The Annals of Probability
2025, Vol. 53, No. 4, 1491–1509
https://doi.org/10.1214/24-AOP1739

TRANSIENCE OF VERTEX-REINFORCED JUMP PROCESSES WITH LONG-RANGE JUMPS

BY MARGHERITA DISERTORI[1,a], FRANZ MERKL[2,b] AND SILKE W. W. ROLLES[3,c]

[1] *Institute for Applied Mathematics & Hausdorff Center for Mathematics, University of Bonn,* [a] *disertori@iam.uni-bonn.de*
[2] *Mathematisches Institut, Ludwig-Maximilians-Universität München,* [b] *merkl@math.lmu.de*
[3] *Department of Mathematics, CIT, Technische Universität München,* [c] *srolles@cit.tum.de*

We show that the vertex-reinforced jump process on the d-dimensional lattice with long-range jumps is transient in any dimension d as long as the initial weights do not decay too fast. The main ingredients in the proof are: an analysis of the corresponding random environment on finite boxes, a comparison with a hierarchical model, and the reduction of the hierarchical model to a nonhomogeneous effective one-dimensional model. For $d \geq 3$ we also prove transience of the vertex-reinforced jump process with possibly long-range jumps as long as nearest-neighbor weights are large enough.

1. Introduction and results. Consider an undirected connected locally finite graph G with vertex set Λ and edge set E_Λ. Every edge $e = \{i, j\} \in E_\Lambda$ is given a positive weight $W_{ij} = W_{ji} = W_e > 0$. The vertex-reinforced jump process $(Y_t)_{t\geq 0}$ (VRJP for short) is a stochastic process in continuous time which starts at time 0 in a vertex $0 \in \Lambda$ and conditionally on $(Y_s)_{s\leq t}$ jumps from $Y_t = i$ along an edge $\{i, j\}$ at rate

$$W_{ij}\left(1 + \int_0^t 1_{\{Y_s=j\}}\, ds\right). \tag{1.1}$$

The process was conceived by Wendelin Werner in 2000 and first studied by Davis and Volkov [6, 7] on trees and finite graphs. For a recent overview of the subject, see [2]. It was shown by Sabot and Tarrès [13] and independently by Angel, Crawford, and Kozma [1] that for any graph of bounded degree there exists $W_1 > 0$ such that the VRJP is recurrent if $W_e \leq W_1$ for all $e \in E_\Lambda$. In contrast, Sabot and Tarrès [13] proved that for any $d \geq 3$, there exists $W_2 < \infty$ such that the VRJP on $\mathbb{Z}^d$ is transient if the weights W_e are equal to a constant $\geq W_2$. Sabot and Zeng in [14] proved the following zero-one law for infinite graphs. If the graph G together with the weights $W = (W_e)_{e\in E_\Lambda}$ is vertex transitive, the VRJP is either almost surely recurrent or it is almost surely transient. Poudevigne [11] proved some monotonicity in the weights W_e which implies that for $\mathbb{Z}^d$, $d \geq 3$, there is a unique phase transition between recurrence and transience when the weights W_e are constant. On $\mathbb{Z}^d$ for $d = 1, 2$, the VRJP with arbitrary constant weights is always recurrent. In one dimension, this was shown for $W_e = 1$ by Davis and Volkov [6] and for general constant W by Sabot and Tarrès [13]. Sabot [12] proved the result in two dimensions. The discrete-time process associated to VRJP on any *locally finite* graph is given by the annealed law of a random walk in random conductances; cf. [13] for finite graphs and [14] for locally finite infinite graphs.

We will consider here VRJP on $\mathbb{Z}^d$ with long-range interactions. This is not a locally finite graph, but VRJP can still be well-defined as long as $\sum_{e\ni i} W_e < \infty$ for all vertices $i \in \Lambda$. If in addition $\sup_{i\in\Lambda} \sum_{e\ni i} W_e < \infty$ holds, then VRJP jumps only finitely often up to any finite time. It seems that the existing results on VRJP have not considered this generalized context.

Received May 2023; revised June 2024.
MSC2020 subject classifications. Primary 60K35, 60K37; secondary 60G60.
Key words and phrases. Vertex-reinforced jump process, long-range random walk, hierarchical model, nonlinear supersymmetric hyperbolic sigma model.

Therefore, we decided not to rely on existing results on infinite graphs, but rather on finite pieces of infinite graphs only.

Also on infinite, possibly not locally finite graphs (Λ, E_Λ) up to the time that the process leaves a given finite set $\Lambda_N \subset \Lambda$ of vertices, the discrete-time process associated to VRJP is a random walk in random conductances. The corresponding random conductance for any edge $e = \{i, j\}$ may depend on Λ_N and can be encoded in the form $c_e = W_{ij} e^{u_i + u_j}$ with a random environment $u = (u_i)_{i \in \Lambda_N}$. The law of u is explicitly given in [13], Theorem 2. We consider wired boundary conditions obtained by adding a "wiring point" ρ to Λ_N, setting $u_\rho := 0$, and connecting any vertex $i \in \Lambda_N$ to ρ with the "wiring weight", synonym "pinning strength",

$$h_i = W_{i\rho} = \sum_{\substack{j \in \Lambda \setminus \Lambda_N: \\ \{i,j\} \in E_\Lambda}} W_{ij}, \tag{1.2}$$

provided that this sum is finite. The corresponding "pinning conductances" are given by $c_{i\rho} = W_{i\rho} e^{u_i + u_\rho} = h_i e^{u_i}$. Leaving Λ_N corresponds to jumping to the wiring point ρ. Let $E_{\Lambda_N} = \{e \in E_\Lambda : e \subseteq \Lambda_N\}$ denote the edge set restricted to Λ_N. The edge set on the extended graph with vertex set $\Lambda_N \cup \{\rho\}$ is

$$E_N = E_{\Lambda_N} \cup \bigl\{\{i, \rho\} : i \in \Lambda_N \text{ and } \{i, j\} \in E_\Lambda \text{ for some } j \in \Lambda \setminus \Lambda_N\bigr\}. \tag{1.3}$$

With these notations, the law $\nu_{W,h}^{\Lambda_N}$ of the random environment u for the VRJP starting in the wiring point ρ is described by

$$\begin{aligned} \nu_{W,h}^{\Lambda_N}(du) &= e^{-\sum_{\{i,j\} \in E_{\Lambda_N}} W_{ij}(\cosh(u_i - u_j) - 1) - \sum_{i \in \Lambda_N} h_i (\cosh u_i - 1)} \\ &\quad \times \sqrt{\sum_{T \in \mathcal{S}} \prod_{\{i,j\} \in T} W_{ij} e^{u_i + u_j}} \prod_{i \in \Lambda_N} \frac{e^{-u_i}}{\sqrt{2\pi}} \, du_i, \end{aligned} \tag{1.4}$$

where $\mathcal{S}$ is the set of spanning trees over the graph $(\Lambda_N \cup \{\rho\}, E_N)$. Let $\mathbb{E}_{W,h}^{\Lambda_N}$ denote the expectation corresponding to $\nu_{W,h}^{\Lambda_N}$.

1.1. *Results on the VRJP with long-range weights in any dimension $d \geq 1$.* In this subsection, we consider the following model.

MODEL 1.1 (d-dimensional lattice with Euclidean long-range interactions). Fix a dimension $d \in \mathbb{N}$ and two parameters $\overline{W} > 0$ and $\alpha > 1$. We consider the d-dimensional lattice $\mathbb{Z}^d$ with edge weights $W_{ij} = W_{ji} = w(\|i - j\|_\infty)$ for $i, j \in \mathbb{Z}^d$, $i \neq j$, where the function $w : [1, \infty) \to (0, \infty)$ is monotonically decreasing and satisfies the lower bound

$$w(x) \geq \overline{W} \frac{(\log_2 x)^\alpha}{x^{2d}} \quad \text{for all } x \geq 1 \tag{1.5}$$

and the summability condition

$$\sum_{i \in \mathbb{Z}^d \setminus \{0\}} w(\|i\|_\infty) < \infty. \tag{1.6}$$

Note that large values of the weights W_{ij} correspond to weak reinforcement, because jumps typically occur after a short time; hence the local time cannot increase a lot before any typical jump. Thus, the lower bound (1.5) can be interpreted as a regime of weak reinforcement. We prove the following transience result.

THEOREM 1.2 (Transience of long-range VRJP). *The discrete time process associated to the VRJP on $\mathbb{Z}^d$ started in 0 with interactions as in Model 1.1 is transient, that is, it visits a.s. any given vertex only finitely often. Slightly stronger, the expected number of visits to any given vertex $x \in \mathbb{Z}^d$ is finite and uniformly bounded in x.*

Discussion. We believe that a logarithmic correction in (1.5) in dimensions $d = 1, 2$ is crucial for this theorem to be valid and not only for its proof. Indeed, for locally finite graphs, Poudevigne's Theorem 4 in [11] shows that the VRJP with weights W is recurrent whenever the corresponding random walk in the *deterministic* conductances W is recurrent. In dimensions $d = 1, 2$, the random walk in conductances $W_{ij} \le \overline{W}\|i - j\|_\infty^{-2d}$ for some $\overline{W} > 0$ is known to be recurrent; see [5] and [3]. Although $\mathbb{Z}^d$ with long-range interactions is not a locally finite graph, one may conjecture that Poudevigne's result could possibly be extended to this case. This would imply recurrence of the corresponding VRJP for $W_{ij} \le \overline{W}\|i - j\|_\infty^{-2d}$.

Strategy of proof. The proof of the transience result Theorem 1.2 relies on an analysis of random walks in random conductances given in Theorem 2.1 below. It requires the transience of the Markovian random walk on $\mathbb{Z}^d$ endowed with the *deterministic* weights W_{ij} from Model 1.1. It is known that for $W_{ij} \ge \overline{W}\|i - j\|_\infty^{-s}$, $i, j \in \mathbb{Z}^d$, with $\overline{W} > 0$, $s < 2d$, the Markovian random walk on the weighted graph is transient. This was shown by Caputo, Faggionato, and Gaudillière [5], Appendix B.1, using characteristic functions and by Bäumler in [3] using electrical networks. In contrast to this, we have a logarithmic correction $W_{ij} \ge \overline{W}(\log\|i - j\|_\infty)^\alpha \|i - j\|_\infty^{-2d}$ with $\alpha > 1$ in Model 1.1. Bäumler's method still applies in this case and is used to prove the following lemma.

