MINIMAL SPANNING TREES AND STEIN’S METHOD

SOURAV CHATTERJEE AND SANCHAYAN SEN

Abstract. Kesten and Lee [23] proved that the total length of a minimal spanning tree on certain random point configurations in $\mathbb{R}^d$ satisfy a central limit theorem. They also raised the question: how to make these results quantitative? However, unlike other functionals studied in geometric probability, the problem of determining the convergence rate in the central limit theorem for Euclidean minimal spanning trees remained open. In this work, we establish bounds on the convergence rate for the Poissonized version of this problem. We also derive bounds on the convergence rate for the analogous problem in the setup of the lattice $\mathbb{Z}^d$. The main ingredients in the proof are (a) a quantification of the Burton-Keane argument for the uniqueness of the infinite open cluster, and (b) Stein’s method of normal approximation.

CONTENTS

1. Introduction 1
2. Main Results 3
3. Notations 5
4. Outline of proof 8
5. Quantification of the Burton-Keane argument 9
6. Proofs 10
  6.1. Proof of Lemma 5.1 11
  6.2. Proof of Theorem 2.1 18
  6.3. Proof of Lemma 5.2 33
  6.4. Proof of Theorem 2.4 35
  6.5. Proof of Theorem 2.5 40
Appendix A. 43
References 49

1. INTRODUCTION

Consider a finite, connected weighted graph $(V, E, w)$ where $(V, E)$ is the underlying graph and $w : E \to [0, \infty)$ is the weight function. A spanning...
tree of \((V,E)\) is a tree which is a connected subgraph of \((V,E)\) with vertex set \(V\). A minimal spanning tree (MST) \(T\) of \((V,E,w)\) satisfies
\[
M(V,E,w) := \sum_{e \in T} w(e) = \min \{ \sum_{e \in T'} w(e) : T' \text{ is a spanning tree of } (V,E) \}.
\]
We will simply write \(M(V,E)\) instead of \(M(V,E,w)\) if the weight function is clear. In this paper, whenever \((V,E)\) is a graph on some random point configuration in \(\mathbb{R}^d\), the weight function will map every edge to its Euclidean length.

Minimal spanning trees and other related functionals are of great interest in geometric probability. For an account of law of large numbers and related asymptotics for these functionals, see e.g. \([3, 4, 8, 12, 34, 35]\). The central limit theorems (CLT) for three such functionals, namely the lengths of the \(k\)-th nearest graph, the Delaunay triangulation and the Voronoi diagram on Poisson point configurations in \([0,1]^2\) were established in \([9]\). Central limit theorems for minimal spanning trees were first proven by Kesten and Lee \([23]\) and Alexander \([7]\). In \([23]\), the CLT for the total weight of an MST on both the complete graph on Poisson points inside \([0,n^{1/d}]^d\) and the complete graph on \(n\) i.i.d. uniformly distributed points inside \([0,1]^d\) were established when \(d \geq 2\) (their results included the case of more general weight functions and not just Euclidean distances). Alexander \([7]\) proved the CLT for the Poissonized problem in two dimensions. Later certain other CLTs related to MSTs were proven in \([24]\) and \([25]\).

Studies related to Euclidean MSTs in several other directions were undertaken in \([10, 13, 28, 29, 30]\). An account of the structural properties of minimal spanning forests (in both Euclidean and non-Euclidean setting) can be found in \([6, 51, 26, 22]\) and the references therein. For an account of the scaling limit of minimal spanning trees, see e.g. \([11, 16, 32]\).

The methods of \([7]\) and \([23]\) cannot be used to get bounds on the rate of convergence to normality in the CLT for Euclidean MSTs. Indeed, Kesten and Lee remark that
\[
"... [A] drawback of our approach is that it is not quantitative. Further ideas are needed to obtain an error estimate in our central limit theorem."
\]
In \([24]\), it was shown that Euclidean MSTs have certain “stabilizing” properties but there was no quantitative bound on how fast this stabilization occurs. This poses the major difficulty in obtaining an error estimate in the CLT.

In this paper we use a variation of Stein’s method, given by an approximation theorem from \([18]\), to connect the problem of bounding the convergence rate in this CLT to the problem of getting upper bounds on the probabilities of certain events in the setup of continuum percolation driven by a Poisson process and thus obtaining an error estimate in this CLT (Theorem \([2.1]\)). Using a similar approach, we also obtain error estimates in the CLT for the
total weights of the MSTs on subgraphs of $\mathbb{Z}^d$ under various assumptions on the edge weights (Theorem 2.4). The percolation theoretic estimates used in the proofs are given in Lemma 5.1 and Lemma 5.2. These are new results about higher dimensional critical percolation, possibly of independent interest to experts in percolation theory. In Theorem 2.5, we present a general CLT satisfied by the MSTs on subgraphs of a vertex-transitive graph.

2. Main Results

We summarize our main results in this section. Define the distance $D(\mu_1, \mu_2)$ between two probability measures $\mu_1$ and $\mu_2$ on $\mathbb{R}$ by the sup norm of the difference between their distribution functions, or equivalently

$$D(\mu_1, \mu_2) := \sup_{x \in \mathbb{R}} |\mu_1(-\infty, x] - \mu_2(-\infty, x]|.$$  

This metric is sometimes called the ‘Kolmogorov distance’. A bound on the Kolmogorov distance between two probability measures is sometimes called a ‘Berry-Esséen bound’.

Recall from the previous section that for a graph $(V,E)$ with weighted edges, $M(V,E)$ denotes the sum of edge weights of a minimal spanning tree of $(V,E)$. Our result on Euclidean minimal spanning trees is the following.

**Theorem 2.1.** Let $\mathcal{P}$ be a Poisson process with intensity one in $\mathbb{R}^d$. Let $(V_n,E_n)$ be the complete graph on $\mathcal{P} \cap [-n,n]^d$ with each edge weighted by its Euclidean length. Let $\mu_n$ be the law of $(M_n - \mathbb{E}(M_n))/\sqrt{\text{Var}(M_n)}$ where $M_n = M(V_n,E_n)$. Let $G$ denote the standard normal distribution on $\mathbb{R}$.

(i) When $d = 2$, there is a positive constant $\beta_1$ such that for every $p > 1$,

$$D(\mu_n, G) \leq c_1 n^{-\frac{\beta_1}{2(p+\beta_1)}}$$

where $c_1$ is a positive constant which depends only on $p$.

(ii) When $d \geq 3$, for every $p > 1$ we have

$$D(\mu_n, G) \leq c_2 (\log n)^{-\frac{d}{8p}}$$

for a positive constant $c_2$ depending only on $p$ and $d$.

**Remark 2.2.** If $\mathcal{P}_\lambda$ is a Poisson process with intensity $\lambda$ in $\mathbb{R}^d$ and $M_n(\lambda)$ is the weight of a minimal spanning tree of the complete graph on $\mathcal{P}_\lambda \cap [-n/\lambda^\frac{1}{d}, n/\lambda^\frac{1}{d}]^d$, then $(M_n(\lambda) - \mathbb{E}M_n(\lambda))/\sqrt{\text{Var}(M_n(\lambda))}$ is distributed as $\mu_n$ where $\mu_n$ is as defined in the statement of Theorem 2.1. For this reason, it is enough to consider only Poisson processes with intensity one.

**Remark 2.3.** It is very likely that the bounds are sub-optimal. Sharper percolation theoretic estimates than the ones obtained in this paper will give better bounds. However, the question of optimal error bounds is probably very difficult.

The following theorem deals with the case of minimal spanning trees on subsets of $\mathbb{Z}^d$. Here $p_c(\mathbb{Z}^d)$ denotes the critical probability of bond percolation in $\mathbb{Z}^d$. 


**Theorem 2.4.** Suppose that we put i.i.d. nonnegative weights (having some non-degenerate distribution $\mu$) on the edges of the lattice $\mathbb{Z}^d$. Let $M_n$ denote the total weight of a minimal spanning tree of the weighted subgraph of $\mathbb{Z}^d$ within the cube $[-n,n]^d$ and let $\nu_n$ be the distribution of $(M_n - \mathbb{E}(M_n))/\sqrt{\text{Var}(M_n)}$. Let $G$ be the standard normal distribution on $\mathbb{R}$.

(i) If $\mu$ has unbounded support and $\int_0^\infty x^{4+\delta} \mu(dx) < \infty$ for some $\delta > 0$, then

\[
\mathcal{D}(\nu_n, G) \leq c_3 n^{-\frac{\beta_2}{2(\beta_2+2)}} \quad \text{if } d = 2 \quad \text{and}
\]

\[
\mathcal{D}(\nu_n, G) \leq c_4 (\log n)^{-\frac{d}{8\delta}} \quad \text{if } d \geq 3
\]

for some positive constants $c_3$, $c_4$ and $\beta_2$ with $c_3$ depending on $\mu$ and $\delta$ and $c_4$ depending on $\mu$, $\delta$ and $d$. The constant $p$ equals $1 + 3/\delta$. Further, if $d \geq 3$ and either $\mu[0,x] = p_c(\mathbb{Z}^d)$ for some unique $x \in \mathbb{R}$ or $\mu[0,x] = p_c(\mathbb{Z}^d)$ for some unique $x \in \mathbb{R}$, then

\[
\lim_n (\log n)^{\frac{d}{8\delta}} \mathcal{D}(\nu_n, G) = 0.
\]

(ii) If $\mu$ has bounded support, then

\[
\mathcal{D}(\nu_n, G) \leq c_5 n^{-\frac{\beta_2}{2(\beta_2+2)}}, \quad \text{if } d = 2 \quad \text{and}
\]

\[
\mathcal{D}(\nu_n, G) \leq c_6 (\log n)^{-\frac{d}{8\delta}}, \quad \text{if } d \geq 3
\]

for some positive constants $c_5$ depending on $\mu$ and $c_6$ depending on $\mu$ and $d$. Further, if $d \geq 3$ and either $\mu[0,x] = p_c(\mathbb{Z}^d)$ for some unique $x \in \mathbb{R}$ or $\mu[0,x] = p_c(\mathbb{Z}^d)$ for some unique $x \in \mathbb{R}$ then

\[
\lim_n (\log n)^{\frac{d}{8\delta}} \mathcal{D}(\nu_n, G) = 0.
\]

(iii) If $d \geq 3$, $\int_0^\infty x^{4+\delta} \mu(dx) < \infty$ for some $\delta > 0$ and $\mu[0,x] > p_c(\mathbb{Z}^d)$ for some $x \in \mathbb{R}$, then for every $\gamma < d/4$, we have

\[
\mathcal{D}(\nu_n, G) \leq c_7 n^{-\gamma}
\]

for some positive constant $c_7$ depending on $\mu$, $d$ and $\gamma$.

