E}^b_{F,\mathcal{N}^+} = \big\{\Gamma\in\mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+} : \forall i\in F, \mathcal{A}_b(\Gamma) \text{ contains a nearest neighbour of } \check{v}_i\big\}.$$

Note that if $F\neq\{1,2,3\}$ this set may include Γ for which some $\mathcal{A}_b(\Gamma)$ also contains a nearest neighbour of $\check{v}_i$ for some $i\in\{1,2,3\}\setminus F$. Inclusion–exclusion over the sets F gives

$$\Big|\sum_{\Gamma\in\mathcal{E}^{-\mathcal{R},b}_{\mathcal{N}^+}}\prod_{vv'\in\Gamma}U_{vv'}\Big| \le \sum_{F\neq\varnothing}\Big|\sum_{\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}}\prod_{vv'\in\Gamma}U_{vv'}\Big|. \tag{4.22}$$

Given $\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}$ we define a subgraph $\Gamma_F\subset\Gamma$ to be the set of bonds $st\in\Gamma$ such that

- st is the bond associated to one of the marked edges $\check{e}_i$ at b, for some $i\in F$, or
- st is the bond associated to one of the marked edges $\check{e}_i$, at $\check{v}_i$ where $i\in F$, or
- st are both vertices in the marked edge $\check{e}_i$ for some $i\in F$.

Let $\mathcal{S}_F$ denote the largest connected subnetwork of $\mathcal{N}^+$ containing b that is covered by Γ_F. Then $\mathcal{S}_F$ is a star-shaped network of degree 3 or less (with branch point b) and Γ_F is a connected graph on $\mathcal{S}_F$. Moreover $\mathcal{S}_F$ contains at most $\#F+1$ special points of $\mathcal{N}^+$ (one of which is b) since Γ contains no bonds in $\mathcal{R}$. Note that the length of branch i of $\mathcal{S}_F$ is at least $\check{n}_i-1$. Let $\mathbb{S}_F(\mathcal{N}^+)$ denote (for fixed F) the set of possible $\mathcal{S}_F$ that can arise as above from graphs $\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}$. It follows that

$$\sum_{\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}}\prod_{vv'\in\Gamma}U_{vv'} = \sum_{\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)}\sum_{\substack{\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}:\\ \mathcal{S}_F(\Gamma)=\mathcal{S}}}\prod_{vv'\in\Gamma}U_{vv'}. \tag{4.23}$$

Now we may proceed as in [17, (6.23)–(6.28)], which we briefly discuss in the following paragraph but direct the reader to [17] for details. For fixed F and $\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)$ we have the notion of a lace on $\mathcal{S}$ containing N bonds and the set of bonds, $\mathcal{C}(L)$, which are compatible with the lace L, as described after (4.11). Similarly, given F, and $\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}$ such that $\mathcal{S}_F(\Gamma)=\mathcal{S}$ we have the lace associated to the subgraph Γ_F, which is a connected graph on $\mathcal{S}$. Thus, as in [17, (6.23)–(6.24)], we can write (4.23) as

$$\sum_{N=1}^{\infty}(-1)^N\sum_{\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)}\sum_{L\in\mathcal{L}^{(N),F}_{\mathcal{S}}}\Big[\prod_{st\in L}(-U_{st})\Big]\Big[\sum_{\substack{\Gamma\in\mathcal{E}^b_{F,\mathcal{N}^+}:\\ \mathcal{S}_F(\Gamma)=\mathcal{S},\\ \boldsymbol{L}(\Gamma_F)=L}}\prod_{vv'\in\Gamma}U_{vv'}\Big], \tag{4.24}$$

where the sum over L is a sum over (a certain subclass of all) laces on $\mathcal{S}$ containing exactly N bonds (for the definition of this subclass see [17, definition prior to (6.23)]). The last two pages of [17] show how to deal with the "messy" final sum in (4.24), by

breaking the sum over Γ into three sets: (i) sets of bonds on $\mathcal{S}$ that are compatible with L; (ii) sets of bonds that live on $\mathcal{N}^+ \setminus \mathcal{S}$; and (iii) sets of bonds st with one endpoint in $\mathcal{S}$ and one in $\mathcal{N}^+ \setminus \mathcal{S}$ for which $\mathcal{S}_F(L \cup \{st\}) = \mathcal{S}$ (in each case bonds in $\mathcal{R}$ are excluded). Using this decomposition we see that (4.24) is equal to