LEMMA 1.3 (Transience of long-range Markovian random walk). *Let $w : [1, \infty) \to (0, \infty)$ be a monotonically decreasing function satisfying* (1.5) *and* (1.6) *for some $d \in \mathbb{N}$, $\overline{W} > 0$, and $\alpha > 1$. Then, the Markovian random walk on $\mathbb{Z}^d$ endowed with the long-range deterministic weights $W_{ij} = w(\|i - j\|_\infty)$ for all $i, j \in \mathbb{Z}^d$ is transient.*

To study the VRJP on $\mathbb{Z}^d$ with weights from Model 1.1, we analyze the process on a sequence of finite boxes $\Lambda_N \subset \mathbb{Z}^d$ with wired boundary conditions, given by the pinning

$$0 \le h_i = \sum_{j \in \mathbb{Z}^d \setminus \Lambda_N} W_{ij} < \infty, \quad i \in \Lambda_N. \tag{1.7}$$

Note that the last formula is a special case of formula (1.2). The following bounds on the environment are the key ingredient of our transience proof. They may be interpreted as long-range order in any dimension $d \ge 1$, provided that the long-range weights do not decay too fast.

THEOREM 1.4 (Bounds on the environment—long-range case). *For $N \in \mathbb{N}$, we consider the box $\Lambda_N := \{0, 1, \ldots, 2^N - 1\}^d \subseteq \mathbb{Z}^d$, $d \ge 1$, with the weights inherited from Model* 1.1 *with parameters $\overline{W} > 0$ and $\alpha > 1$. Then, for all $m \ge 1$ and $\sigma \in \{\pm 1\}$, the bound*

$$\mathbb{E}_{W,h}^{\Lambda_N}\left[e^{\sigma m u_i}\right] \le C \tag{1.8}$$

holds with some constant $C = C(\overline{W}, d, \alpha, m) > 0$, uniformly in $i \in \Lambda_N$ and N.

An explicit expression for C is given in formula (3.24).

1.2. *Results on the VRJP with possibly long-range weights in dimension $d \ge 3$.* In this subsection, we consider the following model.

MODEL 1.5 (d-dimensional lattice with large interactions in high dimension $d \ge 3$). Fix a dimension $d \ge 3$ and $\overline{W} > 0$. We consider the d-dimensional lattice $\mathbb{Z}^d$ with possibly long-range edge weights $W_{ij} = W_{ji} \ge 0$ for $i, j \in \mathbb{Z}^d$ with $i \ne j$, which are lower bounded by weights of simple nearest-neighbor random walk on $\mathbb{Z}^d$

$$W_{ij} \ge \overline{W} 1_{\{\|i - j\|_2 = 1\}} \tag{1.9}$$

and fulfill the uniform summability condition

$$\sup_{i\in\mathbb{Z}^d}\sum_{j\in\mathbb{Z}^d\setminus\{i\}} W_{ij}<\infty. \tag{1.10}$$

THEOREM 1.6 (Transience of VRJP in high dimension). *Consider Model 1.5 with parameters $d\geq 3$ and $\overline{W}$. There is a constant $\overline{W}_1=\overline{W}_1(d)>0$ such that the assumption $\overline{W}\geq\overline{W}_1$ implies that the discrete time process associated to the VRJP starting in 0 for this model is transient, that is, it visits a.s. any given vertex only finitely often. Slightly stronger, the expected number of visits to any given vertex x is finite. Moreover, if the weights W_{ij} are translation invariant this expectation is uniformly bounded in x. One can take $\overline{W}_1=\max\{\overline{W}_2,10^8\}$ with $\overline{W}_2$ as in Theorem 1.7.*

Note that on locally finite graphs the probability that the VRJP is transient is an increasing function of the weights W_{ij} (cf. [11], Theorem 1). If this result was applicable also for nonlocally finite graphs, then transience of VRJP as stated in Theorem 1.6 would follow from the transience result on $\mathbb{Z}^d$ for $d\geq 3$ with nearest-neighbor jumps proved in [13]. However, one may conjecture that this obstruction is only a technical issue, not a fundamental failure of the approach.

For $\mathbb{Z}^d$ in dimension $d\geq 3$, we deduce the following bounds for the environment. As in the long-range case, these bounds are the key ingredient to prove transience.

THEOREM 1.7 (Bounds on the environment—high dimension). *We consider any finite box $\Lambda_N\subset\mathbb{Z}^d$ with the weights inherited from Model 1.5. There exists $\overline{W}_2\geq 2^8$ such that for all $\overline{W}\geq\overline{W}_2$, all $m\in[1,\frac{1}{2}\overline{W}^{\frac{1}{8}}]$, and $\sigma\in\{\pm1\}$, one has*

$$\mathbb{E}^{\Lambda_N}_{W,h}[e^{\sigma m u_i}]\leq 2^{2m+1} \tag{1.11}$$

uniformly in $i\in\Lambda_N$ and the box Λ_N.

Note that in the case of a strict inequality $\overline{W}>\overline{W}_2$, we can take $m>1$, which yields uniform integrability of e^{u_i}. This is formulated in the following Lemma 1.8.

1.3. *Uniform integrability.* A potential strategy to prove Theorems 1.2 and 1.6 uses a recurrence/transience criterion given in [14], Theorem 1, for locally finite graphs. Take a sequence of finite sets $\Lambda_N\uparrow\mathbb{Z}^d$ with $0\in\Lambda_N$. One can couple $u_0^{(N)}$ distributed according to $\nu^{\Lambda_N}_{W,h}$ for all N in such a way that $(e^{u_0^{(N)}})_{N\in\mathbb{N}}$ is a nonnegative martingale, which is uniformly integrable by the following lemma.

LEMMA 1.8 (Uniform integrability). *In Model 1.1, let $\overline{W}>0$. In Model 1.5, let $\overline{W}>\overline{W}_2$ with $\overline{W}_2$ from Theorem 1.7. Under these assumptions, the random variables $e^{\pm u_i}$ distributed according to $\nu^{\Lambda_N}_{W,h}$ are uniformly integrable in $N\in\mathbb{N}$ and $i\in\Lambda_N$.*

Hence, the martingale $(e^{u_0^{(N)}})_{N\in\mathbb{N}}$ converges a.s. and in L^1 to a limit ψ_0. The limit satisfies $\mathbb{E}[\psi_0]=\lim_{N\to\infty}\mathbb{E}^{\Lambda_N}_{W,h}[e^{u_0^{(N)}}]=1$, where the last equation follows by supersymmetry; see [9], (B.3). Because of uniform integrability of $e^{-u_0^{(N)}}$, $N\in\mathbb{N}$, the expectation $\mathbb{E}[\psi_0^{-1}]$ is finite and consequently the limit ψ_0 is almost surely strictly positive. If one extended [14], Theorem 1, to $\mathbb{Z}^d$ with long-range interactions, this would yield transience for the VRJP in an alternative way to the approach in this paper.

How this paper is organized. In Section 2.1 we show how transience of the VRJP follows from the bounds on the random environment given in Theorems 1.4 and 1.7. Uniform integrability from Lemma 1.8 is shown in Section 2.2. In Section 3.1 we analyse an auxiliary model with hierarchical interactions. The corresponding bounds are obtained by reducing the model to an effective one with vertices that represent length scales. Although this effective model is still defined on a complete graph, only interactions between neighboring length scales are relevant, which allows us to treat it essentially like an inhomogeneous one-dimensional model. In Section 3.2, we show how bounds on the corresponding environment can be transferred to bounds on the environment for the Euclidean lattice in any dimension $d \geq 1$ with long-range interactions. In Section 4.1 we introduce the $H^{2|2}$ model, which describes the VRJP random environment as a marginal. Section 4.2 collects some identities generated by supersymmetry. These are applied to deduce a bound on fluctuations of the random environment. Finally, in Section 4.3, we prove the bounds on the environment in dimension $d \geq 3$ from Theorem 1.7 by combining the bound from Section 4.2 with estimates from [9] and a monotonicity result from [11].

The constants $\overline{W}_1$, $\overline{W}_2$, and C keep their meaning throughout the whole paper.

2. Proof of the main results.

2.1. *Transience proof given bounds on the environment.* Our transience proof uses the connection between random walks and electrical networks; cf. [10] for background. First we prove transience for Markovian random walks with long-range jumps.

PROOF OF LEMMA 1.3. By [10], Theorem 2.11, to prove transience, it suffices to construct a unit flow of finite energy from 0 to infinity. Following the same strategy as in the proof of [3], Theorem 1.1, we consider the pairwise disjoint annuli $B_k = ((-2^k, 2^k]^d \setminus (-2^{k-1}, 2^{k-1}]^d) \cap \mathbb{Z}^d$ for $k \in \mathbb{N}$ and $B_0 = \{0\}^d$. We observe $|B_k| \geq 2^{kd}$ for $k \in \mathbb{N}_0$. Note that for $i \in B_k$, $j \in B_{k+1}$, we have $\|i - j\|_\infty \leq 2^{k+2}$ and since w is decreasing and satisfies the lower bound (1.5), it holds

$$W_{ij} = w(\|i - j\|_\infty) \geq w(2^{k+2}) \geq \overline{W} \frac{(k+2)^\alpha}{2^{2d(k+2)}} \geq \overline{W} 2^{-3d} \frac{(k+2)^\alpha}{|B_k||B_{k+1}|}. \tag{2.1}$$

Given $i, j \in \mathbb{Z}^d$, we define $\theta(i, j) = -\theta(j, i) = |B_k|^{-1}|B_{k+1}|^{-1}$ if $i \in B_k$, $j \in B_{k+1}$ for some $k \in \mathbb{N}_0$, and $\theta(i, j) = 0$ otherwise. Then, θ is a unit flow from 0 to infinity, that is, Kirchhoff's node rule $\sum_{j \in \mathbb{Z}^d} \theta(i, j) = \delta_{0i}$ holds for all $i \in \mathbb{Z}^d$. The corresponding energy of the network with conductances W is bounded as follows:

$$\frac{1}{2} \sum_{i,j \in \mathbb{Z}^d} \frac{\theta(i, j)^2}{W_{ij}} = \sum_{k \in \mathbb{N}_0} \sum_{i \in B_k} \sum_{j \in B_{k+1}} \frac{|B_k|^{-2}|B_{k+1}|^{-2}}{W_{ij}} \leq \frac{2^{3d}}{\overline{W}} \sum_{k \in \mathbb{N}_0} (k+2)^{-\alpha} < \infty \tag{2.2}$$

because $\alpha > 1$. □

To prove transience for VRJP, we apply the following criterion for random walks in random conductances. The idea for its proof is due to Christophe Sabot and Pierre Tarrès (proof of [13], Corollary 4, and private communication).