In the proof of Theorem 2.4, we need a bound on the probability of the so called two arm crossing which is uniform in $p$ over an interval containing $p_c(\mathbb{Z}^d)$ (see Lemma 5.2). After finishing our work on this paper, we came across the interesting recent preprint of Cerf [17]. In Proposition 5.2 of [17], Cerf proves one such bound which is an improvement of our Lemma 5.2, but for site percolation instead of bond percolation. This bound follows from arguments in [2] and [20]. It seems to us that it may be straightforward to generalize Cerf’s result to bond percolation; assuming that this generalization is true, the following improvement of Theorem 2.4 — which we state without proof — should hold.
Theorem 2.4*. Assume the setup of Theorem 2.4 Assume that either \( \mu \) has unbounded support and \( \int_0^\infty x^{4+\delta} \mu(dx) < \infty \) for some \( \delta > 0 \) or that \( \mu \) has bounded support. Let \( p \) equal \( 1 + 3/\delta \) (resp. 1) if \( \mu \) has unbounded support but finite \((4+\delta)\)th moment (resp. bounded support). Then for \( d \geq 2 \),

\[
D(\nu_n, G) \leq c_3^*(\log n)^{1/p} n^{1/(4(2p+1))}
\]

for some constant \( c_3^* \) depending only on \( d \) and \( \delta \). Further, if \( d \geq 3 \) and either \( \mu[0, x] = p_c(\mathbb{Z}^d) \) for some unique \( x \in \mathbb{R} \) or \( \mu[0, x] = p_c(\mathbb{Z}^d) \) for some unique \( x \in \mathbb{R} \), then

\[
\lim_n n^{1/(4(2p+1))) D(\nu_n, G)/(\log n)^{1/p} = 0.
\]

Similarly, any improvement over Lemma 5.1 will yield a sharper upper bound in (2.2).

Our approach can be used to give a simple proof of asymptotic normality of the total weight of the minimal spanning tree under a very general assumption on the underlying graph. We present this result in the following theorem. The advantage of this approach is that we can get the convergence rate in the central limit theorem whenever we can prove the percolation theoretic estimates analogous to the ones used in the proofs of Theorem 2.1 and Theorem 2.4.

For a graph \( G = (V, E) \) and a vertex \( v \in V \), we will write \( S(v, r) \) to denote the subgraph of \( G \) spanned by the set of all vertices \( v' \in V \) such that \( d_G(v', v) \leq r \) where \( d_G \) denotes the graph distance of \( G \).

Theorem 2.5. Let \( G = (V, E) \) be a (I) connected, infinite, locally finite, vertex-transitive graph.

Consider a sequence of finite connected subgraphs \( G_n = (V_n, E_n) \) such that

(II) \(|V_n| \to \infty \) and

(III) \(|\{v \in V_n : S(v, r) \not\subseteq G_n\}| = o(|V_n|) \) for every \( r > 0 \).

Consider i.i.d. nonnegative weights associated with the edges of \( G \) where the weights follow some non-degenerate distribution \( \mu \) having finite \((4+\delta)\)-th moment (for some \( \delta > 0 \)). Let \( M_n \) be the total weight of a minimal spanning tree of \( G_n \). Then

(i) \( \text{Var}(M_n) = \Theta(|V_n|) \) and

(ii) \( (M_n - E(M_n))/\sqrt{\text{Var}(M_n)} \xrightarrow{d} N(0,1) \).

3. Notations

We will use some notations frequently throughout this paper. For convenience we collect them together in this section.

If \( x \) is a point in \( \mathbb{R}^d \) and \( A \subset \mathbb{R}^d \), then we define \( x + A := \{x + y : y \in A\} \).

If \( r > 0 \), \( S(x, r) \) will denote the closed \( L^2 \) ball of radius \( r \) centered at \( x \) and \( B(x, r) \) will denote the closed \( L^\infty \) ball of radius \( r \) centered at \( x \), i.e. \( B(x, r) = x + [-r, r]^d \). When \( x \) is the origin, we will simply write \( B(r) \)
instead of \( B(0, r) \). For any cube \( B \), we refer to its center as \( c(B) \). We will denote by \( d(\cdot, \cdot) \), the metric induced by the \( L^2 \) norm in \( \mathbb{R}^d \).

When working with the lattice \( \mathbb{Z}^d \), \( B(x, r) \) will denote the set of all lattice points inside \( x + [-r, r]^d \) and \( B(r) \) will stand for \( B(0, r) \). For a subset \( V \) of \( \mathbb{Z}^d \), let \( G(V) \) denote the subgraph of \( \mathbb{Z}^d \) induced by \( V \). We will sometimes make abuse of notation by referring to \( G(V) \) as \( V \). With this convention \( B(x, r) \) will sometimes mean \( G(B(x, r)) \) and the meaning will be clear from the context. For a cube \( Q \) in \( \mathbb{Z}^d \), \( \partial Q \) will denote the set of all vertices in \( Q \) which are adjacent to at least one vertex not in \( Q \).

Suppose we have a Poisson process \( \mathcal{P} \) in \( \mathbb{R}^d \). For \( A \subset \mathbb{R}^d \) having nonempty interior, consider the complete graph on \( \mathcal{P} \cap A \). As explained earlier the edge weights in this situation are the Euclidean distances between the endpoints. We will often write \( M(\mathcal{P} \cap A) \) to mean the sum of edge weights of the minimal spanning tree on this graph or simply \( M(A) \) when there is no scope for confusion. Similarly for a finite subset \( X \) of \( \mathbb{R}^d \), \( M(X) \) will mean the total weight of the minimal spanning tree on the complete graph on \( X \) having Euclidean distance as edge weights.

For a connected finite graph \( G = (V, E) \) having edge weights \( \{X_e\}_{e \in E} \), \( M(\{X_e\}_{e \in E}) \) will denote the total weight of the minimal spanning tree on \( (V, E) \), we will write this simply as \( M(G) \) when the edge weights are clear. For any \( e \in E \), \( G - e \) will denote the graph \( (V, E - \{e\}) \) and \( M(G - e) \) will mean the total weight of the minimal spanning tree on \( (V, E - \{e\}) \) (provided \( G - e \) is connected).

For \( A \subset \mathbb{R}^d \) and \( r > 0 \), we define

\[
A^{(r)} := \{x \in \mathbb{R}^d : d(x, A) \leq r \}
\]

where \( d \) is the Euclidean metric. Let us also define

\[
A^{(r)} := \{x \in A : d(x, \partial A) \leq r \}.
\]

Let \( \mathcal{P} \) be a Poisson process in \( \mathbb{R}^d \) and let \( A \) be a subset of \( \mathbb{R}^d \). Then \( \mathcal{C} = \{x_i\}_{i \in I} \subset \mathcal{P} \cap A \) will be called an \( r \)-cluster in \( A \) (or just \( r \)-cluster) if \( \bigcup_{i \in I} S(x_i, r) \) forms a connected component of \( (\mathcal{P} \cap A)^{(r)} \). We will refer to the region \( \bigcup_{i \in I} S(x_i, r) \) occupied by \( \mathcal{C} \) as \( R(\mathcal{C}) \). We say that two \( r \)-clusters \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) in \( A \subset \mathbb{R}^d \) are disjoint if \( R(\mathcal{C}_1) \) and \( R(\mathcal{C}_2) \) are. We emphasize that the occupied regions must be disjoint in \( \mathbb{R}^d \), it is not enough to have their restrictions to \( A \) to be disjoint. We will write configuration to mean a locally finite subset of \( \mathbb{R}^d \). For \( A \subset \mathbb{R}^d \), \( \mathcal{X}(A) \) will denote the space of all locally finite subsets of \( A \).

For two compact sets \( K_1, K_2 \subset \mathbb{R}^d \) with \( K_1 \subset K_2 \), a positive integer \( k \) and a positive real \( r \), we write \( K_1 \leftarrow^k_r K_2 \) if there exist at least \( k \) disjoint \( r \)-clusters \( \mathcal{C}_1, \ldots, \mathcal{C}_k \) in \( K_2 - K_1 \) such that

\[
\mathcal{C}_j \cap K_1^{(r)} \neq \emptyset \text{ and } R(\mathcal{C}_j) \cap \partial K_2^{(r)} \neq \emptyset, \text{ for } j = 1, \ldots, k.
\]
We will write \( K_1 \xrightarrow[k] r K_2 \), if there exist at least \( k \) disjoint \( r \)-clusters \( C_1, \ldots, C_k \) in \( (K_2 - K_1) \) such that
\[
C_j \cap K_1^{(2r)} \neq \emptyset \text{ and } R(C_j) \cap \partial K_2^{(r)} \neq \emptyset, \text{ for } j = 1, \ldots, k.
\]

If \( K_1, K_2 \) and \( K_3 \) are cubes in \( \mathbb{R}^d \) such that \( K_1 \subset K_2 \), \( c(K_1) \in K_3 \) and \( \partial K_2 \cap K_3 \neq \emptyset \), we write \( K_1 \xrightarrow[k] r K_2 \) in \( K_3 \), if there exist at least \( k \) disjoint \( r \)-clusters \( C_1, \ldots, C_k \) in \( (K_2 - K_1) \cap K_3 \) such that
\[
C_j \cap K_1^{(r)} \neq \emptyset \text{ and } R(C_j) \cap \partial K_2^{(r)} \cap K_3 \neq \emptyset, \text{ for } j = 1, \ldots, k.
\]

Similarly we can define \( K_1 \xrightarrow[k] r K_2 \) in \( K_3 \).

For \( p \in [0, 1] \), consider i.i.d. Bernoulli\((p)\) random variables \( \{X_e\}_{e \in \mathbb{Z}^d} \) associated with edges of \( \mathbb{Z}^d \), i.e. \( \mathbb{P}(X_e = 1) = p = 1 - \mathbb{P}(X_e = 0) \). We call an edge \( e \) open (resp. closed) at level \( p \) if \( X_e = 1 \) (resp. \( X_e = 0 \)). Given a subgraph \( G = (V, E) \) of \( \mathbb{Z}^d \) and \( V' \subset V \), we say that \( V' \) forms a \( p \)-cluster in \( G \) if there is a path consisting of open edges in \( E \) between any two vertices in \( V' \) and \( V' \) is a maximal subset of \( V \) in this regard.

For two cubes \( Q_1 \subset Q_2 \) in \( \mathbb{Z}^d \), denote by \( Q_2 - Q_1 \), the subgraph \( (V, E) \) of \( Q_2 \) with \( E = \{ \text{all edges in } Q_2 \text{ except the ones with both endpoints in } Q_1 \} \) and \( V = \{ v : v \text{ is an endpoint of } e \text{ for some } e \in E \} \).

For two cubes \( Q_1 \subset Q_2 \) in \( \mathbb{Z}^d \) and \( p \in [0, 1] \), \( Q_1 \xrightarrow[k] p Q_2 \) will mean that there exist at least \( k \) disjoint \( p \)-clusters in \( Q_2 - Q_1 \) which intersect both \( \partial Q_1 \) and \( \partial Q_2 \). If \( Q_2, Q_3 \) are cubes in \( \mathbb{Z}^d \) such that \( Q_3 \cap \partial Q_2 \neq \emptyset, x \in Q_3 \) and \( B(x, 1) \subset Q_2 \), then we will write \( B(x, 1) \xrightarrow[k] p Q_2 \) in \( Q_3 \) if there exists \( k \) disjoint \( p \)-clusters in \( (Q_2 - B(x, 1)) \cap Q_3 \) each intersecting \( \partial B(x, 1) \) and \( \partial Q_2 \).