$$\sum_{N=1}^{\infty}(-1)^N \sum_{\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)} \sum_{L\in\mathcal{L}_{\mathcal{S}}^{(N),F}} \Big[\prod_{st\in L}(-U_{st})\Big]\Big[\sum_{\substack{\Gamma\in\mathcal{G}_{\mathcal{S}}^{-\mathcal{R},\mathrm{con}}:\\ \boldsymbol{L}(\Gamma)=L}} \prod_{st\in\Gamma} U_{st}\Big]\Big[\sum_{\Gamma'\in\mathcal{G}_{\mathcal{N}^+\setminus\mathcal{S}}^{-\mathcal{R}}} \prod_{st\in\Gamma'} U_{st}\Big]$$
$$\times\Big[\sum_{\substack{\Gamma^*\in\mathcal{G}_{\mathcal{S},\mathcal{N}^+\setminus\mathcal{S}}^{-\mathcal{R}}:\\ \mathcal{S}_F(L\cup\Gamma^*)=\mathcal{S}}} \prod_{st\in\Gamma} U_{st}\Big]$$
$$= \sum_{N=1}^{\infty}(-1)^N \sum_{\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)} \sum_{L\in\mathcal{L}_{\mathcal{S}}^{(N),F}} \Big[\prod_{st\in L}(-U_{st})\Big]\Big[\prod_{st\in\mathcal{C}(L)}[1+U_{st}]\Big]\Big[\prod_{st\in\mathcal{N}^+\setminus\mathcal{S}}[1+U_{st}]\Big] \tag{4.25}$$
$$\times \prod_{\substack{s\in\mathcal{S},t\in\mathcal{N}^+\setminus\mathcal{S}:\\ \mathcal{S}_F(L\cup\{st\})=\mathcal{S}}}[1+U_{st}]. \tag{4.26}$$

We bound the absolute value of the above by simply removing the factors $(-1)^N$ (everything else is non-negative). Then we can ignore the last product (4.26) (bound it by 1) and get an upper bound. Similarly we can discard any part of the last product in (4.25) to get an upper bound. For the latter we throw away all st such that s and t are on different connected components of $\mathcal{N}^+ \setminus \mathcal{S}$. We deduce from (4.21), (4.22), and the above that

$$\sum_{\check{\boldsymbol{y}}} |\phi^b_{\mathcal{N}^+}(\check{\boldsymbol{y}})|$$
$$\le \sum_{F\neq\varnothing}\sum_{N=1}^{\infty}\sum_{\check{\boldsymbol{y}}}\sum_{\omega\in\Omega_{\mathcal{N}^+}(\check{\boldsymbol{y}})} W(\omega) \sum_{\substack{(R_s)_{s\in\mathcal{N}^+}:\\ R_s\ni\omega(s)\forall s\in\mathcal{N}^+}} \prod_{t\in\mathcal{N}^+} W(R_t)$$
$$\times \sum_{\mathcal{S}\in\mathbb{S}_F(\mathcal{N}^+)} \Big\{\sum_{L\in\mathcal{L}_{\mathcal{S}}^{(N),F}} \Big[\prod_{st\in L}(-U_{s,t})\Big]\Big[\prod_{st\in\mathcal{C}(L)}[1+U_{st}]\Big]\Big\} \prod_{j=1}^{\Delta_{\mathcal{N}^+\setminus\mathcal{S}}} \Big[\prod_{st\in(\mathcal{N}^+\setminus\mathcal{S})_j}[1+U_{st}]\Big], \tag{4.27}$$

where $\Delta_{\mathcal{N}^+\setminus\mathcal{S}}$ denotes the number of disjoint components of $\mathcal{N}^+\setminus\mathcal{S}$ and the components are denoted by $(\mathcal{N}^+\setminus\mathcal{S})_j$. Here, the components $\mathcal{S}$, and $(\mathcal{N}^+\setminus\mathcal{S})_j$ for all j have now been decoupled, because there are no U_{st} terms where s and t are on different components. Recalling (4.11), the term in curly brackets in (4.27) (in combination with the part of ω and the trees $R_\cdot$ corresponding to $\mathcal{S}$) is the quantity that gives rise to $\pi^{(N)}_{\vec{m}}$ (where the m_i are the lengths of the branches of $\mathcal{S}$) except that we are summing over a restricted set of laces containing N bonds. But we can also sum over all $L \in \mathcal{L}^{(N)}(\mathcal{S})$, the set of laces on $\mathcal{S}$ with exactly N bonds, to get an upper bound. This gives rise to a bound

$$\sum_{\check{\boldsymbol{y}}}\Big|\phi^b_{\mathcal{N}^+}(\check{\boldsymbol{y}})\Big| \le \sum_{F\neq\varnothing}\sum_{N=1}^{\infty}\sum_{\substack{\vec{m}:\\ m_i\ge\check{n}_i-1\forall i\in F}}\sum_{\vec{u}} \pi^{(N)}_{\vec{m}}(\vec{u})C(r_+), \tag{4.28}$$

where we note that the sum over $\vec{m}$ arises from the sum over $\mathcal{S}$ seen in previous expressions, and the constant $C(r_+)$ arises from the generalisation of (2.25) noted in Remark 2.12. Finally use Proposition 4.17 to see that (4.28) is at most

$$C \sum_{F \neq \varnothing} \sum_{N=1}^{\infty} \sum_{\substack{\vec{m}: \\ m_i \geq \check{n}_i - 1 \\ \forall i \in F}} B_N(\vec{m}) \leq \sum_{i=1}^{3} \frac{C}{\check{n}_i^{\frac{d-8}{2}}}.$$

This proves the result. □

Acknowledgments The authors thank an anonymous referee for many comments that helped improve the paper.