THEOREM 2.1 (Transience criterion). *Consider an undirected connected graph $G = (\Lambda, E_\Lambda)$, not necessarily of finite degree, a vertex $0 \in \Lambda$, and an increasing sequence $\Lambda_N \uparrow \Lambda$ of finite vertex sets containing 0. Let $G_N = (\Lambda_N \cup \{\rho\}, E_N)$, where ρ is an additional wiring point and G_N is obtained from G as follows. We restrict the vertex set to Λ_N, add ρ, and use wired boundary conditions to define the edge set E_N as in (1.3).*

Consider a stochastic process $X = (X_n)_{n\in\mathbb{N}_0}$ on $G = (\Lambda, E_\Lambda)$ starting in 0 and taking only nearest-neighbor jumps. We assume that for each N, the law of the process before it hits $\Lambda \setminus \Lambda_N$ equals the annealed law of a random walk in some random conductances $c_e = c_e^{(N)} > 0$, $e \in E_N$, on G_N before it hits ρ. The law of $(c_e^{(N)})_{e\in E_N}$ may depend on N. We denote the expectation averaging over the conductances by $\mathbb{E}^N$. Let $x \in \Lambda$. Assume that there are deterministic conductances $W_e > 0$, $e \in E_\Lambda$, such that the following hold:

1. *For all $i \in \Lambda$, one has $\sum_{e\in E_\Lambda : i\in e} W_e < \infty$.*
2. *The random walk on the weighted graph (G, W) is transient.*
3. *There is a constant $K_x > 0$ such that for all $N \in \mathbb{N}$ and all $e \in E_N$*

$$\sum_{f\in E_N : x\in f} \mathbb{E}^N\left[\frac{c_f}{c_e}\right] \le \frac{K_x}{W_e}, \tag{2.3}$$

where $W_{i\rho} = W_{i\rho}^{(N)}$ is defined via the wired boundary conditions as in (1.2).

Let $\mathcal{R}^N(W, x \leftrightarrow \rho)$ denote the effective resistance between x and ρ in the network G_N with deterministic conductances $W = (W_e)_{e\in E_N}$. Then the expected number of visits in x of X before exiting Λ_N is bounded by $K_x \mathcal{R}^N(W, x \leftrightarrow \rho)$, which is bounded uniformly in N. As a consequence, the expected number of visits of x by the process X on G is finite with the same bound.

PROOF. Let $E^{\mathrm{rw}}_{\Lambda,0}$ denote the expectation with respect to the process X, which starts in 0, on G. For any $y \in \Lambda_N$, let $P^N_{c,y}$ and $E^N_{c,y}$ be the probability measure and expectation, respectively, underlying the Markovian random walk X starting in y on G_N in given conductances c. Let τ_x^+ denote the first return time to x, and τ_x, τ_ρ the hitting time of x, ρ, respectively. Let $N_x := \sum_{n=0}^\infty 1_{\{X_n = x, n<\tau_\rho\}}$ denote the number of visits to x before visiting ρ. The expected number of visits of X to x equals

$$E^{\mathrm{rw}}_{\Lambda,0}\left[\sum_{n=0}^\infty 1_{\{X_n = x\}}\right] = \lim_{N\to\infty} \mathbb{E}^N[E^N_{c,0}[N_x]]. \tag{2.4}$$

Let N be large enough that $x \in \Lambda_N$. Conditionally on $\tau_x < \tau_\rho$, since $c_e > 0$ for all $e \in E_N$, the number of visits to x before hitting ρ in fixed conductances c is geometric with mean

$$E^N_{c,x}[N_x] = P^N_{c,x}(\tau_\rho < \tau_x^+)^{-1} = c(x)\mathcal{R}^N(c, x \leftrightarrow \rho), \tag{2.5}$$

where $c(x) = \sum_{f\in E_N : x\in f} c_f$ and $\mathcal{R}^N(c, x \leftrightarrow \rho)$ denotes the effective resistance between x and ρ in the network G_N with conductances c; the expression of the escape probability in terms of the effective resistance is described in [10], formula (2.5). We obtain

$$E^N_{c,0}[N_x] = P^N_{c,0}(\tau_x < \tau_\rho)E^N_{c,x}[N_x] \le c(x)\mathcal{R}^N(c, x \leftrightarrow \rho). \tag{2.6}$$

By [10], formula (2.7), there is a unit flow $\theta = (\theta(i,j))_{i,j\in\Lambda_N : \{i,j\}\in E_N}$ from x to ρ in this network, that is, $\theta(i,j) = -\theta(j,i)$, $\sum_{j\in\Lambda_N : \{i,j\}\in E_N} \theta(i,j) = \delta_{ix} - \delta_{i\rho}$, such that the effective resistance in the network G_N with deterministic conductances W is given by

$$\mathcal{R}^N(W, x \leftrightarrow \rho) = \sum_{\{i,j\}\in E_N} W_{ij}^{-1}\theta(i,j)^2. \tag{2.7}$$

By Thomson's principle [10], Section 2.4, with the same flow θ, we obtain an upper bound for the effective resistance with random conductances

$$\mathcal{R}^N(c, x \leftrightarrow \rho) \le \sum_{\{i,j\}\in E_N} c_{ij}^{-1}\theta(i,j)^2. \tag{2.8}$$

Using first (2.6) and then assumption (2.3), we obtain the following bound for the expected number of visits in x before exiting Λ_N:

$$\begin{aligned}\mathbb{E}^N[E^N_{c,0}[N_x]] &\le \mathbb{E}^N[c(x)\mathcal{R}^N(c, x\leftrightarrow\rho)]\\ &\le \sum_{\{i,j\}\in E_N}\sum_{\substack{f\in E_N:\\ x\in f}}\mathbb{E}^N\left[\frac{c_f}{c_{ij}}\right]\theta(i,j)^2 \le K_x\sum_{\{i,j\}\in E_N}W_{ij}^{-1}\theta(i,j)^2\\ &= K_x\mathcal{R}^N(W, x\leftrightarrow\rho)\uparrow_{N\to\infty}K_x\mathcal{R}(W, x\leftrightarrow\infty)\end{aligned}\tag{2.9}$$

with a finite limit $\mathcal{R}(W, x\leftrightarrow\infty)$ because of the transience assumption. In particular, the estimate in (2.9) is uniform in N. Finally, (2.4) allows us to conclude. □

Using the bounds in Theorems 1.4 and 1.7, we apply the criterion from Theorem 2.1 to deduce transience of VRJP.

PROOF OF THEOREM 1.2. We use Model 1.1 on $\mathbb{Z}^d$ with the weights W_{ij}, $i, j\in\mathbb{Z}^d$, and the boxes $\Lambda_N=\{0,\dots,2^N-1\}^d$ replaced by shifted boxes $\tilde\Lambda_N=\Lambda_N+(a_N,\dots,a_N)$ with $a_N\in\mathbb{Z}$ suitably chosen such that $\tilde\Lambda_N\uparrow\mathbb{Z}^d$ and $0\in\tilde\Lambda_N$. We take the graphs $G_N=(\tilde\Lambda_N\cup\{\rho\}, E_N)$ with wired boundary conditions as specified in Theorem 2.1. Let $W_{i\rho}=W_{i\rho}^{(N)}$ be as in formula (1.2). By [13], Theorem 2, the discrete-time process associated with the VRJP on G_N starting at ρ has the same distribution as the random walk in random conductances $c_{ij}=W_{ij}e^{u_i+u_j}$, $i, j\in\tilde\Lambda_N\cup\{\rho\}$, on G_N, where u_i, $i\in\tilde\Lambda_N$, are jointly distributed according to the law $\nu_{W,h}^{\tilde\Lambda_N}$ given in (1.4). Changing the starting point from ρ to 0, one may take the same random conductances $W_{ij}e^{u_i+u_j}$, but with respect to the modified measure $e^{u_0}\,d\nu_{W,h}^{\tilde\Lambda_N}$ having the density $e^{u_0-u_\rho}=e^{u_0}$ with respect to $\nu_{W,h}^{\tilde\Lambda_N}$. This can be seen from formula (1.4) using the shift $\tilde u_i=u_i-u_0$ and noting that a factor e^{-u_0} is replaced by e^{-u_ρ} when changing the product $\prod_{i\in\Lambda_N}e^{-u_i}$ to $\prod_{i\in\Lambda_N\cup\{\rho\}\setminus\{0\}}e^{-u_i}$.

We verify the Assumptions 1–3 of Theorem 2.1. Assumption 1 holds by hypothesis (1.6). Assumption 2 is true by Lemma 1.3. Finally, Assumption 3 is verified as follows. Let $N\in\mathbb{N}$. The expectation $\mathbb{E}^N$ in Theorem 2.1 is with respect to the environment for the VRJP starting in 0, not in ρ. Hence, given $x\in\Lambda_N$ and $e=\{i,j\}\in E_N$, it remains to estimate

$$\sum_{\substack{f\in E_N:\\ x\in f}}\mathbb{E}^N\left[\frac{c_f}{c_e}\right]=\sum_{\substack{f\in E_N:\\ x\in f}}\mathbb{E}_{W,h}^{\tilde\Lambda_N}\left[\frac{c_f}{c_e}e^{u_0}\right]=\sum_{k\in(\tilde\Lambda_N\cup\{\rho\})\setminus\{x\}}\frac{W_{xk}}{W_{ij}}\mathbb{E}_{W,h}^{\tilde\Lambda_N}\left[\frac{e^{u_x+u_k}}{e^{u_i+u_j}}e^{u_0}\right].\tag{2.10}$$