For an edge \( \{x, y\} \) in \( \mathbb{Z}^d \) and a cube \( Q \) containing both \( x \) and \( y \), \( \{x, y\} \xrightarrow[k] p Q \) will mean that the \( p \)-clusters in \( Q \) containing \( x \) and \( y \) are disjoint and that they both intersect \( \partial Q \). Similarly we can define \( \{x, y\} \xrightarrow[k] p Q - \{x, y\} \) to be the event that the \( p \)-clusters in \( Q - \{x, y\} \) containing \( x \) and \( y \) intersect \( \partial Q \) and are disjoint.

To ease notations, most constants in this paper will be denoted by \( c, c', C \) etc. and their values may change from line to line, these constants may depend on parameters like the dimension and often we will not mention this dependence explicitly; none of these constants will depend on the quantity “\( n \)”, used to index infinite sequences. Specific constants will have a subscript as for example \( c_1, c_2 \) etc. Also we will not concern ourselves with quantities like \( n/2 \) or \( n^\alpha \) not being integers and freely write statements such as “log \( n \) many” objects when really it should be \( \lceil \log n \rceil \) or \( \lfloor \log n \rfloor \) many objects as the case may be. Since we are interested in asymptotic bounds, this will not affect our results.
4. Outline of Proof

We briefly sketch here the main ideas in the proof. Let us start with the lattice case. Consider the cube $B(n)$ and take an edge $e$ in $B(n)$. Let $X_f$ denote the weight associated with the edge $f$ in $B(n)$. Heuristically, we expect $M(B(n))$ to satisfy a CLT if the change in $M(B(n))$ due to the replacement of $X_e$ by an independent identically distributed observation $X'_e$ “is not observed far away from $e$”. A quantitative formulation of this vague statement will give us a convergence rate in the CLT.

To this end, fix $\alpha \in (0, 1)$ and take an edge $e = \{x_1, x_2\}$ in $B(n)$ such that $d(x_1, \partial B(n)) \geq \lfloor n^{\alpha}\rfloor$. Let $X = \{X_f : f$ is an edge of $B(n)\}$ and $\tilde{X} = \{X_f : f$ is an edge in $B(x_1, n^\alpha)\}$. Let $M(X^e)$ (resp. $M(\tilde{X}^e)$) be the weight of the MST of $B(n)$ (resp. $B(x_1, n^\alpha)$) when $X_e$ is replaced by $X'_e$. Define

$$\Delta_e M = M(X) - M(X^e) \text{ and } \tilde{\Delta}_e M = M(\tilde{X}) - M(\tilde{X}^e).$$

Then, an application of an approximation theorem from [18] (Theorem 6.6) reduces the problem to getting an upper bound on $\mathbb{E}|\Delta_e M - \tilde{\Delta}_e M|$. The actual calculations are given in Section 6.2 and 6.4. This is the precise formulation of the heuristics explained above.

Next, it is easily seen that getting a bound on $\mathbb{E}|\Delta_e M - \tilde{\Delta}_e M|$ amounts to proving an upper bound on $\mathbb{E} |\partial_e M|$ where

$$\partial_e M := [M(B(n)) - M(B(n) - e)] - [M(B(x_1, n^\alpha)) - M(B(x_1, n^\alpha) - e)].$$

From the properties of MSTs, it follows that

$$M(B(n)) - M(B(n) - e) = X_e - X_e \vee Y \text{ and }$$

$$M(B(x_1, n^\alpha)) - M(B(x_1, n^\alpha) - e) = X_e - X_e \vee \tilde{Y}$$

where $Y$ (resp. $\tilde{Y}$) is the maximum weight associated with the edges in the path, $\Gamma_1$ (resp. $\Gamma_2$) connecting $x_1$ and $x_2$ in an MST of $B(n) - e$ (resp. $B(x_1, n^\alpha) - e$). Thus, $\mathbb{E} |\partial_e M| \leq \mathbb{E} |\tilde{Y} - Y|$. Note that, by the well-known minimax property of paths in MSTs (Lemma 6.10), $(\tilde{Y} - Y)$ is always nonnegative and is positive only when parts of $\Gamma_1$ lie outside $B(x_1, n^\alpha)$ (Figure 1). From this, we complete the proof by arguing that $\mathbb{E} |\partial_e M|$ can be bounded in terms of the probability of the “two-arms event” in the setup of bond percolation in $B(x_1, n^\alpha)$.

For Euclidean MST, we start by dividing $B(n)$ into cubes $\{Q \in Q\}$ of disjoint interior having side length $s \in [1, 2]$. Fix a cube $Q$ with $d(c(Q), \partial B(n)) \geq n^\alpha$. Let $M(B(n)^Q)$ (resp. $M(B(c(Q), n^\alpha)^Q)$) denote the weight of the MST on the points inside $B(n)$ (resp. $B(c(Q), n^\alpha)$) when the configuration inside $Q$ is replaced by an independent Poisson configuration. Our aim here is to get a bound on $\mathbb{E}|\Delta_Q M - \tilde{\Delta}_Q M|$ where

$$\Delta_Q M = M(B(n)) - M(B(n)^Q) \text{ and }$$

$$\tilde{\Delta}_Q M = M(B(c(Q), n^\alpha)) - M(B(c(Q), n^\alpha)^Q).$$
This can also be reduced to getting a bound on the probability of the analogue of the two-arms event in the setup of continuum percolation. However, since all possible edges between points are permitted, this step requires a little work. We achieve this by introducing the concept of a “wall” (Definition 6.8) and then using the add and delete algorithm from [23]. We will omit the details of these steps from the proof sketch.

5. Quantification of the Burton-Keane argument

The key ingredients in the proofs of Theorem 2.1 and Theorem 2.4 are some percolation theoretic estimates which are also interesting in their own rights. We state them in the following Lemmas.

Lemma 5.1. Assume $d \geq 3$ and let $\mathcal{P}$ be a Poisson process having intensity one in $\mathbb{R}^d$. Then for every fixed $r$ in a compact interval $[r_1, r_2]$ and $a \in (1/2, \log \log n)$, we have

$$
\mathbb{P}(B(a) \rightarrow B(n)) \leq \frac{c_8 \exp(c_9 a^{d-1}) a^{d/2}}{(\log n)^{d/2}}
$$

for some constants $c_8$ and $c_9$ depending only on $r_1$, $r_2$ and $d$. The same bound holds if we replace $B(a)$ by $B(a)^{(r)}$ or $B(a)^{(r)} \cup S(x, r)$ for some $x \in B(a)^{(r)}$.

The next Lemma states the lattice analogue of (5.1).

Lemma 5.2. Consider the lattice $\mathbb{Z}^d$ where $d \geq 3$. Denote the vertices adjacent to the origin by $e_1, \ldots, e_{2d}$. Then for every fixed $p$ in a compact
interval $[p_1, p_2] \subset (0, 1)$,

\[ (5.2) \quad P\left(\{0, e_i\} \leftrightarrow B(n)\right) \leq c_{10}(\log n)^{-\frac{d}{2}}, \quad \text{for } 1 \leq i \leq 2d \]

for some constant $c_{10}$ depending only on $p_1$, $p_2$ and $d$. The same bound holds if we replace the edge $\{0, e_i\}$ by the cube $B(1)$.

**Remark 5.3.** Note that we can actually get an exponentially decaying bound in (5.1) when $r_2 < r_c := r_c(1, \mathbb{R}^d)$ where $r_c(\lambda, A)$ is the critical radius for continuum percolation in $A \subset \mathbb{R}^d$ driven by a Poisson process with intensity $\lambda$ (see e.g. [7] or [14]). Also if $r_c'(\lambda) := \lim_{L \to \infty} r_c(\lambda, \mathbb{R}^2 \times [0, L]^{d-2})$ then it is possible to get exponential decay in (5.1) if $r_1 > r_c' := r_c'(1)$. So the bound in (5.1) is really useful when $[r_c, r_c'] \subset [r_1, r_2]$. We remark here that the results of [36] show that $r_c(\lambda, \mathbb{R}^d) = \lim_{L \to \infty} r_c(\lambda, \mathbb{R}^2 \times [0, L]^{d-2})$. That $r_c(\lambda, \mathbb{R}^d) = r_c'(\lambda)$ was not proven in this paper, but we will not need it.

The same is true for Lemma 5.2, the exponential decay in (5.2) is standard when $p_c(\mathbb{Z}^d) \notin [p_1, p_2]$.

**Remark 5.4.** The two lemmas stated above may be seen as quantifications of the statement that the infinite open cluster is unique. This uniqueness theorem was first proved by Aizenman, Kesten and Newman [2] for percolation on lattices. A new and very elegant proof was given by Burton and Keane [15], which has now become the standard textbook proof of the theorem. Unlike the original argument of Aizenman, Kesten and Newman, the Burton–Keane argument admits a wide array of applications and generalizations due to its simplicity and robustness. The AKN argument is known to have a quantitative version (see [17]), while the Burton–Keane argument, due to its use of translation-invariance, is not expected to be quantifiable. The proofs of the two lemmas presented above show that it is actually possible to quantify the Burton–Keane argument. This may have wider applicability in other contexts, where the Burton–Keane argument works but the AKN argument does not.

6. Proofs

Throughout this section, all Poisson processes will have intensity one. We will write $r_c(A)$ to mean $r_c(1, A)$ and $r_c$ will denote $r_c(1, \mathbb{R}^d)$. In what follows $d$ will denote the dimension of the ambient space ($\mathbb{Z}^d$ or $\mathbb{R}^d$).

Before beginning the proof of Lemma 5.1 we recall the following simple fact.

**Lemma 6.1.** Let $X_1, \ldots, X_n$ be independent random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in some measurable space $(X, \mathcal{S})$. Let $f : X^n \to \mathbb{R}$ be a bounded measurable function. Then for any $A_1, \ldots, A_k \subset \{1, \ldots, n\}$ such that $A_i$ are pairwise disjoint, we have

\[ (6.1) \quad \text{Var}(f(X_1, \ldots, X_n)) \geq \sum_{i=1}^{k} \text{Var}\left[ \mathbb{E}(f(X_1, \ldots, X_n) | \{X_j\}_{j \in A_i}) \right]. \]
Proof. Without loss of generality, assume \( \mathbb{E}(f(X_1, \ldots, X_n)) = 0 \). Let
\[
H = \{ g \in L^2(\Omega, \mathcal{A}, \mathbb{P}) : \int g = 0 \}
\]
and
\[
H_i = \{ g \in H : g \text{ is } \sigma(\{X_j\}_{j \in A_i}) \text{ measurable} \}.
\]
Then under the natural inner product \( H \) is a Hilbert space and \( H_i \) are closed orthogonal subspaces of \( H \). Further \( \mathbb{E}(f(X_1, \ldots, X_n) | \{X_j\}_{j \in A_i}) \) is the projection of \( f(X_1, \ldots, X_n) \) on \( H_i \) and the result follows. \( \square \)

6.1. **Proof of Lemma 5.1.** Let us first prove the bound for \( \mathbb{P}(B(a) \overset{2}{\leftrightarrow} B(n)) \), the arguments are similar when we replace \( B(a) \) by the other sets. Fix \( r \in [r_1, r_2] \). We write \( \mathbb{R}^d \) as a union of cubes
\[
\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^d} B_k \text{ where } B_k = 2ak + B(a).
\]
Since \( \mathbb{P}(\mathcal{P} \cap \partial B_k \neq \emptyset \text{ for some } k \in \mathbb{Z}^d) = 0 \) we will assume that no Poisson point lies in any of the common interfaces shared by two cubes.