Funding The work of Mark Holmes was supported by Future Fellowship FT160100166 from the Australian Research Council. Ed Perkins acknowledges the support of NSERC Discovery Grant RGPIN-2019-03928. Alex Fribergh acknowledges the support of NSERC Discovery Grant RGPIN-2020-05024. Manuel Cabezas is supported by Proyecto Fondecyt #1201090.

Data availability Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

1. Banova, T.: Convergence of the voter model to historical Brownian motion in $d > 2$ dimensions. PhD thesis in progress (2022)
2. Ben Arous, G., Cabezas, M., Fribergh, A.: The ant on lattice trees. Preprint (2022)
3. Bramson, M., Cox, J.T., Le Gall, J.-F.: Super-Brownian limits of voter model clusters. Ann. Probab. **29**, 1001–1032 (2001)
4. Cox, J.T., Durrett, R., Perkins, E.: Rescaled voter models converge to cuper-Brownian motion. Ann. Probab. **28**(1), 185–234 (2000)
5. Croydon, D.: Hausdorff measure of arcs and Brownian motion on Brownian spatial trees. Ann. Probab. **37**, 946–978 (2009)
6. Dawson, D.A., Perkins, E.: Historical processes. Mem. Am. Math. Soc. **93**, 454 (1991)
7. Derbez, E., Slade, G.: The scaling limit of lattice trees in high dimensions. Commun. Math. Phys. **193**, 69–104 (1998)
8. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, Hoboken (2005). https://doi.org/10.1002/9780470316658
9. Fitzner, R., van der Hofstad, R.: NoBLE for lattice trees and lattice animals. J. Stat. Phys. **185**, 13 (2021)
10. Hofstad, R.v.d., Holmes, M.: The survival probability and r-point functions in high dimensions. Ann. Math. **178**(2), 665–685 (2013)

11. Hofstad, R.v.d., Holmes, M., Perkins, E.: A criterion for convergence to super-Brownian motion on path space. Ann. Probab. **45**, 278–376 (2017)
12. Hofstad, R.v.d., Holmes, M., Slade, G.: An extension of the inductive approach to the lace expansion. Electron. Commun. Probab. **13**, 291–301 (2008)
13. Hofstad, R.v.d., Sakai, A.: Convergence of the critical finite-range contact process to super-Brownian motion above the upper critical dimension: the higher-point functions. Electron. J. Probab. **15**, 801–894 (2010)
14. Hofstad, R.v.d., Slade, G.: A generalised inductive approach to the lace expansion. Probab. Theory Relat. Fields **122**, 389–430 (2002)
15. Hofstad, R.v.d., Slade, G.: The lace expansion on a tree with application to networks of self-avoiding walks. Adv. Appl. Math. **30**(3), 471–528 (2003)
16. Hofstad, R.v.d., Slade, G.: Convergence of critical oriented percolation to super-Brownian motion above $4 + 1$ dimensions. Ann. Inst. H. Poincaré Probab. Stat. **39**(3), 413–485 (2003)
17. Holmes, M.: Convergence of lattice trees to super-Brownian motion above the critical dimension. Electron. J. Probab. **13**(23), 671–755 (2008)
18. Holmes, M.: Backbone scaling for critical lattice trees in high dimensions. J. Phys. A Math. Theor. **49**, 314001 (2016)
19. Holmes, M., Perkins, E.: Weak convergence of measure-valued processes and r-point functions. Ann. Probab. **35**(5), 1769–1782 (2007)
20. Holmes, M., Perkins, E.: On the range of lattice models in high dimensions. Prob. Theory Relat. Fields **176**, 941–1009 (2020)
21. Holmes, M., Perkins, E.: Random skeletons in high dimensional lattice trees. Preprint (2022)
22. Le Gall, J.-F.: Spatial Branching Processes, Random Snakes, and Partial Differential Equations. Birkhäuser, Basel (1999)
23. Perkins, E.: A space-time property of a class of measure-valued branching diffusions. Trans. Am. Math. Soc. **305**, 743–795 (1988)
24. Perkins, E.: Dawson–Watanabe superprocesses and measure-valued diffusions. In: Lectures on Probability Theory and Statistics, no. 1781, Ecole d'Eté de Probabilités de Saint Flour 1999. Springer, Berlin (2002)
25. Slade, G.: The lace expansion and its applications. In: Lecture Notes in Mathematics, no. 1879. Ecole d'Eté de Probabilités de Saint Flour 2004. Springer, Berlin (2006)

Communicated by M. Hairer