We apply first Hölder's inequality and then the bound (1.8) in Theorem 1.4 to obtain

$$\begin{aligned}&\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{u_x+u_k-u_i-u_j+u_0}]\\ &\quad\le\left(\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{5u_x}]\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{5u_k}]\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{-5u_i}]\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{-5u_j}]\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{5u_0}]\right)^{\frac15}\le C\end{aligned}\tag{2.11}$$

with the constant C from there; for $k=\rho$, note that $\mathbb{E}_{W,h}^{\tilde\Lambda_N}[e^{5u_\rho}]=1\le C$. Note that the theorem is applied to the translated box $\tilde\Lambda_N$ rather than the original box Λ_N, using translational invariance of the weights. Substituting the bound (2.11) in (2.10), we obtain

$$\sum_{f\in E_N:x\in f}\mathbb{E}^N\left[\frac{c_f}{c_e}\right]\le C\sum_{k\in(\tilde\Lambda_N\cup\{\rho\})\setminus\{x\}}\frac{W_{xk}}{W_{ij}}=\frac{K}{W_{ij}},\tag{2.12}$$

with $K = C\sum_{k\in\mathbb{Z}^d\setminus\{x\}} W_{xk} < \infty$; by translation invariance, the constant K does not depend on x. Having checked all its hypotheses, we apply Theorem 2.1. The limit $\mathcal{R}(W, x \leftrightarrow \infty) = \lim_{N\to\infty} \mathcal{R}^N(W, x \leftrightarrow \rho)$ does not depend on the choice of the boxes $\tilde{\Lambda}_N \uparrow \mathbb{Z}^d$. Finally, since the weights W_{ij} are translation invariant, it also does not depend on x. We conclude that the expected number of visits to x is uniformly bounded by $K\mathcal{R}(W, 0 \leftrightarrow \infty)$. □

PROOF OF THEOREM 1.6. The proof is parallel to the proof of Theorem 1.2, but with some modifications. We explain these modifications. This time, we use Model 1.5 and take the shifted box $\tilde{\Lambda}_N$ again, but keep only edges $\{i, j\}$ with $W_{ij} > 0$. We apply Theorem 2.1 in the same way as before. This time, we verify its three assumptions as follows. Assumption 1 holds by hypothesis (1.10). Assumption 2 follows from Rayleigh's monotonicity principle [10], Section 2.4, and transience of nearest-neighbor simple random walk in dimensions $d \ge 3$. Finally, Assumption 3 is verified using Theorem 1.7 as follows. We assume $\overline{W} \ge \overline{W}_1$ with $\overline{W}_1 = \max\{\overline{W}_2, 10^8\}$ and $\overline{W}_2$ as in Theorem 1.7, which implies $\frac{1}{2}\overline{W}^{1/8} \ge 5 =: m$. The steps leading to estimate (2.10) and (2.11) remain unchanged, with the exception that the constant C is replaced by $2^{2m+1} = 2^{11}$; cf. (1.11). The steps leading to (2.12) remain unchanged, with the exception that this time the constant $K_x := 2^{11}\sum_{k\in\mathbb{Z}^d\setminus\{x\}} W_{xk} < \infty$ may depend on x. However, it is uniformly bounded in x by (1.10): $K := \sup_{x\in\mathbb{Z}_d} K_x < \infty$. By the same argument as in the proof of Theorem 1.2, we conclude that the expected number of visits to x is bounded by $K\mathcal{R}(W, x \leftrightarrow \infty)$. If the weights W_{ij} are translation invariant, this bound does not depend on x. □

2.2. *Uniform integrability and monotonicity.* Lemma 1.8 is a consequence of Theorems 1.4 and 1.7, as the following proof shows.

PROOF OF LEMMA 1.8. Note that for $\sigma \in \{\pm 1\}$ and some $m > 1$ one has $\mathbb{E}^{\Lambda_N}_{W,h}[e^{\sigma m u_i}] \le \max\{C, 2^{2m+1}\}$ for all boxes $\Lambda_N \subset \mathbb{Z}^d$ of side length 2^N and $i \in \Lambda_N$. This follows from Theorem 1.4 together with translation invariance in the case of Model 1.1 and from Theorem 1.7 in case of Model 1.5. Given an arbitrary finite set $\Lambda' \subset \mathbb{Z}^d$ there exists a box Λ_N of side length 2^N with $\Lambda' \subseteq \Lambda_N$ for N large enough. We endow both of them with wired boundary conditions. Given $i \in \Lambda'$, according to [14], Theorem 1, one can couple $u_i^{(\Lambda')}$ and $u_i^{(\Lambda_N)}$ such that $e^{u_i^{(\Lambda')}}, e^{u_i^{(\Lambda_N)}}$ is a one-step martingale. Note that e^u is called ψ in [14]. Since $x \mapsto x^{\sigma m}$ is convex for $m > 1$, Jensen's inequality gives $\mathbb{E}^{\Lambda'}_{W,h}[e^{\sigma m u_i}] \le \mathbb{E}^{\Lambda_N}_{W,h}[e^{\sigma m u_i}] \le \max\{C, 2^{2m+1}\}$ and the claim follows. □

We need the following monotonicity result for the environment in the weights.

FACT 2.2 (Poudevigne's monotonicity theorem [11], Theorem 6). Consider two families of nonnegative edge weights and pinnings W^+, h^+ and W^-, h^- on the same vertex set Λ, satisfying $W^+ \ge W^-$ and $h^+ \ge h^-$ componentwise. For any convex function $f : [0, \infty) \to \mathbb{R}$ and all $i \in \Lambda$, one has

$$\mathbb{E}^{\Lambda}_{W^+,h^+}\big[f(e^{u_i})\big] \le \mathbb{E}^{\Lambda}_{W^-,h^-}\big[f(e^{u_i})\big]. \tag{2.13}$$

In particular, for $m \ge 1$ and $\sigma \in \{\pm 1\}$,

$$\mathbb{E}^{\Lambda}_{W^+,h^+}\big[e^{\sigma m u_i}\big] \le \mathbb{E}^{\Lambda}_{W^-,h^-}\big[e^{\sigma m u_i}\big]. \tag{2.14}$$

Recall that ρ denotes the wiring point. In Poudevigne's notation, the joint law of $G^{\pm}(\rho, i)/G^{\pm}(\rho, \rho)$, $i \in \Lambda$, from [11], Theorem 6, coincides with the joint law of e^{u_i}, $i \in \Lambda$, with respect to $\nu^{\Lambda}_{W^{\pm},h^{\pm}}$. Taking the convex function $f(x) = e^{\sigma m u}$, $m \ge 1$, yields (2.14). Using this monotonicity, we derive bounds for the environment in the coming sections.

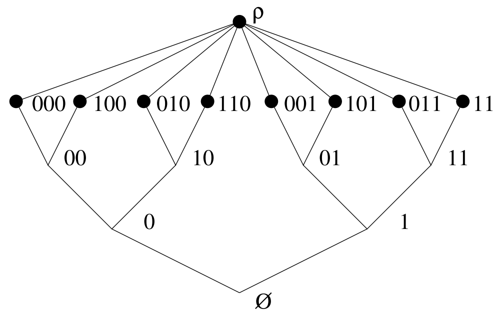

FIG. 1. *Hierarchical graph for $N = 3$ together with the pinning point. In Model 3.1 any two leaves of the binary tree have a positive interaction.*

3. Bounding environments for hierarchical and long-range weights.

3.1. *Hierarchical weights.* To prove Theorem 1.4 we compare a box in the Euclidean lattice with long range interactions with the corresponding box in a hierarchical lattice as follows.

MODEL 3.1 (Complete graph with hierarchical interactions). Fix $N \in \mathbb{N}$. We consider the vertex set consisting of the set of leaves $\Lambda_N = \{0, 1\}^N$ of the binary tree $\mathcal{T} = \bigcup_{n=0}^{N} \{0, 1\}^n$. For $i = (i_0, \dots, i_{N-1})$, $j = (j_0, \dots, j_{N-1}) \in \{0, 1\}^N$, the hierarchical distance is defined by $d_H(i, j) = \min(\{l \in \{0, \dots, N-1\} : i_k = j_k \text{ for all } k \geq l\} \cup \{N\})$; it is the distance to the least common ancestor in $\mathcal{T}$. We endow $\mathcal{T}$ with hierarchical weights

$$W_{ij}^H = w^H\big(d_H(i, j)\big), \quad i, j \in \Lambda_N, i \neq j, \tag{3.1}$$

with a weight function $w^H : \mathbb{N} \to (0, \infty)$ and uniform pinning $W_{i\rho} = h^H > 0$ for all $i \in \Lambda_N$.

An illustration of $\mathcal{T}$ is given in Figure 1.

THEOREM 3.2 (Bounds on the environment—hierarchical case). *Let $\overline{W}^H > 0$ and $\alpha > 1$. Consider Model 3.1 and assume that the weight function w^H and the pinning h^H satisfy*

$$w^H(l) \geq \overline{W}^H 2^{-2l} l^\alpha \quad \textit{for } l \in \mathbb{N}, \qquad h^H \geq \overline{W}^H 2^{-(2+N)} (N+1)^\alpha. \tag{3.2}$$

Then, for all $i \in \Lambda_N$, $\sigma \in \{\pm 1\}$, and all $m \geq 1$, one has

$$\mathbb{E}_{W^H, h^H}^{\Lambda_N}\big[e^{\sigma m u_i}\big] \leq c_H \tag{3.3}$$

with a constant $c_H = c_H(\overline{W}^H, \alpha, m)$, uniformly in N. An explicit form of c_H is given in (3.12), below.

The proof of this theorem is based on the following general lemma, which holds on any finite graph. The underlying strategy has been suggested by an anonymous Associate Editor.