Consider a sequence \( a_n \to \infty \) such that \( a_n = o(n) \) but \( a_n = \Omega((\log \log n)^2) \) (so that \( a_n \) is large compared to \( a \)). Define
\[
E = \{ \exists \text{ exactly one } r\text{-cluster } C \text{ in } B(n) \text{ such that } \partial(C) \text{ intersects both } \partial(B(n)) \text{ and } \partial(B(a_n)) \}.
\]
Define \( f : \prod_{k \in \mathbb{Z}^d} \mathcal{X}(B_k) \to \mathbb{R} \) by
\[
f(\{\omega_k\}) = \mathbb{I}_E(\cup_k \omega_k).
\]
Note that \( E \) (and also \( f \)) depends only on those \( \omega_k \)'s for which \( B_k \cap B(n) \neq \emptyset \). Writing \( X_k \) for \( \mathcal{P} \cap B_k \), we have from Lemma 6.1,
\[
\text{(6.2) } \text{Var}(f(\{X_k\})) \geq \sum_{i \in \mathcal{I}} \text{Var}[\mathbb{E}(f(\{X_k\}) | X_i)]
\]
where
\[
\mathcal{I} = \left\{ k \in \mathbb{Z}^d : B_k \cap B\left(\frac{a_n}{3}\right) \neq \emptyset \right\}.
\]

Consider another Poisson process \( \mathcal{P}' \) independent of \( \mathcal{P} \) and set \( X'_k = \mathcal{P}' \cap B_k \) for \( k \in \mathbb{Z}^d \). We define
\[
\{X_k\}' := \{X_k\}_{k \neq i} \cup X'_i,
\]
also let
\[
S_i := \{ \omega_i \in \mathcal{X}(B_i) : B_{i(2r)} \subset \omega_i^{(r)} \} \quad \text{and}
\]
\[
G_i := \{ \omega_i \in \mathcal{X}(B_i) : B_{i(2r)} \cap \omega_i^{(r)} = \emptyset \}.
\]
Then for any fixed \( i \in \mathcal{I} \)

\[
\text{Var} \left[ \mathbb{E} \left( f(\{X_k\} | X_i) \right) \right] = \frac{1}{2} \mathbb{E} \left[ \left( \mathbb{E} \left( f(\{X_k\}) | X_i \right) - \mathbb{E} \left( f(\{X_k\}|i) | X_i \right) \right)^2 \right],
\]

(6.3) \[ \geq \frac{1}{2} \mathbb{E} \left[ \left( \mathbb{E} \left( f(\{X_k\}) - f(\{X_k\}|i) | X_i, X_i' \right) \right)^2 \cdot \mathbb{I}(X_i \in \mathcal{S}_i, X_i' \in \mathcal{G}_i) \right]. \]

Note that for any \( \omega_i \in \mathcal{S}_i \) and \( \omega'_i \in \mathcal{G}_i \)

\[ \{\omega \in \mathcal{X}(B(n) - B_i) : \mathbb{I}_E(\omega \cup \omega') = 1\} \subset \{\omega \in \mathcal{X}(B(n) - B_i) : \mathbb{I}_E(\omega \cup \omega_i) = 1\}. \]

Therefore, if \( X_i \in \mathcal{S}_i \) and \( X_i' \in \mathcal{G}_i \), then

(6.4) \[ f(\{X_k\}) - f(\{X_k\}|i) \geq 0. \]

Now, for any \( \omega \in \mathcal{X}(B(n) - B_i) \) for which the event \( A_i := \{B_i \leftrightarrow B(n), \text{ any } r\text{-cluster } \mathcal{C} \text{ in } B(n) - B_i \text{ for which } R(\mathcal{C}) \text{ intersects both } \partial B(a_n) \text{ and } \partial (B(n)(r)) \text{ has a point in } B_i^{(r)}\} \) is true, \( \mathbb{I}_E(\omega \cup \omega_i) = 1 \) when \( \omega_i \in \mathcal{S}_i \) and \( \mathbb{I}_E(\omega \cup \omega'_i) = 0 \) when \( \omega'_i \in \mathcal{G}_i \).

Consequently, if \( X_i \in \mathcal{S}_i \) and \( X_i' \in \mathcal{G}_i \) and \( \mathbb{I}_{A_i}(\cup_{k \neq i} X_k) = 1, \) then

\[ f(\{X_k\}) - f(\{X_k\}|i) = 1. \]

Hence from (6.3) and (6.4) we get

(6.5) \[ \text{Var} \left[ \mathbb{E} \left( f(\{X_k\} | X_i) \right) \right] \geq \frac{1}{2} \mathbb{P}(A_i)^2 \mathbb{P}(X_i \in \mathcal{S}_i) \mathbb{P}(X_i' \in \mathcal{G}_i) \]

\[ \geq \frac{1}{2} \mathbb{P}(A_i)^2 \exp(-ca^{d-1}), \]

the constant depends on \( d \) and \( r_1 \) only.

For \( i \in \mathcal{I} \) we also have

(6.6) \[ \mathbb{P}(A_i) \geq \mathbb{P}(B_i \leftrightarrow B(n); \text{ any } r\text{-cluster } \mathcal{C} \text{ in } B(n) - B_i \text{ for which } R(\mathcal{C}) \text{ intersects both } \partial B(c(B_i), 2a_n) \text{ and } \partial (B(n)(r)) \text{ has a point in } B_i^{(r)} \]

\[ \geq \mathbb{P}(B_i \leftrightarrow B(c(B_i), 2n) ; \text{ if } \mathcal{C} \text{ is an } r\text{-cluster in } B(c(B_i), 2n) - B_i \text{ then every connected component of } \left( \mathcal{C} \cap B \left( c(B_i), \frac{n}{2} \right) \right)^{(r)} \text{ which intersects both } \partial B(c(B_i), 2a_n) \text{ and } \partial \left( B \left( c(B_i), \frac{n}{2} \right)^{(r)} \right) \text{ also intersects } \partial B_i. \]
Define the event

\[ F = \{ B_0 \leftrightarrow_{2r} B(2n), \text{ if } C \text{ is an } r\text{-cluster in } B(2n) - B_0 \} \]

then every connected component of \((C \cap B(n/2))^{(r)}\) which intersects both \(\partial B(2a_n)\) and \(\partial \left( B \left( \frac{n}{2} \right)^{(r)} \right)\) also intersects \(\partial B_0\).

From (6.2), (6.5), (6.6) and translational invariance, we get

\[ \mathbb{P}(F) \leq c \exp(c'a^{d-1}) \sqrt{|\mathcal{I}|} \leq c'' \exp(c'a^{d-1})a^{d/2} \frac{d^2}{a_n^2}. \]

Here we have used the fact that \(\text{Var}(f(\{X_k\})) \leq 1/4\) and \(|\mathcal{I}| = \Theta((a_n/a)^d)\).

On the event \(\{ B_0 \leftrightarrow_{2r} B(2n) \} \cap F^c\), we can find \(k \geq 2\) disjoint \(r\)-clusters \(C_1, \ldots, C_k\) in \(B(2n) - B_0\) and an \(r\)-cluster \(\overline{C}\) (which may be the same as one of the \(r\)-clusters \(C_1, \ldots, C_k\) in \(B(2n) - B_0\) such that

(i) each of \(C_1, \ldots, C_k\) has a point in \(B_0^{(r)}\) and a point in \(B(2n)_{(2r)}\),

(ii) there is an \(r\)-cluster in \(B(n/2) - B_0\), call it \(\overline{C}'\), which is contained in \(\overline{C} \cap B(n/2)\), such that \(\overline{C}'\) has a point in \(B(2a_n)^{(r)}\) and a point in \(B(n/2)_{(2r)}\) but does not have a point in \(B_0^{(r)}\).

If \(k \geq 3\), then the event \(\{ B(2a_n) \rightarrow_{3r} B(n/2) \}\) is trivially true. If \(k = 2\), then we can find two disjoint \(r\)-clusters \(C_1'\) and \(C_2'\) in \(B(n/2) - B_0\) which are contained in \(C_1 \cap B(n/2)\) and \(C_2 \cap B(n/2)\) respectively such that \(C_1'\) and \(C_2'\) satisfy the requirements for \(\{ B_0 \leftrightarrow_{2r} B(n/2) \}\) to be true. Further, \(\overline{C}'\) is different from \(C_1'\) and \(C_2'\) since \(\overline{C}'\) does not have a point in \(B_0^{(r)}\). Hence the restrictions of \(\overline{C}\), \(C_1'\) and \(C_2'\) to \(B(n/2) - B(2a_n)\) will contain three disjoint \(r\)-clusters satisfying the requirements for \(\{ B(2a_n) \rightarrow_{3r} B(n/2) \}\) to be true.

Hence we have

\[ \mathbb{P}(B_0 \leftrightarrow_{2r} B(2n)) \leq \mathbb{P}(F) + \mathbb{P}(B(2a_n) \rightarrow_{3r} B \left( \frac{n}{2} \right)). \]

All we need now is an upper bound for the second term on the right side. We would like to apply a Burton-Keane type argument to get a bound for this term.

Assume that \(C_1\), \(C_2\) and \(C_3\) are three disjoint \(r\)-clusters in \(B(n/2) - B(2a_n)\) such that \(R(C_j)\) intersects both \(B(n/2)^{(r)}\) and \(B(2a_n)^{(r)}\) and let \(x_j\) be the point in \(C_j\) closest to \(B(2a_n)\) for \(j = 1, 2, 3\).

If \(x_j \in B(2a_n)^{(r)}\) for every \(j\) then \(B(2a_n) \leftrightarrow_{3r} B(n/2)\) holds true and if \(x_j \in B(2a_n)^{(2r)} - B(2a_n)^{(r)}\) for every \(j\) then \(B(2a_n)^{(r)} \leftrightarrow_{3r} B(n/2)\) holds true.
Assume now that the event
\[ E = \left\{ B(2a_n) - \frac{3}{r} B \left( \frac{n}{2} \right) \right\} \]
\[ \cap \left\{ B(2a_n) - \frac{3}{r} B \left( \frac{n}{2} \right) \right\} \cup \left\{ B(2a_n)^{(r)} - \frac{3}{r} B \left( \frac{n}{2} \right) \right\}^c \]
is true, then the number of \( x_i \)'s in \( B(2a_n)^{(r)} - B(2a_n) \) is one or two.

Let us assume that \( x_1, x_2 \in B(2a_n)^{(r)} \) and \( x_3 \in B(2a_n)^{(2r)} - B(2a_n)^{(r)} \) (the other possibilities can be handled similarly). We can find a sequence of points \( z_1^{(j)}, \ldots, z_{k_j}^{(j)} \) in \( C_j \) for \( j = 1, 2 \) such that

(i) \( z_1^{(j)} \in B(2a_n)^{(r)} \) and \( z_i^{(j)} \notin B(2a_n)^{(r)} \) if \( i \geq 2 \),

(ii) \( z_{k_j}^{(j)} \in B(n/2)(2r) \),

(iii) \( d(z_i^{(j)}), z_{i+1}^{(j)} \) \( \leq 2r \) for \( 1 \leq i \leq k_j - 1 \) and

(iv) \( d(z_i^{(j)}), z_{i'}^{(j)} \) > \( 2r \) whenever \( i' \geq i + 2 \).

Let \( C_j^1(\subset C_j) \) be the \( r \)-cluster in \( B(n/2) - B(2a_n)^{(r)} \) containing \( \{z_2^{(j)}, \ldots, z_{k_j}^{(j)}\} \).

Note that
\[ \max_{j=1,2} d(z_1^{(j)}, z_2^{(j)}) > r, \]
because otherwise the event \( \{ B(2a_n)^{(r)} - \frac{3}{r} B(n/2) \} \) will be true (the \( r \)-clusters \( C_1^j, C_2^j \) and \( C_3 \) will satisfy the requirements). If \( \min_{j=1,2} d(z_1^{(j)}, z_2^{(j)}) \leq r \) then \( E_1(z_1^{(1)}), E_1(z_1^{(2)}) \) holds where
\[ E_1(x) := \left\{ B(2a_n)^{(r)} \cup S(x, r) - \frac{3}{r} B \left( \frac{n}{2} \right) \right\} \]
for \( x \in B(2a_n)^{(r)} \) and if \( \min_{j=1,2} d(z_1^{(j)}, z_2^{(j)}) > r \) then the event
\[ E_2(z_1^{(1)}, z_1^{(2)}) := \left\{ B(2a_n)^{(r)} \cup S(z_1^{(1)}, r) \cup S(z_1^{(2)}, r) - \frac{3}{r} B \left( \frac{n}{2} \right) \right\}, \]
holds; in each case, \( C_3 \) and the appropriate \( r \)-clusters containing the points \( \{z_2^{(j)}, \ldots, z_{k_j}^{(j)}\} \) (\( j = 1, 2 \)) satisfying the requirements. Hence

\[ \mathbb{P} \left( B(2a_n) - \frac{3}{r} B \left( \frac{n}{2} \right) \right) \]
\[ \leq \mathbb{P} \left( B(2a_n) - \frac{3}{r} B \left( \frac{n}{2} \right) \right) + \mathbb{P} \left( B(2a_n)^{(r)} - \frac{3}{r} B \left( \frac{n}{2} \right) \right) \]
\[ + \mathbb{P} \left( \exists x, y \in \mathcal{P} \cap (B(2a_n)^{(r)} - B(2a_n)) \text{ such that } \right. \]
\[ x \neq y \text{ and } E_2(x, y) \text{ holds} \]
\[ + \mathbb{P} \left( \exists x \in \mathcal{P} \cap (B(2a_n)^{(r)} - B(2a_n)) \text{ such that } \right. \]
\[ E_1(x) \text{ holds} \].
This gives

\begin{equation}
\Pr \left( B(2a_n) \overset{3}{\rightarrow} B \left( \frac{n}{2} \right) \right) \\
\leq \Pr \left( B(2a_n) \overset{3}{\leftrightarrow} B \left( \frac{n}{2} \right) \right) + \Pr \left( B(2a_n)^{(r)} \overset{3}{\rightarrow} B \left( \frac{n}{2} \right) \right) \\
+ \mathbb{E} |\mathcal{P} \cap (B(2a_n)^{(r)} - B(2a_n))| \sup_1 \Pr \left( E_2(x, y) \right) \\
+ \mathbb{E} |\mathcal{P} \cap (B(2a_n)^{(r)} - B(2a_n))| \sup_2 \Pr \left( E_1(x) \right)
\end{equation}

where \( \sup_1 \) (resp. \( \sup_2 \)) is supremum taken over all \( x, y \) (resp. \( x \)) in \( B(2a_n)^{(r)} - B(2a_n) \). The following lemma helps us in estimating \( \Pr(E_2(x, y)) \) and \( \Pr(E_1(x)) \).

**Lemma 6.2.** Fix \( r \in [r_1, r_2] \) and two nonnegative numbers \( x \) and \( y \) such that \( x + y > 2r_2 \). Then there exist positive constants \( c \) and \( c' \) depending only on \( r_1, r_2 \) and the dimension \( d \) such that for every \( m > 100(x + y) \)

\begin{equation}
\Pr(B(x)^{(y)} \overset{3}{\leftrightarrow} B(m)) \leq c \cdot \exp \left( c'(x + y) \right) \frac{(x + y)^d}{m}.
\end{equation}

For \( z_1, z_2 \in B(x)^{(y)} \), the same bound holds for \( \Pr(B(x)^{(y)} \cup S(z_1, r) \overset{3}{\leftrightarrow} B(m)) \) and \( \Pr(B(x)^{(y)} \cup S(z_1, r) \cup S(z_2, r) \overset{3}{\leftrightarrow} B(m)) \).

**Proof.** Let \( K \subset B(m) \) be a translate of \( B(x)^{(y)} \). We will say that \( K \) is a trifurcation box in \( B(m) \) (in short “\( K \) T-box in \( B(m) \)” at level \( r \) (Figure 2).

![Figure 2. K is a trifurcation box in B(m).](image-url)
if

(i) there is an \( r \)-cluster \( C \) in \( B(m) \) with \( C \cap K \neq \emptyset \) and

(ii) \( C \cap K^c \) contains at least three disjoint \( r \)-clusters in \( B(m) - K \) each having a point in \( B(m)(2r) \).

Then,

\[
\Pr(B(x)(y) \text{T-box in } B(m)) \geq \Pr(B(x)(y) \xleftarrow{3/r} B(m)) \cdot \Pr(B(x)(y) \text{T-box in } B(m) | B(x)(y) \xleftarrow{3/r} B(m)).
\] (6.12)

Now, given any \( \eta \in X(B(m) - B(x)(y)) \) for which the event \( \mathcal{E} := \{B(x)(y) \xleftarrow{3/r} B(m)\} \) is true, we can ensure that the event \( \{B(x)(y) \text{T-box in } B(m)\} \) happens just by placing enough Poisson points inside \( B(x)(y) \) so that at least three of the \( r \)-clusters in \( B(m) - B(x)(y) \) satisfying the requirements for \( \mathcal{E} \) to be true get connected to form a single component. Since this can be done by placing at least one Poisson point in each of at most \( 6d^{3/2}(x + y)/r_1 \) cubes (of side length \( r_1/\sqrt{d} \)) inside \( B(x)(y) \)

\[
\Pr(B(x)(y) \text{T-box in } B(m) | B(x)(y) \xleftarrow{3/r} B(m)) \geq \exp(-c(x + y))
\]

for a positive universal constant \( c \) depending only on \( r_1 \) and \( d \). Plugging this into (6.12), we get

\[
\Pr(B(x)(y) \xleftarrow{3/r} B(m)) \leq \exp(c(x + y)) \cdot \Pr(B(x)(y) \text{T-box in } B(m)).
\] (6.13)

Let us define

\[
\mathcal{T} := \{j \in \mathbb{Z}^d : 4(x + y)j + B(x)(y) \subset B(m/4)\}
\]

and for \( j \in \mathcal{T} \) denote \( 4(x + y)j + B(x)(y) \) by \( K_j \). The following Lemma can be proven by following the arguments in the proof of Lemma 3.2 in [27].

**Lemma 6.3.** Let \( R \) be a finite non empty subset of a set \( S \). Assume further that

(I) for every \( r \in R \), there exist pairwise disjoint subsets (which we call “branches”) \( C_r^{(1)}, \ldots, C_r^{(m_r)} \) of \( S \) such that

(Ia) \( m_r \geq 3 \),

(Ib) \( r \notin C_r^{(i)} \) for \( i \leq m_r \) and

(Ic) \( |C_r^{(i)}| \geq k \) for \( i \leq m_r \);
(II) for all \( r, r' \in R \), either

\((IIa)\quad \left( \bigcup_{j \leq m_r} C_r^{(j)} \cup \{ r \} \right) \cap \left( \bigcup_{i \leq m_{r'}} C_{r'}^{(i)} \cup \{ r' \} \right) = \emptyset \) or

\((IIb)\quad \left( \bigcup_{j \leq m_r} C_r^{(j)} \cup \{ r \} \right) - C_r^{(j_0)} \subset C_r^{(i_0)} \) and

\[ \left( \bigcup_{i \leq m_{r'}} C_{r'}^{(i)} \cup \{ r' \} \right) - C_{r'}^{(i_0)} \subset C_{r'}^{(j_0)} \] for some \( i_0 \leq m_{r'} \) and \( j_0 \leq m_r \).

Then \(|S| \geq k|R|\).

We have the following

**Lemma 6.4.** For some positive constant \( c \) depending on \( r_2 \)

\[ |\{ \mathcal{P} \cap B(m/2) \}| \geq cm |\{ j \in \mathcal{T} : K_j \text{ T-box in } B(m/2) \}|. \]  

**Proof.** Set \( S = \mathcal{P} \cap B(m/2) \). If \( K_j \) is a trifurcation box in \( B(m/2) \) for some \( j \in \mathcal{T} \), then there is an \( r \)-cluster \( C_j \) in \( B(m/2) \) such that there is a point \( r_j \) in \( C_j \cap K_j \). Further, \( C_j \cap B(m/2) - K_j \) contains \( m_j (\geq 3) \) disjoint \( r \)-clusters, say \( C_j^{(1)}, \ldots, C_j^{(m_j)} \) each having a point in \( B(m/2)(2r) \), call these clusters the “branches” of \( r_j \). Set \( R = \{ r_j : j \in \mathcal{T}, K_j \text{ T-box in } B(m/2) \} \).

For any \( r_j, r_{j'} \in R \), condition \((IIa)\) of Lemma 6.3 holds if \( C_j \) and \( C_{j'} \) are disjoint and condition \((IIb)\) holds otherwise. Also

\[ |C_j^{(i)}| \geq \frac{m/4 - 2r_2}{2r_2} \geq cm \]

for every \( r_j \in R \) and \( i \leq m_j \). Hence an application of Lemma 6.3 yields the result. \( \square \)

Taking expectation in (6.14),

\[ \frac{m^{d-1}}{c^{2d}} \geq \sum_{j \in \mathcal{T}} \mathbb{P}(K_j \text{ T-box in } B(m/2)) \]

\[ \geq \sum_{j \in \mathcal{T}} \mathbb{P}(K_j \text{ T-box in } 4(x + y)j + B(m)). \]

By translational invariance and the fact that \( |\mathcal{T}| : (x + y)^d = \Theta(m^d) \), we get

(6.15) \[ c'm^{d-1} \geq \frac{m^d}{(x + y)^d} \mathbb{P}(B(x)^{(y)} \text{ T-box in } B(m)) \]

and (6.11) follows if we plug this in (6.13).