LEMMA 3.3. *Let $G = (\Lambda, E_\Lambda)$ be a finite undirected connected graph endowed with positive weights $W = (W_e)_{e \in E_\Lambda}$ on the edges. Fix a pinning point $\rho \in \Lambda$. We write $\mathbb{E}_W^{\Lambda,\rho}$ for the expectation with respect to the probability measure $\nu_{\tilde{W},h}^{\Lambda \setminus \{\rho\}}$ (cf. (1.4)), with weights $\tilde{W}_{ij} = W_{ij}$ for edges $\{i, j\} \subseteq \Lambda$ and pinnings $h_i = W_{i\rho}$ if $\{i, \rho\}$ is an edge and $h_i = 0$ otherwise.*

Let $\pi = (i = \pi_1, \dots, \pi_{N+1} = \rho)$ be a self-avoiding path from some vertex $i \in \Lambda$ to ρ. Then, for all $m \geq 1$ and $\sigma \in \{\pm 1\}$, one has

$$\mathbb{E}_W^{\Lambda,\rho}[e^{\sigma m u_i}] \leq \exp\left(\frac{(2m+1)^2}{8} \sum_{k=1}^N W_{\pi_k \pi_{k+1}}^{-1}\right). \tag{3.4}$$

PROOF. Let $T \subseteq E_\Lambda$ be a spanning tree of G such that all edges $\{\pi_k, \pi_{k+1}\}$ of the path π are contained in T. Define $W^- = (W_e^-)_{e\in E_\Lambda}$ by $W_e^- := 1_{\{e\in T\}} W_e$. By Fact 2.2, one has

$$\mathbb{E}_W^{\Lambda,\rho}[e^{\sigma m u_i}] \leq \mathbb{E}_{W^-}^{\Lambda,\rho}[e^{\sigma m u_i}]. \tag{3.5}$$

Because π is a path in the tree T, the increments $u_{\pi_k} - u_{\pi_{k+1}}$, $1 \leq k \leq N$, along the path are independent with respect to $\nu_{W^-}^{\Lambda,\rho}$ and distributed according to

$$\nu_{W^-}^{\Lambda,\rho}(u_{\pi_k} - u_{\pi_{k+1}} \in A) = \sqrt{\frac{W_{\pi_k\pi_{k+1}}}{2\pi}} \int_A e^{-\frac{1}{2}u} e^{-W_{\pi_k\pi_{k+1}}[\cosh u - 1]}\, du \tag{3.6}$$

for any Borel set $A \subseteq \mathbb{R}$. Consequently, for $\sigma \in \{\pm 1\}$, using $u_i = u_{\pi_1}$ and $u_{\pi_{N+1}} = u_\rho = 0$, we obtain

$$\mathbb{E}_{W^-}^{\Lambda,\rho}[e^{\sigma m u_i}] = \prod_{k=1}^N \mathbb{E}_{W^-}^{\Lambda,\rho}[e^{\sigma m (u_{\pi_k} - u_{\pi_{k+1}})}]. \tag{3.7}$$

For any $k \in \{1, \dots, N\}$, abbreviating $w := W_{\pi_k\pi_{k+1}}$, we estimate

$$\begin{aligned} \mathbb{E}_{W^-}^{\Lambda,\rho}[e^{\sigma m (u_{\pi_k} - u_{\pi_{k+1}})}] &= \sqrt{\frac{w}{2\pi}} \int_{\mathbb{R}} e^{\sigma m u} e^{-\frac{1}{2}u} e^{-w[\cosh u - 1]}\, du \\ &\leq \sqrt{\frac{w}{2\pi}} \int_{\mathbb{R}} e^{\sigma m u} e^{-\frac{1}{2}u} e^{-\frac{1}{2}wu^2}\, du = \exp\frac{(2\sigma m - 1)^2}{8w} \\ &\leq \exp\frac{(2m+1)^2}{8w}. \end{aligned} \tag{3.8}$$

Inserting this in (3.7) and the result in (3.5), the claim follows. □

Next, we apply this lemma to deduce the bounds for the hierarchical model, using an antichain in the tree $\mathcal{T}$.

PROOF OF THEOREM 3.2. By uniform pinning and the symmetry of the weights induced by the hierarchical structure, the distribution of u_i is the same for all $i \in \Lambda_N$. Therefore, it is enough to prove (3.3) for the vertex $i_1 := 0^N := (0, \dots, 0)$. Corollary 2.3 of [8] tells us that the distribution of $e^{u_{i_1}}$ in this model coincides with the distribution of $e^{u_{i_1}}$ in the corresponding effective model on the complete graph with vertex set $A = \{j, i_1, \dots, i_N\} \subseteq \mathcal{T}$, where $j := (1, 0^{N-1}) = (1, 0, \dots, 0) \in \{0,1\}^N$ and $i_l := (1, 0^{N-l}) \in \{0,1\}^{N-l+1}$ for $2 \leq l \leq N$. This is a maximal antichain in the language of Section 2.3 of [8] (cf. Figure 2). Let us review this effective model. Weights and pinning on the antichain, viewed as a complete graph, are given in [8], (2.13) and (2.14). However, only the values of the following nearest-neighbor weights and pinnings are relevant for the subsequent argument, while the values of all other weights do not play any role: $W_{i_{l-1}, i_l} = 2^{(l-2)+(l-1)} w^H(l) = 2^{2l-3} w^H(l)$, $h_{i_l} := 2^{l-1} h^H$, $2 \leq l \leq N$, $h_{i_1} := h^H$. Here, 2^{l-1} equals the number of leaves in $\mathcal{T}$ above the vertex $i_l = (1, 0^{N-l})$. Corollary 2.3 of [8] yields

$$\mathbb{E}_{W^H, h^H}^{\Lambda_N}[e^{m\sigma u_{i_1}}] = \mathbb{E}_{W,h}^{A}[e^{m\sigma u_{i_1}}]. \tag{3.9}$$

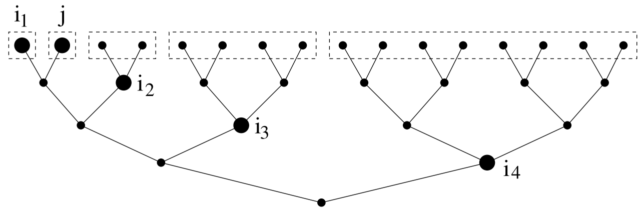

FIG. 2. *The maximal antichain A.*

We regard the pinning point ρ as an additional vertex at the end of the antichain $j, i_1, \dots, i_N$, writing $i_{N+1} := \rho$. By assumption (3.2), one has

$$W_{i_{l-1}i_l} \geq \frac{1}{8}\overline{W}^H l^\alpha \tag{3.10}$$

for $l \in \{2, \dots, N+1\}$. Indeed, in the case $l \leq N$, this follows from $W_{i_{l-1}i_l} = 2^{2l-3}w^H(l) \geq \frac{1}{8}\overline{W}^H l^\alpha$, and in the case $l = N+1$ from

$$W_{i_N,i_{N+1}} = h_{i_N} = 2^{N-1}h^H \geq \frac{1}{8}\overline{W}^H (N+1)^\alpha. \tag{3.11}$$

We remark that the assumption on h^H in (3.2) is chosen such that the pinning case $l = N+1$ gives a lower bound of the same form as the other cases $l \leq N$. Hence, we can apply Lemma 3.3 to the complete graph with vertex set $\{j, i_1, \dots, i_{N+1}\}$ and the path $\pi = (i_1, \dots, i_{N+1})$ to obtain

$$\mathbb{E}^A_{W,h}[e^{m\sigma u_{i_1}}] \leq \exp\left(\frac{(2m+1)^2}{\overline{W}^H}\sum_{l=2}^\infty l^{-\alpha}\right) =: c_H = c_H(\overline{W}^H, \alpha, m). \tag{3.12}$$

□

3.2. *Long-range weights*. Fix an arbitrary dimension $d \geq 1$ and $N \in \mathbb{N}$. Take the box $\Lambda_N := \{0, 1, \dots, 2^N - 1\}^d$. In order to compare Λ_N with $\{0,1\}^{Nd}$, we define (cf. Figure 3)

$$\phi : \Lambda_N \to \{0,1\}^{Nd}, \quad \phi(i_0, \dots, i_{d-1}) = (z_0, \dots, z_{Nd-1}), \tag{3.13}$$

where $z_n \in \{0,1\}$ denotes the $\lfloor n/d \rfloor$th digit in the binary expansion $i_l = \sum_{k=0}^{N-1} z_{dk+l}2^k$, with $l = n \bmod d$. The map ϕ is a bijection. The elements of $\{0,1\}^{Nd}$ can be visualized as the leaves of a binary tree. On this set, the hierarchical distance, that is, the distance to the least common ancestor in the binary tree, is given by

$$d_H(z, z') = \min(\{n \in \{0, \dots, Nd-1\} : z_m = z'_m \text{ for all } m \geq n\} \cup \{Nd\}). \tag{3.14}$$

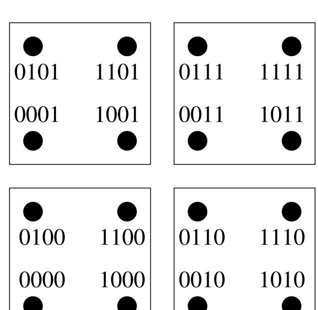

FIG. 3. *Illustration of the map ϕ for $N = 2$ and $d = 2$.*

On Λ_N, the natural definition of hierarchical distance would be $\lceil d_H(\phi(i),\phi(j))/d\rceil$. With this choice, in Figure 3, any two points in the same small box would have distance 1 from each other, while points in different boxes would have distance 2. In the following lemma, we compare the maximum norm distance with the hierarchical distance.

LEMMA 3.4 (Comparison of d_H and $\|\cdot\|_\infty$). *For all $i, j \in \Lambda_N$, one has*

$$2\cdot 2^{d_H(\phi(i),\phi(j))/d} \geq 2^{\lceil d_H(\phi(i),\phi(j))/d\rceil} > \|i-j\|_\infty. \tag{3.15}$$

PROOF. For $i = j$, the claim is true. We take $i \neq j$ in Λ_N and set $z = \phi(i)$, $z' = \phi(j)$, and $r = d_H(z, z')$. One has

$$\|i-j\|_\infty = \max_{0\leq l\leq d-1} |i_l - j_l| = \max_{0\leq l\leq d-1}\left|\sum_{k=0}^{N-1}(z_{dk+l} - z'_{dk+l})2^k\right|. \tag{3.16}$$

The only nonzero contributions in the above sum come at most from k satisfying $dk + l \leq r - 1$ which implies $k \leq (r-1)/d$. It follows that

$$\begin{aligned} \|i-j\|_\infty &\leq \max_{0\leq l\leq d-1} \sum_{k=0}^{\lfloor (r-1)/d\rfloor} |z_{dk+l} - z'_{dk+l}| 2^k \\ &\leq \sum_{k=0}^{\lfloor (r-1)/d\rfloor} 2^k < 2^{\lfloor (r-1)/d\rfloor + 1} = 2^{\lceil r/d\rceil}, \end{aligned} \tag{3.17}$$

where we use $|z_{dk+l} - z'_{dk+l}| \leq 1$ and $r \in \mathbb{N}_0$. $\square$

Next, we show how the bound on the environment for the Euclidean box with long-range interactions follows from Theorem 3.2.

PROOF OF THEOREM 1.4. For $m \geq 1$, we apply Fact 2.2 to compare the Euclidean long-range model, that is, Model 1.1, but restricted to the box Λ_N with wired boundary conditions, with a hierarchical model defined on the same box. Remember that in the Euclidean model, we take long-range weights $W^+_{ij} := W_{ij} = w(\|i-j\|_\infty)$ as in Model 1.1 and pinning $h^+_j := h_j$ given in (1.7). In the hierarchical model, we take weights

$$W^-_{ij} := W^H_{ij} = w^H\big(d_H(\phi(i),\phi(j))\big) \quad \text{with } w^H(l) := w\big(2^{\lceil l/d\rceil}\big) \tag{3.18}$$

and uniform pinning given by $h^- := h^H := \min_{j\in\Lambda_N} h_j$. By Lemma 3.4 and monotonicity of w,

$$W_{ij} = w\big(\|i-j\|_\infty\big) \geq w\big(2^{\lceil d_H(\phi(i),\phi(j))/d\rceil}\big) = W^H_{ij}, \quad h_i \geq h^H, \tag{3.19}$$

for $i, j \in \Lambda_N$. By monotonicity (2.14), for all $i \in \Lambda_N$, $\sigma \in \{\pm 1\}$, and $m \geq 1$, one has

$$\mathbb{E}^{\Lambda_N}_{W,h}\big[e^{\sigma m u_i}\big] \leq \mathbb{E}^{\Lambda_N}_{W^H,h^H}\big[e^{\sigma m u_i}\big]. \tag{3.20}$$

We identify Λ_N with $\{0,1\}^{Nd}$ via the bijection $\phi : \Lambda_N \to \{0,1\}^{Nd}$ from (3.13). As a result we obtain Model 3.1 with N replaced by Nd. We apply Theorem 3.2 with $\overline{W}^H = \overline{W} d^{-\alpha} 2^{-2d}$ and

$$w^H(l) = w\big(2^{\lceil l/d\rceil}\big) \geq \overline{W}\frac{(\lceil l/d\rceil)^\alpha}{2^{2d\lceil l/d\rceil}} \geq \overline{W}\frac{(l/d)^\alpha}{2^{2d(1+l/d)}} = \overline{W}^H 2^{-2l} l^\alpha, \tag{3.21}$$

where we used (1.5). We estimate the pinning using that w is decreasing

$$
\begin{aligned}
h^H &= \min_{i\in\Lambda_N} h_i = \min_{i\in\Lambda_N} \sum_{j\in\mathbb{Z}^d\setminus\Lambda_N} w(\|i-j\|_\infty) \\
&\geq \sum_{\substack{j\in\mathbb{Z}^d:\\ 2^N\leq\|j\|_\infty<2^{N+1}}} w(\|j\|_\infty) \geq w(2^{N+1}) \sum_{\substack{j\in\mathbb{Z}^d:\\ 2^N\leq\|j\|_\infty<2^{N+1}}} 1.
\end{aligned}
\tag{3.22}
$$

The last sum has at least $2^{(N+1)d}$ summands, which is a lower bound on the cardinality of the set $\{z\in\mathbb{Z}: 2^N\leq|z|<2^{N+1}\}\times\{j\in\mathbb{Z}^{d-1}: \|j\|_\infty<2^{N+1}\}$. Hence, we obtain

$$
\begin{aligned}
h^H &\geq 2^{(N+1)d} w(2^{N+1}) \geq 2^{(N+1)d}\overline{W}\frac{(N+1)^\alpha}{2^{2(N+1)d}} \\
&\geq \overline{W} d^{-\alpha} 2^{-(2+2d+Nd)}(Nd+1)^\alpha \\
&= \overline{W}^H 2^{-(2+Nd)}(Nd+1)^\alpha.
\end{aligned}
\tag{3.23}
$$

Note that we drop a factor of 2^{2+d} in the last inequality, uniformly in N. However, keeping it would not improve the result. Thus, assumption (3.2) of Theorem 3.2 is satisfied with N replaced by Nd and the estimate (1.8) follows with the constant

$$
\begin{aligned}
C(\overline{W}, d, \alpha, m) &:= c_H(\overline{W} d^{-\alpha} 2^{-2d}, \alpha, m) \\
&= \exp\left(\frac{d^\alpha 2^{2d}(2m+1)^2}{\overline{W}} \sum_{l=2}^\infty l^{-\alpha}\right) > 1.
\end{aligned}
\tag{3.24}
$$

$\square$

Discussion. Note that Theorem 3.2 applies in the special case $W_{ij}\propto 2^{-(d+2)\lceil d_H(\phi(i),\phi(j))/d\rceil}$ in hierarchical dimension $d\geq 3$, which corresponds to $w^H(l)\propto 2^{-(d+2)\lceil l/d\rceil}$. In this case, the Green's function for the Markov jump process on the corresponding infinite hierarchical lattice with jump rates W_{ij} decays as $G(i,j)=\text{const}\cdot 2^{-(d-2)\lceil d_H(\phi(i),\phi(j))/d\rceil}$, $i\neq j$ (see [4], formula (1.3)); note that our $\lceil d_H(\phi(i),\phi(j))/d\rceil$ corresponds to $\log_2|i-j|_H$ in the case $L=2$ in this reference, that is, $G(i,j)=\text{const}/|i-j|_H^{d-2}$. This case is interesting since it mimics the decay of the inverse Laplacian $-\Delta^{-1}$ on the Euclidean lattice in dimensions $d\geq 3$.

4. Bounding environments in high dimension. The proof of Theorem 1.7, which is the goal of this section, is based on the description of the random environment for VRJP in terms of the nonlinear supersymmetric hyperbolic sigma model, $H^{2|2}$ model for short. The key tools are a bound for the expectation of $(\cosh(u_i-u_j))^m$, which is proved in a slightly more general form in formula (4.18) below, and an estimate from [9].

4.1. *The nonlinear supersymmetric hyperbolic sigma model.* The supersymmetric hyperbolic sigma $H^{2|2}$ model was introduced by Zirnbauer [16] in the context of quantum diffusion. It can be seen as a statistical mechanical spin model where the spins take values in the supermanifold $H^{2|2}$. This is the hyperboloid $H^2=\{(x,y,z)\in\mathbb{R}^3: x^2+y^2-z^2=-1, z>0\}$ extended by two Grassmann, that is, anticommuting, variables ξ, η with the constraint $z^2=1+x^2+y^2+2\xi\eta$. For more mathematical background, see [15], the appendix of [8], and [9]. Changing to horospherical coordinates $(x,y,\xi,\eta)\mapsto(u,s,\overline{\psi},\psi)$; cf. [9], equation (2.7), where u, s are real and $\overline{\psi}$, ψ are Grassmann variables, the model can be defined as follows.

We consider a finite undirected graph with vertex set Λ and edge set E_Λ. Every edge $e=\{i,j\}\in E_\Lambda$ is given a positive weight $W_{ij}=W_{ji}=W_e>0$. To every vertex $i\in\Lambda$, we associate a pinning strength $h_i \geq 0$ in such a way that every connected component of the graph contains at least one vertex i with $h_i > 0$. It is convenient to treat pinning and interaction in a unified way. Therefore, we extend the vertex set Λ by a special vertex $\rho \notin \Lambda$ called "pinning point" or "wiring point". In addition, we introduce pinning edges connecting the pinning point with the other vertices. Since we will study the dependence of the field at any pair of points, we will also add edges connecting vertices without direct interaction. Therefore, we consider the undirected *complete* graph $G = (\Lambda \cup \{\rho\}, E)$ without direct loops. Every edge $e \in E$ is given a weight $W_e \geq 0$ as follows. For $e \in E_\Lambda$, $W_e > 0$ is the weight defined above. The weight of pinning edges $e = \{i, \rho\}$, $i \in \Lambda$, is defined to be the pinning strength $W_e = h_i$. All other edges e are given weight $W_e = 0$. The subgraph consisting of edges with positive weights is denoted by

$$G_+ = (\Lambda \cup \{\rho\}, E_+) \quad \text{where } E_+ := \{e \in E : W_e > 0\}. \tag{4.1}$$

Note that E_+ consists of E_Λ and the pinning edges with positive weight, and our assumption on the pinning is equivalent to require that the graph G_+ is connected.

We assign to every vertex i two real-valued variables u_i and s_i and two Grassmann variables $\overline{\psi}_i$ and ψ_i. We set $u_\rho = s_\rho := 0 =: \overline{\psi}_\rho = \psi_\rho$. Following [9], equations (2.9)–(2.10), for all $e = \{i, j\}$ with $i, j \in \Lambda \cup \{\rho\}$ with $i \neq j$, we set

$$B_{ij} = B_e = B_e(u, s) := \cosh(u_i - u_j) + \frac{1}{2}(s_i - s_j)^2 e^{u_i + u_j}, \tag{4.2}$$

$$S_{ij} = S_e = S_e(u, s, \overline{\psi}, \psi) := B_e + (\overline{\psi}_i - \overline{\psi}_j)(\psi_i - \psi_j) e^{u_i + u_j}. \tag{4.3}$$

For each vertex $i \in \Lambda$, let

$$B_i := B_{i\rho} = \cosh u_i + \frac{1}{2} s_i^2 e^{u_i}, \qquad S_i := S_{i\rho} = B_i + \overline{\psi}_i \psi_i e^{u_i}. \tag{4.4}$$

The action of the nonlinear hyperbolic sigma model in horospherical coordinates is given by (cf. [9], equation (2.8) with $z_i = S_i$)

$$A = A(u, s, \overline{\psi}, \psi) := \sum_{e \in E_\Lambda} W_e(S_e - 1) + \sum_{i \in \Lambda} h_i(S_i - 1) = \sum_{e \in E_+} W_e(S_e - 1). \tag{4.5}$$