The same type of arguments work when \( B(x)^{(y)} \) is replaced by the other sets, so we do not repeat them. Thus we have proven Lemma 6.2. \( \square \)

From (6.7), (6.8), (6.10) and Lemma 6.2 we get

(6.16) \[ \mathbb{P}(B_0 \overset{2}{\leftrightarrow} B(2n)) \leq c \left( \exp(c'' a^{d-1}) a^{d/2} + \exp(c'' a_n) a_n^{3d-2} \right). \]
We choose \( a_n \) so that \( c^na_n = \frac{1}{2} \log n \), plug this into (5.16) and finally replace \( n \) by \( n/2 \) to get (5.1).

If we replace \( B(a) \) in (5.1) by, say, \( K = B(a)^{(r)} \cup S(x, r) \), define \( B_k := 2(a + 2r)k + K \) so that the sets \( B_k \) remain disjoint. Define \( \mathcal{I} \) as before and think of \( f \) as a function of the configurations inside \( \{B_k\}_{k \in \mathcal{I}} \) and the configuration in the complement of \( \bigcup_{k \in \mathcal{I}} B_k \). The rest of the proof can be carried out by following the same arguments as before.

6.2. Proof of Theorem 2.1. First let us state a Lemma which collects the estimates in different regimes together.

**Lemma 6.5.** For positive numbers \( r_1, r_2 \) satisfying \( r_1 < r_c(d) \) and \( r_2 > r'_c := \lim_{L \to \infty} r_c(\mathbb{R}^d_+ \times [0, L]^{d-2}) \), we have the following estimates.

(i) When \( d = 2 \) and \( a \in (1/2, \log n) \),

\[
\mathbb{P}(B(a)^{(r)} \to B(n)) \leq \begin{cases} 
    c_{11} \exp(c_{12}a) \exp(-c_{13}n), & \text{if } r \leq r_1, \\
    c_{14}/n^{\beta_1}, & \text{if } r_1 < r \leq (\log n)^2, 
\end{cases}
\]

for a universal constant \( \beta_1 \). The same bound holds for \( \mathbb{P}(B(a)^{(r)} \to B(n) \text{ in } K) \) for a square \( K \) which contains the origin and \( K \cap \partial B(n) \neq \emptyset \).

(ii) When \( d \geq 3 \) and \( a \in (1/2, \log \log n) \),

\[
\mathbb{P}(B(a)^{(r)} \to B(n)) \leq \begin{cases} 
    c_{16} \exp(c_{17}a) \exp(-c_{18}n) & \text{if } r \leq r_1, \\
    c_{19}a^{3d-1} \exp(c_{20}d^{-1} \log n)/r^{d-1} & \text{if } r \in [r_1, r_2], \\
    c_{20} \exp(-c_{21}n) & \text{if } r_2 \leq r \leq n/8. 
\end{cases}
\]

The constants appearing here depend only on \( r_1, r_2 \) and \( d \). When \( r \leq r_1 \) or \( r_2 \leq r \leq n/8 \) and \( d \geq 3 \), the same bounds hold for \( \mathbb{P}(B(a)^{(r)} \to B(n) \text{ in } K) \) for a cube \( K \) which contains the origin and \( K \cap \partial B(n) \neq \emptyset \).

**Proof.** Note that for any \( d \geq 2 \),

\[
\{B(a)^{(r)} \to B(n)\} \subset \{B(a) \leftarrow \frac{1}{r} \to B(n)\}
\]

\[
\subset \{B(a) \leftarrow \frac{1}{r} \to B(n)\} \cup \{B(a)^{(r)} \leftarrow \frac{1}{r} \to B(n)\}.
\]

That the last inclusion holds can be seen as follows. Consider an \( r \)-cluster \( \mathcal{C} \) in \( B(n) - B(a) \) which has a point in both \( B(a)^{(2r)} \) and \( B(n)^{(2r)} \) and let \( x \in \mathcal{C} \) be the point closest to \( B(a) \). If \( x \in B(a)^{(r)} \) then \( \{B(a) \leftarrow \frac{1}{r} \to B(n)\} \) is true and if \( x \in B(a)^{(2r)} \) then \( \{B(a)^{(r)} \leftarrow \frac{1}{r} \to B(n)\} \) is true.

For any \( r \leq r_1 \), \( \{B(a) \leftarrow \frac{1}{r_1} \to B(n)\} \subset \{B(a) \leftarrow \frac{1}{r_1} \to B(n)\} \) and a similar statement holds if we replace \( B(a) \) by \( B(a)^{(r)} \). If we fix a configuration in \( B(n) - B(a) \) (resp. \( B(n) - B(a)^{(r)} \)) for which \( \{B(a) \leftarrow \frac{1}{r_1} \to B(n)\} \) (resp. \( \{B(a)^{(r)} \leftarrow \frac{1}{r_1} \to B(n)\} \)) holds, we can connect any of the corresponding clusters to the origin by placing at least one Poisson point in at most \( c(a + r_2)/r_1 \).
many cubes inside $B(a)$ (resp. $B(a)^{(r)}$) each of side length $\min(2a, r_1/\sqrt{d})$. Thus, if $\mu_P$ is the probability measure corresponding to a Poisson process of intensity one conditioned to have a point at the origin, then

$$\mu_P(\text{diameter}(C_0) \geq n \text{ at level } r_1 | B(a) \overset{1}{\leftrightarrow}_{r_1} B(n)) \geq c \exp(-c'a),$$

$C_0$ being the occupied component containing the origin. A similar inequality holds for $\mu_P(\text{diameter}(C_0) \geq n \text{ at level } r_1 | B(a)^{(r)} \overset{1}{\leftrightarrow}_{r_1} B(n))$. Hence, from (6.18), we get

$$\mathbb{P}(B(a) \overset{2}{\leftrightarrow}_r B(n)) \leq \mathbb{P}(B(a) \overset{1}{\leftrightarrow}_{r_1} B(n)) + \mathbb{P}(B(a)^{(r)} \overset{1}{\leftrightarrow}_{r_1} B(n))$$

$$\leq c \exp(c'a) \mu_P(\text{diameter}(C_0) \geq n \text{ at level } r_1)$$

$$\leq c \exp(c'a) \exp(-c''n).$$

The last inequality is just an application of (3.60) in [27].

When $r \in [r_1, r_c]$ and $d = 2$

$$\mathbb{P}(B(a) \overset{2}{\leftrightarrow}_r B(n)) \leq \mathbb{P}(B(a) \overset{1}{\leftrightarrow}_{r_c} B(n)) + \mathbb{P}(B(a)^{(r)} \overset{1}{\leftrightarrow}_{r_c} B(n))$$

$$\leq c/n^\theta, \text{ for some } \theta > 0.$$

To see that the last inequality holds, first note that

$$\mathbb{P}(\exists \text{ a vacant left-right crossing of } [0, \ell] \times [0, 3\ell] \text{ at level } r_c) \geq \frac{1}{(9e)^{122}}$$

for $\ell \geq r_c$ by arguments similar to the ones in the proof of Theorem 4.5 of [27]. This inequality together with Lemma 4.4 of [27] and the RSW lemma for vacant crossings (see [33] or Theorem 4.2 in [27]) will yield

$$\mathbb{P}(\exists \text{ a vacant left-right crossing of } [0, 3\ell] \times [0, \ell] \text{ at level } r_c) \geq \delta$$

for a positive constant $\delta$ and every $\ell$ bigger than a fixed threshold $\ell_0$. It then follows from standard arguments that with probability at least $1 - c/n^\theta$, a vacant circuit around $B(a + r_c)$ exists in $B(n)$ at level $r_c$. Hence we get the desired upper bound on $\mathbb{P}(B(a) \overset{2}{\leftrightarrow}_r B(n))$ for $r \in [r_1, r_c]$.

When $r \geq r_2$, the polynomial decay of $\mathbb{P}(B(a) \overset{2}{\leftrightarrow}_r B(n))$ follows from the existence of occupied “circuits” at level $r_c$ around $B(a)$. The argument for this is also standard, we will give an outline in Appendix [A].

Next we turn to the case when dimension is bigger than two. Fix $r \in [r_1, r_2]$ and assume $\{B(a) \overset{2}{\leftrightarrow}_r B(n)\}$ holds. Take any two disjoint clusters $C_1$ and $C_2$ in $B(n) - B(a)$ each having a point in $B(a)^{(2r)}$ and $B(n)^{(2r)}$ and let $x_j \in C_j$ be the point closest to $B(a)$. If $x_j \in B(a)^{(r)}$ for $j = 1, 2$ then the event $\{B(a) \overset{2}{\leftrightarrow}_r B(n)\}$ is true and if $x_j \in B(a)^{(2r)} - B(a)^{(r)}$ for $j = 1, 2$ then the event $\{B(a)^{(r)} \overset{2}{\leftrightarrow}_r B(n)\}$ is true.
Now, assume that the event
\[ \{ B(a) \xrightarrow{2}{r} B(n) \} \cap \left[ \{ B(a) \leftrightarrow r B(n) \} \cup \{ B(a) \leftrightarrow r B(n) \} \right]^c \]
is true. Then each of the sets \( B(a)^{(r)} \) and \( B(a)^{(2r)} - B(a)^{(r)} \) contain exactly one of the points \( x_1 \) and \( x_2 \).

By arguments similar to the ones leading to (6.10), we can show that in this case the event
\[ E := \{ \exists x \in \mathcal{P} \cap (B(a)^{(r)} - B(a)) \text{ such that } S(x, r) \cup B(a)^{(r)} \leftrightarrow r B(n) \} \]
is true. For any realization \( \eta = \{ \eta_1, \ldots, \eta_\ell \} \) of \( \mathcal{P} \cap (B(a)^{(r)} - B(a)) \),
\[ E \subset \bigcup_{j=1}^\ell \{ S(\eta_j, r) \cup B(a)^{(r)} \leftrightarrow r B(n) \}. \]

Hence from Lemma 5.1,
\[ \mathbb{P}(E) \leq \frac{c_8 \exp(c_9 a^{d-1}) a^{d/2}}{(\log n)^{2a}} \mathbb{P}(\mathcal{P} \cap (B(a)^{(r)} - B(a))) \]
\[ \leq \frac{c \exp(c_9 a^{d-1}) a^{3d/2 - 1}}{(\log n)^{d}}. \]

From our earlier discussion and another application of Lemma 5.1
\[ \mathbb{P}(B(a) \xrightarrow{2}{r} B(n)) \leq \mathbb{P}(B(a) \leftrightarrow r B(n)) \]
\[ + \mathbb{P}(B(a)^{(r)} \leftrightarrow r B(n)) + \mathbb{P}(E) \]
\[ \leq c \cdot a^{3d/2 - 1} \exp(c_9 a^{d-1}) \]
\[ \frac{1}{(\log n)^{d}}. \]

When \( n/8 \geq r \geq r_2 \), the exponential decay in (6.17) can be proven using the same type of arguments as in the lattice case (e.g. Lemma 7.89 in [21]). A brief sketch of the proof highlighting the parts which are slightly different in the continuum case is given in Appendix A.