The superintegration form of the model is given by ([9], equations (2.12)–(2.13))

$$\langle f \rangle = \langle f \rangle_{W,h}^{\Lambda} := \int_{\mathbb{R}^\Lambda \times \mathbb{R}^\Lambda} \prod_{i \in \Lambda} \frac{e^{-u_i}}{2\pi} \, du_i \, ds_i \, \partial_{\overline{\psi}_i} \partial_{\psi_i} f e^{-A} \tag{4.6}$$

for any function of the variables $u_i, s_i, \overline{\psi}_i, \psi_i$ such that the integral exists. For any integrable function $f = f(u, s)$ which does not depend on the Grassmann variables, one has

$$\langle f \rangle = \mathbb{E}[f] = \mathbb{E}_{W,h}^{\Lambda}[f] := \int_{\mathbb{R}^\Lambda \times \mathbb{R}^\Lambda} f(u, s) \mu_{W,h}^{\Lambda}(du \, ds), \tag{4.7}$$

where $\mu_{W,h}^{\Lambda}$ is the probability measure on $\mathbb{R}^\Lambda \times \mathbb{R}^\Lambda$ given by

$$\begin{aligned} \mu(du\, ds) = \mu_{W,h}^{\Lambda}(du\, ds) &:= e^{-\sum_{e \in E_\Lambda} W_e(B_e - 1) - \sum_{i \in \Lambda} h_i(B_i - 1)} \det D \prod_{i \in \Lambda} \frac{e^{-u_i}}{2\pi} \, du_i \, ds_i \\ &= e^{-\sum_{e \in E_+} W_e(B_e - 1)} \det D \prod_{i \in \Lambda} \frac{e^{-u_i}}{2\pi} \, du_i \, ds_i, \end{aligned} \tag{4.8}$$

and $D=D(u)=(D_{ij})_{i,j\in\Lambda}$ is the weighted Laplacian matrix on G_+ with entries

$$D_{ij}=\begin{cases}-W_{ij}e^{u_i+u_j} & \text{if } \{i,j\}\in E_+,\\ \displaystyle\sum_{k\in\Lambda:\{i,k\}\in E_+} W_{ik}e^{u_i+u_k} & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}\tag{4.9}$$

Note that for $i,j\in\Lambda$ the condition $\{i,j\}\in E_+$ is equivalent to $\{i,j\}\in E_\Lambda$, and that the diagonal entries can also be written in the form $D_{ii}=\sum_{k\in\Lambda:\{i,k\}\in E_\Lambda} W_{ik}e^{u_i+u_k}+h_ie^{u_i}$. Our connectedness assumption on the graph G_+ guarantees that the matrix D is positive definite. In particular, it is invertible. The fact that $\mu^\Lambda_{W,h}$ is a probability measure was shown in [9] using supersymmetry. Since the s-variables, conditionally on u, are normally distributed with covariance D^{-1}, and by the matrix-tree theorem $\det D=\sum_{T\in\mathcal{S}}\prod_{\{i,j\}\in T}W_{ij}e^{u_i+u_j}$, the u-marginal of $\mu^\Lambda_{W,h}$ is given by (1.4). In particular, for functions depending only on u, the expectations with respect to $\mu^\Lambda_{W,h}$ and $\nu^\Lambda_{W,h}$ coincide, which allows us to use the same notation $\mathbb{E}^\Lambda_{W,h}$.

4.2. *A Ward identity and some applications*. In this section, $G=(\Lambda\cup\{\rho\},E)$ is the complete graph with weights $W_e\geq 0$, $e\in E$, such that G_+ is connected; cf. formula (4.1). We use the $H^{2|2}$-model with Grassmann variables still present. Note that real functions of even elements of the Grassmann algebra are defined as formal Taylor series in the Grassmann variables, which are nilpotent. For example, for an edge $e=\{i,j\}$, the expression $f(S_e)$ is defined by $f(S_e)=f(B_e+n_e):=f(B_e)+f'(B_e)n_e$ with $n_e=n_{ij}=(\overline{\psi}_i-\overline{\psi}_j)(\psi_i-\psi_j)e^{u_i+u_j}$. This Taylor expansion stops after the first order term because $n_e^2=0$ by the anticommutativity of the Grassmann variables.

We will use the following localization result from [9].

LEMMA 4.1 (Ward identity). *Let $f_e:[1,\infty)\to\mathbb{R}$, $e\in E$, be a family of continuously differentiable functions such that each f_e and its derivative f_e' are bounded by polynomials. Then it holds*

$$\left\langle\prod_{e\in E}f_e(S_e)\right\rangle=\prod_{e\in E}f_e(1).\tag{4.10}$$

PROOF. This is a special case of the localization result in [9], Prop. 2 in the appendix. The key points are that any S_e is annihilated by the supersymmetry operator used in the proposition and $S_{ij}=1$ holds whenever $u_k=s_k=\overline{\psi}_k=\psi_k=0$ for $k\in\{i,j\}$. □

Notation. The following notation will be technically convenient. The two vertices incident to an edge $e\in E$ are denoted by e_+ and e_-. This gives every edge a direction $e_+\to e_-$, which has no physical meaning, but is used for bookkeeping purposes. Let

$$F\in\{-1,0,1\}^{\Lambda\times E},\quad F_{k,e}=1_{\{e_+=k\}}-1_{\{e_-=k\}}\quad\text{for } k\in\Lambda, e\in E\tag{4.11}$$

denote the signed incidence matrix of G with ρ removed, but keeping the pinning edges. We consider the diagonal matrix

$$\mathcal{W}=\mathcal{W}(u):=\operatorname{diag}(\mathcal{W}_e,e\in E)\in\mathbb{R}^{E\times E}\quad\text{with } \mathcal{W}_e:=W_ee^{u_{e_+}+u_{e_-}}.\tag{4.12}$$

With these notations, we reformulate the weighted Laplacian matrix D from (4.9) as

$$D=F\mathcal{W}F^t\in\mathbb{R}^{\Lambda\times\Lambda}.\tag{4.13}$$

We will also need the following matrices:

$$Q = Q^G(u,s) := \operatorname{diag}(Q_e, e \in E) \quad \text{with } Q_e := \frac{e^{u_{e_+}+u_{e_-}}}{B_e(u,s)}, \tag{4.14}$$

$$\mathcal{G} = \mathcal{G}_W^G(u,s) := \sqrt{Q} F^t D^{-1} F \sqrt{Q} \in \mathbb{R}^{E \times E}. \tag{4.15}$$

Let $m_e \geq 0$, $e \in E$. We set

$$M := \operatorname{diag}(m_e, e \in E). \tag{4.16}$$

LEMMA 4.2. *The following hold*:

$$\mathbb{E}\Big[\prod_{e\in E} B_e^{m_e} \cdot \det(\mathrm{Id} - M\mathcal{G})\Big] = 1. \tag{4.17}$$

If in addition $0 \leq m_e < W_e$ *for* $e \in E_+$ *defined in* (4.1) *and* $m_e = 0$ *otherwise*, *the equality implies*

$$\mathbb{E}\Big[\prod_{e\in E_+} \big(\cosh(u_{e_+} - u_{e_-})\big)^{m_e}\Big] \leq \mathbb{E}\Big[\prod_{e\in E_+} B_e^{m_e}\Big] \leq \prod_{e\in E_+} \frac{1}{1 - \frac{m_e}{W_e}}. \tag{4.18}$$

PROOF. Ward identity (4.10) with $f_e(x) = x^{m_e}$ reads

$$\Big\langle \prod_{e\in E} S_e^{m_e} \Big\rangle = 1. \tag{4.19}$$

We calculate for $e = \{i,j\}$, using $(\overline{\psi}^t F_{\cdot e} F_{e\cdot}^t \psi)^2 = 0$ and $(\overline{\psi}^t F_{\cdot e} m_e Q_e F_{e\cdot}^t \psi)^2 = 0$,

$$\begin{aligned} S_e^{m_e} &= \big(B_e + (\overline{\psi}_i - \overline{\psi}_j)(\psi_i - \psi_j)e^{u_i+u_j}\big)^{m_e} = \big(B_e + e^{u_i+u_j}\overline{\psi}^t F_{\cdot e} F_{e\cdot}^t \psi\big)^{m_e} \\ &= B_e^{m_e} + m_e B_e^{m_e-1} e^{u_i+u_j} \overline{\psi}^t F_{\cdot e} F_{e\cdot}^t \psi = B_e^{m_e}\Big(1 + \frac{m_e}{B_e} e^{u_i+u_j} \overline{\psi}^t F_{\cdot e} F_{e\cdot}^t \psi\Big) \\ &= B_e^{m_e}\big(1 + \overline{\psi}^t F_{\cdot e} m_e Q_e F_{e\cdot}^t \psi\big) = B_e^{m_e} \exp\big(\overline{\psi}^t F_{\cdot e} m_e Q_e F_{e\cdot}^t \psi\big). \end{aligned} \tag{4.20}$$

Hence, taking a product over e and using $\sum_{e\in E} F_{\cdot e} m_e Q_e F_{e\cdot}^t = FMQF^t$, we obtain

$$1 = \Big\langle \prod_{e\in E} S_e^{m_e} \Big\rangle = \Big\langle \exp\big(\overline{\psi}^t FMQF^t \psi\big) \prod_{e\in E} B_e^{m_e} \Big\rangle. \tag{4.21}$$

The Grassmann part in the action A in (4.5) is given by

$$\begin{aligned} \sum_{\{i,j\}\in E_+} W_{ij}(\overline{\psi}_i - \overline{\psi}_j)(\psi_i - \psi_j)e^{u_i+u_j} &= \sum_{e\in E_+} \overline{\psi}^t F_{\cdot e} \mathcal{W}_e F_{e\cdot}^t \psi \\ &= \overline{\psi}^t F \mathcal{W} F^t \psi = \overline{\psi}^t D \psi. \end{aligned} \tag{4.22}$$

Using the definition (4.8) of μ, we can rewrite (4.21) as

$$\begin{aligned} 1 &= \int_{\mathbb{R}^\Lambda \times \mathbb{R}^\Lambda} \prod_{i\in\Lambda} \partial_{\overline{\psi}_i} \partial_{\psi_i} \exp\big(-\overline{\psi}^t (D - FMQF^t)\psi\big) \prod_{e\in E} B_e^{m_e} \frac{\mu(du\,ds)}{\det D} \\ &= \mathbb{E}\Big[\prod_{e\in E} B_e^{m_e} \frac{\det(D - FMQF^t)}{\det D}\Big]. \end{aligned} \tag{4.23}$$

Using the identity $\det(\mathrm{Id}+AB)=\det(\mathrm{Id}+BA)$, which holds for arbitrary rectangular matrices $A\in\mathbb{R}^{m\times n}$, $B\in\mathbb{R}^{n\times m}$, and the fact $MQ=\sqrt{Q}M\sqrt{Q}$, we obtain

$$\begin{aligned}\frac{\det(D-FMQF^t)}{\det D}&=\det\bigl(\mathrm{Id}-D^{-1}FMQF^t\bigr)\\&=\det\bigl(\mathrm{Id}-M\sqrt{Q}F^tD^{-1}F\sqrt{Q}\bigr)=\det(\mathrm{Id}-M\mathcal{G}),\end{aligned}\tag{4.24}$$

where in the second expression the identity matrix Id is indexed by vertices, while in the third and fourth expression it is indexed by edges. We conclude claim (4.17).