Finally let us mention that the bounds for \( \mathbb{P}(B(a) \xrightarrow{2}{r} B(n) \text{ in } K) \) follow from similar arguments and we do not repeat them.

Let us now recall that the Kantorovich-Wasserstein distance between two probability measures \( \mu_1 \) and \( \mu_2 \) on \( \mathbb{R} \) is given by
\[ \mathcal{W}(\mu_1, \mu_2) := \sup \left\{ \left| \int f \, d\mu_1 - \int f \, d\mu_2 \right| : f \text{ Lipschitz with } \|f\|_{\text{Lip}} \leq 1 \right\}. \]

Convergence in this metric implies weak convergence.

We now state our main approximation theorem from [18]. To state the theorem we need some notations, we will use these notations repeatedly in this paper.
Let \( \mathcal{X} \) be a measure space and \( X = (X_1, \ldots, X_n) \) and \( X' = (X'_1, \ldots, X'_n) \) be two identically distributed \( \mathcal{X}^n \) valued random vectors such that the components of \( X \) and \( X' \) are independent. For every \( A \subset [n] := \{1, \ldots, n\} \) define the random vector \( X^A \) with components

\[
X_i^A = \begin{cases} 
X'_i, & \text{if } i \in A, \\
X_i, & \text{if } i \notin A.
\end{cases}
\]

We will simply write \( X^j \) instead of \( X^\{j\} \). Now, for a measurable function \( f : \mathcal{X}^n \rightarrow \mathbb{R} \) define

\[
\Delta_j f(X) := f(X) - f(X^j)
\]

and for every \( A \subset [n] \), let

\[
T_A := \sum_{j \notin A} \Delta_j f(X) \Delta_j f(X^A).
\]

Finally, let

\[
T = \frac{1}{2} \sum_{A \subset [n]} \frac{T_A}{\binom{n}{|A|}(n - |A|)}.
\]

Then we have

**Theorem 6.6.** \([18]\) Let all terms be defined as above and let \( W = f(X) \) with \( \sigma^2 := \text{Var}(W) < \infty \). Then \( \mathbb{E}T = \sigma^2 \) and

\[
(6.21) \quad W(\mu, G) \leq \frac{[\text{Var}(\mathbb{E}(T|W))]^{1/2}}{\sigma^2} + \frac{1}{2\sigma^3} \sum_{j=1}^{n} \mathbb{E}|\Delta_j f(X)|^3
\]

where \( \mu \) is the law of \( (W - \mathbb{E}W)/\sigma \).

Note that \( \text{Var}(\mathbb{E}(T|W)) \leq \text{Var}(T) \). This is the bound we will work with.

If \( \mu \) is a probability measure on \( \mathbb{R} \) and \( G \) is the standard Gaussian measure on \( \mathbb{R} \), then we have

\[
(6.22) \quad D(\mu, G) \leq 2\sqrt{W(\mu, G)}.
\]

A proof of this simple inequality can be found in \([19]\). We will work with the metric \( W(\cdot, \cdot) \) to measure distance from normality and then switch to the metric \( D(\cdot, \cdot) \) by an application of \((6.22)\).

With this, we are ready to prove **Theorem 2.1**

**Proof of Theorem 2.1.** Choose an integer \( k \) such that \( (n - 1)/2 \geq k \geq (n - 2)/4 \) and let \( s = n/(2k + 1) \). Write \( \mathbb{R}^d \) as the union of cubes,

\[
\mathbb{R}^d = \bigcup_{j \in \mathbb{Z}^d} B_j \text{ where } B_j := 2sj + B(s).
\]

Let \( B(n) = \bigcup_{j \in \mathbb{L}} B_j \), clearly \( \ell := |\mathbb{L}| = \Theta(n^d) \).

Consider two independent Poisson process \( \mathcal{P} \) and \( \mathcal{P}' \) (having intensity one) in \( \mathbb{R}^d \). We will apply \((6.21)\) with \( X_j := \mathcal{P} \cap B_j, X'_j := \mathcal{P}' \cap B_j, \)
\( X := \{ X_j \}_{j \in \mathcal{L}} \), \( X' := \{ X'_j \}_{j \in \mathcal{L}} \) and the function \( f : \prod_{i \in \mathcal{L}} X(B_i) \to \mathbb{R} \) given by
\[
f(\{ \omega_i \}_{i \in \mathcal{L}}) = M \left( \bigcup_{i \in \mathcal{L}} \omega_i \right).
\]
To simplify notations, we will not distinguish between the functions \( f \) and \( M \). In particular, we will write \( M(X^A) \) and \( \Delta_j M(X) \) to mean \( f(X^A) \) and \( \Delta_j f(X) \) respectively.

When dealing with Euclidean minimal spanning trees, we would like to have a criterion which ensures that if we fix a small cube, then there are no “long” edges in the MST with one endpoint inside that cube. Kesten and Lee [23] used the idea of a separating set to meet this purpose.

**Definition 6.7 (Separating set [23])**. Fix \( \delta > 0 \) and \( x \in \mathbb{R}^d \). A separating set \( S \) of width \( \delta \) for the cube \( B(x, a) \) is a finite set in the annulus \( B(x, 1 + a + \delta) \setminus B(x, 1 + a) \) such that any line segment with one endpoint in \( \partial B(x, 1 + a + \delta) \) and other endpoint in \( \partial B(x, 1 + a + \delta) \) passes within distance \( 1/3 \) of some point of \( S \).

In a similar fashion they also define “\( n \)-separating sets” . They showed that separating sets have the following

**Property A (Lemma 3, [23])**. If \( S \subset B(x, 1 + a + \delta) \setminus B(x, 1 + a) \) is a separating set for \( B(x, a) \), then in any MST of the complete graph on a finite subset \( A \) of \( \mathbb{R}^d \) containing \( S \), there are no edges between vertices in \( B(x, a) \) and vertices in \( \mathbb{R}^d \setminus B(x, a + 2 + \delta) \)

and a similar statement holds for \( n \)-separating sets.

Further, they showed that the tail of \( \tilde{\delta} \) decays exponentially fast where \( \tilde{\delta} \) is the infimum of all \( \delta \) such that a Poisson process in \( \mathbb{R}^d \) has a separating set for \( B(x, a) \) of width \( \delta \) (Lemma 5 in [23]). This is enough to prove that the central limit theorem holds. Since we are interested in the rate of convergence, we will relax the criterion in the definition of separating sets to get a faster convergence rate.

**Definition 6.8.** Assume that \( b > a \) are positive numbers and \( x \in \mathbb{R}^d \) and let \( K \) be a cube containing \( B(x, a) \). Further assume that \( K \cap \partial B(x, b) \neq \emptyset \).

We say that a subset \( \mathcal{W} \) of \( \mathbb{R}^d \) contains a \( K \)-wall around \( B(x, a) \) in \( B(x, b) \) if the following holds :

For any line segment \( p_1p_2 \) of length \( \ell \) with
\[
p_1 \in \partial B(x, a) \text{ and } p_2 \in K \cap \partial B(x, b), \text{ the set } K \cap \mathcal{W} \cap S(p_1, 3\ell/4) \cap S(p_2, 3\ell/4) \cap \{ B(x, b) - B(x, a) \}
\]
is nonempty.

If \( B(x, b) \subset K \), we will simply say \( \mathcal{W} \) contains a wall around \( B(x, a) \) in \( B(x, b) \).

The following simple Lemma shows that walls satisfy Property A.
Lemma 6.9. Let \( \omega \) be a finite set of points in \( K \subset \mathbb{R}^d \) and consider the complete graph \((V, E)\) on \( \omega \) with edge weights being the Euclidean length of edges. If \( \omega \) contains a \( K \)-wall around \( B(x, a) \) in \( B(x, b) \), then an edge in \( E \) with one endpoint in \( B(x, a) \) and other endpoint in \( B(x, b)^c \) is not included in any MST of \((V, E)\).

Before proving this Lemma, let us recall a simple fact about minimal spanning trees which we will use repeatedly.

Lemma 6.10. Consider a weighted finite graph \( G = (V, E, w) \). Let \( T \) be a minimal spanning tree of \( G \). Then any path \((x_0, \ldots, x_n)\) with \( x_i \in V \) and \( \{x_i, x_{i+1}\} \in E \) is minimax if all of the edges \( \{x_i, x_{i+1}\} \) are contained in \( T \), i.e.

\[
\max_i w(\{x_i, x_{i+1}\}) \leq \max_j w(\{x'_j, x'_{j+1}\})
\]

for any path \((x'_0, \ldots, x'_m)\) with \( \{x'_j, x'_{j+1}\} \in E \) and \( x_0 = x'_0 \) and \( x_n = x'_m \).

Proof. This is just a restatement of Lemma 2 in [23]. □

Proof of Lemma 6.9. Let \( y_1, y_2 \) be two points in \( \omega \) such that \( y_1 \in B(x, a) \) and \( y_2 \in B(x, b)^c \). Assume that \( p_1 \in \partial B(x, a) \) and \( p_2 \in \partial B(x, b) \) are points on the line segment \( \overline{y_1y_2} \). We will denote the length of a line segment \( \overline{x_1x_2} \) by \( \ell(\overline{x_1x_2}) \) and let \( \ell := \ell(\overline{p_1p_2}) \).

Since \( \omega \) contains a \( K \)-wall around \( B(x, a) \) in \( B(x, b) \), we can find a point \( z \in \omega \cap S(p_1, 3\ell/4) \cap S(p_2, 3\ell/4) \cap (B(x, b) - B(x, a)) \). Then

\[
\ell(\overline{y_1z}) \leq \ell(\overline{y_1p_1}) + \ell(\overline{p_1z})
\]

\[
\leq \ell(\overline{y_1p_1}) + 3\ell/4
\]

\[
< \ell(\overline{y_1p_1}) + \ell(\overline{p_1y_2}) = \ell(\overline{y_1y_2})
\]

and similarly \( \ell(\overline{zy_2}) < \ell(\overline{y_1y_2}) \).

Hence, it follows from Lemma 6.10 that \( \overline{y_1y_2} \) will not be included in any minimal spanning tree of \((V, E)\). □
Next we show that a wall exists in a large annulus with high probability.

**Lemma 6.11.** For any $a_0 > 0$, there exist constants $c$ and $c'$ depending only on $a_0$ and $d$ such that for every $a \leq a_0$ for which $B(x, a) \subset B(n)$,

\[
\mathbb{P}(\mathcal{P} \text{ does not contain a } B(n)-\text{wall around } B(x, a) \text{ in } B(x, b)) \leq c(\max(1, b))^{d-1} \exp(-c'b^d)
\]

for any $b$ satisfying $b > a$ and $B(n) \cap \partial B(x, b) \neq \emptyset$.