Assume now $0\le m_e<W_e$ for $e\in E_+$ and $m_e=0$ otherwise. We calculate

$$\det(\mathrm{Id}-M\mathcal{G})\det D=\det\bigl(D-FMQF^t\bigr)=\det\bigl(F(\mathcal{W}-MQ)F^t\bigr).\tag{4.25}$$

Note that $F(\mathcal{W}-MQ)F^t$ is a discrete Laplacian with weights $\mathcal{W}_e-m_eQ_e$ rather than $\mathcal{W}_e$. Edges in E_+ have a positive weight $\mathcal{W}_e-m_eQ_e=\mathcal{W}_e(1-\frac{m_e}{W_eB_e})>0$ since $B_e\ge 1$. Using the matrix tree theorem and writing $\mathcal{S}$ for the set of spanning trees of the graph G_+, we rewrite the last determinant as follows:

$$\det\bigl(F(\mathcal{W}-MQ)F^t\bigr)=\sum_{T\in\mathcal{S}}\prod_{e\in T}(\mathcal{W}_e-m_eQ_e)=\sum_{T\in\mathcal{S}}\prod_{e\in T}\mathcal{W}_e\biggl(1-\frac{m_e}{W_eB_e}\biggr).\tag{4.26}$$

For each $T\in\mathcal{S}$, using $B_e\ge 1$ again, we have

$$\prod_{e\in T}\biggl(1-\frac{m_e}{W_eB_e}\biggr)\ge\prod_{e\in E_+}\biggl(1-\frac{m_e}{W_eB_e}\biggr)\ge\prod_{e\in E_+}\biggl(1-\frac{m_e}{W_e}\biggr).\tag{4.27}$$

Therefore, we obtain

$$\det\bigl(F(\mathcal{W}-MQ)F^t\bigr)\ge\prod_{e\in E_+}\biggl(1-\frac{m_e}{W_e}\biggr)\sum_{T\in\mathcal{S}}\prod_{e\in T}\mathcal{W}_e=\prod_{e\in E_+}\biggl(1-\frac{m_e}{W_e}\biggr)\det D.\tag{4.28}$$

It follows from (4.25) that

$$\det(\mathrm{Id}-M\mathcal{G})\ge\prod_{e\in E_+}\biggl(1-\frac{m_e}{W_e}\biggr).\tag{4.29}$$

Inserting this in (4.17) and using $\cosh(u_{e_+}-u_{e_-})\le B_e$, the claim (4.18) follows. □

4.3. *Proof of Theorem* 1.7. Take $\overline{W}_2$ large enough, to be specified below, and let $\overline{W}\ge\overline{W}_2$. By assumption, $W_{ij}\ge\overline{W}1_{\{\|i-j\|_2=1\}}=:W_{ij}^{\mathrm{nn}}$ and $h_i=W_{i\rho}\ge W_{i\rho}^{\mathrm{nn}}=:h_i^{\mathrm{nn}}$ where $W_{i\rho}^{\mathrm{nn}}$ is defined as in (1.2). Therefore, for $m\ge 1$, by Fact 2.2, we can compare our model with weights W, h with the nearest neighbor model with weights W^{nn}, h^{nn}:

$$\mathbb{E}_{W,h}^{\Lambda_N}\bigl[e^{\sigma m u_i}\bigr]\le\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[e^{\sigma m u_i}\bigr].\tag{4.30}$$

Note that we have $h_i^{\mathrm{nn}}=0$ for $i\notin\partial\Lambda_N$, while $h_i^{\mathrm{nn}}=W_{i\rho}>0$ for $i\in\partial\Lambda_N$. Therefore, we take $j\in\partial\Lambda_N$ and use the Cauchy–Schwarz inequality to obtain

$$\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[e^{\sigma m u_i}\bigr]\le\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[e^{2\sigma m(u_i-u_j)}\bigr]^{\frac12}\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[e^{2\sigma m u_j}\bigr]^{\frac12}.\tag{4.31}$$

Assuming now that $\overline{W}_2$ is so large that $\overline{W}_2\ge 2^8$ and [9], Theorem 1, is applicable, we have

$$\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[e^{2\sigma m(u_i-u_j)}\bigr]\le 2^{2m}\mathbb{E}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}^{\Lambda_N}\bigl[\bigl(\cosh(u_i-u_j)\bigr)^{2m}\bigr]\le 2^{2m+1}\tag{4.32}$$

for all $m\in[1,\frac12\overline{W}^{\frac18}]\neq\varnothing$. Note that this result holds for any choice of the pinning and in any dimension $d\ge 3$.

The next bound uses $u_\rho = 0$, the estimate (4.18) from Lemma 4.2 which is valid on any finite graph, and the inequality $W_{j\rho} \geq \overline{W}$, which follows from $j \in \partial\Lambda_N$, to obtain

$$\begin{aligned}\mathbb{E}^{\Lambda_N}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}\big[e^{2\sigma m u_j}\big] &\leq 2^{2m}\mathbb{E}^{\Lambda_N}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}\big[(\cosh u_j)^{2m}\big] = 2^{2m}\mathbb{E}^{\Lambda_N}_{W^{\mathrm{nn}},h^{\mathrm{nn}}}\big[\big(\cosh(u_j - u_\rho)\big)^{2m}\big] \\ &\leq \frac{2^{2m}}{1-\frac{2m}{W_{j\rho}}} \leq \frac{2^{2m}}{1-\frac{2m}{\overline{W}}} \leq 2^{2m+1},\end{aligned} \tag{4.33}$$

where the last step is a consequence of $2m \leq \overline{W}^{\frac{1}{8}}$ and $\overline{W} \geq 2^{8/7}$. Substituting this and (4.32) in (4.31), the result (1.11) now follows. $\square$

Note that it would suffice to do this proof for $d = 3$. Indeed for $d \geq 4$ one can embed $\mathbb{Z}^3$ in $\mathbb{Z}^d$ and use Poudevigne's monotonicity (Fact 2.2) together with the result in $d = 3$.

Acknowledgments. We are grateful to Malek Abdesselam for suggesting we compare the hierarchical model with the long-range model on $\mathbb{Z}^d$. We also thank Christophe Sabot, Pierre Tarrès, and Xiaolin Zeng for useful discussions. We would like to thank an anonymous Associate Editor for suggesting a modification of the main results and a major simplification of their proof. We would also like to thank an anonymous referee for careful reading and constructive remarks.

Funding. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within SPP 2265 Random Geometric Systems (Project number 443757604) and the Hausdorff Center for Mathematics (project ID 390685813).

REFERENCES

[1] ANGEL, O., CRAWFORD, N. and KOZMA, G. (2014). Localization for linearly edge reinforced random walks. *Duke Math. J.* **163** 889–921. MR3189433 https://doi.org/10.1215/00127094-2644357

[2] BAUERSCHMIDT, R. and HELMUTH, T. (2023). Spin systems with hyperbolic symmetry: A survey. In *ICM—International Congress of Mathematicians*, *Vol*. 5. *Sections* 9–11 3986–4008. EMS Press, Berlin. MR4680390

[3] BÄUMLER, J. (2023). Recurrence and transience of symmetric random walks with long-range jumps. *Electron. J. Probab.* **28** Paper No. 106, 24. MR4632146 https://doi.org/10.1214/23-ejp998

[4] BRYDGES, D., EVANS, S. N. and IMBRIE, J. Z. (1992). Self-avoiding walk on a hierarchical lattice in four dimensions. *Ann. Probab.* **20** 82–124. MR1143413

[5] CAPUTO, P., FAGGIONATO, A. and GAUDILLIÈRE, A. (2009). Recurrence and transience for long range reversible random walks on a random point process. *Electron. J. Probab.* **14** 2580–2616. MR2570012 https://doi.org/10.1214/EJP.v14-721

[6] DAVIS, B. and VOLKOV, S. (2002). Continuous time vertex-reinforced jump processes. *Probab. Theory Related Fields* **123** 281–300. MR1900324 https://doi.org/10.1007/s004400100189

[7] DAVIS, B. and VOLKOV, S. (2004). Vertex-reinforced jump processes on trees and finite graphs. *Probab. Theory Related Fields* **128** 42–62. MR2027294 https://doi.org/10.1007/s00440-003-0286-y

[8] DISERTORI, M., MERKL, F. and ROLLES, S. W. W. (2022). The non-linear supersymmetric hyperbolic sigma model on a complete graph with hierarchical interactions. *ALEA Lat. Amer. J. Probab. Math. Stat.* **19** 1629–1648. MR4517733 https://doi.org/10.30757/alea.v19-62

[9] DISERTORI, M., SPENCER, T. and ZIRNBAUER, M. R. (2010). Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model. *Comm. Math. Phys.* **300** 435–486. MR2728731 https://doi.org/10.1007/s00220-010-1117-5

[10] LYONS, R. and PERES, Y. (2016). *Probability on Trees and Networks*. *Cambridge Series in Statistical and Probabilistic Mathematics* **42**. Cambridge Univ. Press, New York. MR3616205 https://doi.org/10.1017/9781316672815

[11] POUDEVIGNE-AUBOIRON, R. (2024). Monotonicity and phase transition for the VRJP and the ERRW. *J. Eur. Math. Soc.* (*JEMS*) **26** 789–816. MR4721024 https://doi.org/10.4171/jems/1298

[12] SABOT, C. (2021). Polynomial localization of the 2D-vertex reinforced jump process. *Electron. Commun. Probab.* **26** Paper No. 1, 9. MR4218029 https://doi.org/10.1214/20-ecp356

[13] SABOT, C. and TARRÈS, P. (2015). Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. *J. Eur. Math. Soc.* (*JEMS*) **17** 2353–2378. MR3420510 https://doi.org/10.4171/JEMS/559

[14] SABOT, C. and ZENG, X. (2019). A random Schrödinger operator associated with the vertex reinforced jump process on infinite graphs. *J. Amer. Math. Soc.* **32** 311–349. MR3904155 https://doi.org/10.1090/jams/906

[15] SWAN, A. (2020). Superprobability on Graphs. PhD thesis, Univ. Cambridge.

[16] ZIRNBAUER, M. R. (1991). Fourier analysis on a hyperbolic supermanifold with constant curvature. *Comm. Math. Phys.* **141** 503–522. MR1134935