**Proof.** It suffices to prove the claim for large values of $b$, so let us start with the assumption $b > 4a_0 + 16$.

Cover $B(n) \cap \partial B(x, b)$ by $(d - 1)$ dimensional cubes, $\{Q^1_i\}_{i \leq m_1}$ of diameter one, this can be done in a way so that the total number of cubes $m_1$ is at most $c b^{d-1}$. Similarly cover $\partial B(x, a)$ by $(d - 1)$ dimensional cubes $\{Q^2_i\}_{i \leq m_2}$ of diameter $\min(1, 2a \sqrt{d-1})$ so that the total number of cubes, $m_2$ is at most $c \max(1, a_0^{d-1})$.

Let $p_1', p_2'$ be two points on $\partial B(x, a)$ and $B(n) \cap \partial B(x, b)$ respectively and let $z' = (p_1' + p_2')/2$ be the midpoint of $p_1' p_2'$. Let $p_1$ and $p_2$ be the midpoints of the cubes $Q^1_1$ and $Q^2_1$ such that $p_1' \in Q^1_1$ and $p_2' \in Q^2_1$. Let $z = (p_1 + p_2)/2$.

Consider $y' \in S(z', b/8)$. Then $\|y' - z'\|_\infty \leq b/8$ and hence

\[
\|z' - x\|_\infty - \frac{b}{8} \leq \|y' - x\|_\infty \leq \|z' - x\|_\infty + \frac{b}{8}.
\]

Now,

\[
\|z' - x\|_\infty + \frac{b}{8} = \left\| \frac{p_1' + p_2'}{2} - \frac{2x}{2} \right\|_\infty + \frac{b}{8} \leq \frac{a_0 + b}{2} + \frac{b}{8} < b.
\]

Also

\[
\|z' - x\|_\infty - \frac{b}{8} \geq \frac{b - a_0}{2} - \frac{b}{8} > a.
\]

Hence $S(z', b/8) \subset B(x, b) - B(x, a)$. Further, if $d(y, z) \leq b/16$, then

\[
d(y, z') \leq \frac{b}{16} + d(z, z') = \frac{b}{16} + \|p_1 + p_2 - \frac{p_1' + p_2'}{2}\|_{L^2} \leq \frac{b}{16} + 1 \leq \frac{b}{8},
\]

so $S(z, b/16) \subset S(z', b/8) \subset B(x, b) - B(x, a)$.

If $y' \in S(z', b/8)$, then

\[
d(y', p'_1) \leq d(y', z') + d(z', p'_1) \leq \frac{b}{8} + \frac{\ell(p'_1 p'_2)}{2} \leq \frac{3\ell(p'_1 p'_2)}{4}.
\]
The last inequality holds since
\[ \ell(p'_1, p'_2) \geq b - a \geq b - a_0 \geq b/2. \]
By a similar argument \( d(y', p'_2) \leq 3\ell(p'_1, p'_2)/4 \). Hence
\[ S \left( z', \frac{b}{8} \right) \subset S \left( p'_1, \frac{3\ell(p'_1, p'_2)}{4} \right) \cap S \left( p'_2, \frac{3\ell(p'_1, p'_2)}{4} \right) \cap (B(x, b) - B(x, a)). \]

Finally, we note that \( \mathbb{E} b(S(z, b/16) \cap B(n)) \geq c'b^d \). So we can conclude that
\[ \mathbb{P}(\mathcal{P} \text{ does not contain a } B(n)-\text{wall around } B(x, a) \text{ in } B(x, b)) \]
\[ \leq \mathbb{P}\left( \text{ For some } i \leq m_1, j \leq m_2, \mathcal{P} \cap B(n) \cap S(p_1 + p_2, \frac{b}{16}) = \emptyset \right) \]
\[ \leq c \max(1, a_0^{d-1})b^{d-1} \exp(-c'b^d), \]
the last inequality follows from the union bound. This proves the claim. \( \square \)

The next Lemma puts an upper bound on how much the weight of the MST changes when some points are removed.

**Lemma 6.12.** Let \( \mathcal{A} \) and \( \mathcal{B} \) be finite sets of points in \( \mathbb{R}^d \) and let \( K \) be a cube in \( \mathbb{R}^d \) such that \( \mathcal{A} \subset B(x, a) \subset K \) and \( \mathcal{B} \subset K - B(x, a) \). If \( \mathcal{B} \) contains a \( K \)-wall around \( B(x, a) \) in \( B(x, b) \), then
\[ |M(\mathcal{A} \cup \mathcal{B}) - M(\mathcal{B})| \leq c|\mathcal{A}|b \]
for some constant \( c \) depending only on \( d \). If such a wall does not exist, then
\[ |M(\mathcal{A} \cup \mathcal{B}) - M(\mathcal{B})| \leq c|\mathcal{A}|\text{diameter}(K). \]

The proof of Lemma 6.12 is similar to the proof of Lemma 7 in \[23\]. In Lemma 7 in \[23\], existence of a separating set has been assumed which is slightly stronger than existence of a wall but only Property A of separating sets has been used in the proof, so the same arguments will go through. We omit the details.

Lemma 6.12 gives us control over the tails of \( |M(\mathcal{A} \cup \mathcal{B}) - M(\mathcal{B})| \), using this we can show that all moments of this quantity are finite when the configuration comes from a Poisson process.

**Lemma 6.13.** For \( x \in \mathbb{R}^d \), \( 0 < a \leq a_0 \) and \( n \geq \max(2a_0, 1) \) for which \( B(x, a) \subset B(n) \),
\[ \mathbb{E}(|M(\mathcal{P} \cap B(n)) - M(\mathcal{P} \cap [B(n) - B(x, a)])|^q) \leq C_q \] for every \( q \geq 1 \).
The constant \( C_q \) depends only on \( a_0, d \) and \( q \).

**Proof.** Define a random variable \( Z \) as follows: if there does not exist a \( b \geq a \) such that \( \partial B(x, b) \cap B(n) \neq \emptyset \) and \( \mathcal{P} \) contains a \( B(n) \)-wall around \( B(x, a) \) in \( B(x, b) \), set \( Z = 2\sqrt{dn} \); otherwise define \( Z \) to be the infimum of all such \( b \).
From Lemma 6.11

\[ E(Z^q) = \int_0^{\sqrt{\lambda_n}} qu^{q-1}\mathbb{P}(Z > u)du \]
\[ \leq a_0^q + c \int_0^n qu^{q-1}\max(1,u)^{d-1}\exp(-c'u^d)\,du \]
\[ + cq(2n\sqrt{d})^{q+d-1}\exp(-c'n^d), \]

the right side is bounded by a constant depending only on \(a_0, d\) and \(q\). Now, from Lemma 6.12

\[ \mathbb{E}(|M(P \cap B(n)) - M(P \cap (B(n) - B(x,a)))|^q) \]
\[ \leq c\mathbb{E}(Z \cdot |P \cap B(x,a)|)^q \leq \frac{c}{2}\mathbb{E}[Z^{2q} + (|P \cap B(x,a)|)^{2q}], \]

and this finishes the proof. \(\square\)

Going back to (6.21),

\[ \Delta_j M(X) = [M(X) - M(P \cap (B(n) - B_j))] - [M(X_j) - M(P \cap (B(n) - B_j))] \]
for every \(j \in \mathcal{L}\). From Lemma 6.13 it follows that

(6.24) \[ \mathbb{E}(|\Delta_j M(X)|^q) \leq C'_q, \forall q \geq 1 \]

for constants \(C'_q\) depending only on \(d\) and \(q\). Now,

(6.25)

\[ [\text{Var}(\mathbb{E}(T|W))] \]
\[ \leq \text{Var}(T) = \frac{1}{4} \text{Var} \left[ \sum_{A \subseteq \mathcal{L}} \sum_{j \in \mathcal{L} - A} \frac{\Delta_j M(X) \Delta_j M(X^A)}{|A|(|\ell - |A|)|} \right] \]
\[ = \frac{1}{4} \sum_{A \subseteq \mathcal{L}} \sum_{j \in \mathcal{L} - A} \sum_{A' \subseteq \mathcal{L} - A'} \text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_j M(X) \Delta_j M(X^A')) \frac{\ell}{|A|(|\ell - |A|)|} \frac{\ell}{|A'|(|\ell - |A'|)|} \]

From (6.24), we see that

(6.26) \[ |\text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_j M(X) \Delta_j M(X^A'))| \leq C'_q. \]

Let us first focus on \(d \geq 3\). We plan to show that that the covariance term appearing in the numerator on the right side of (6.25) is small when \(j\) and \(j'\) are “far away”. With this in mind, we break up the summation on the
right side of (6.25) into two parts $\sum_1$ and $\sum_2$: $\sum_1$ denotes sum over all $(j, j', A, A') \in \mathcal{E}(\alpha) \ (\alpha \in (0, 1))$ where

$$\mathcal{E}(\alpha) := \{(j, j', A, A') : A, A' \subseteq \mathcal{L}; \ j \in \mathcal{L} - A, \ j' \in \mathcal{L} - A' \ \text{and either} \ \|j - j'\|_\infty \leq n^\alpha \ \text{or} \ \|2sj\|_\infty > (n - n^\alpha) \}$$

and $\sum_2$ denotes sum over the remaining terms, i.e. all $(j, j', A, A') \in \mathcal{F}(\alpha) := \{(j, j', A, A') : A, A' \subseteq \mathcal{L}; \ j \in \mathcal{L} - A, \ j' \in \mathcal{L} - A' \} - \mathcal{E}(\alpha)$.

Let $\mathcal{E}_{1,2}^{(\alpha)}$ be the collection of all $(j, j')$ for which $(j, j', A, A') \in \mathcal{E}(\alpha)$ for some $A, A' \subseteq \mathcal{L}$. Then from (6.26),

$$\sum_1 \text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})) \left(\frac{\ell}{|A|} \frac{\ell}{|A'|}\right) \leq C'_4 \sum_{(j, j') \in \mathcal{E}_{1,2}^{(\alpha)}} \sum_{A \not\subseteq j'j} \left(\frac{\ell}{|A|} \frac{\ell}{|A'|}\right)^{-1}
= C'_4 \sum_{(j, j') \in \mathcal{E}_{1,2}^{(\alpha)}} \sum_{k, k' = 0}^{l-1} \sum_{|A| = k, |A'| = k'} \left(\frac{\ell}{|A|} \frac{\ell}{|A'|}\right)^{-1}
= C'_4 |\mathcal{E}_{1,2}^{(\alpha)}| \leq c(n^{2d-1} \cdot n^{\alpha} + n^d \cdot n^{\alpha d}) \leq c'n^{2d-1+\alpha}.$$

For $(j, j') \notin \mathcal{E}_{1,2}^{(\alpha)}$, $\|2sj - 2sj'\|_\infty > 2sn^\alpha \geq 2n^\alpha$ so the cubes $B(2sj, n^\alpha)$ and $B(2sj', n^\alpha)$ are disjoint. As a result, the restrictions of $\mathcal{P}$ (and of $\mathcal{P}'$) to these cubes are independent. Let us now define

$$\tilde{\Delta}_j M(X) := M(X \cap B(2sj, n^\alpha)) - M(X^j \cap B(2sj, n^\alpha))$$

for every $j$ with $\|2sj\|_\infty \leq (n - n^\alpha)$. Whenever $(j, j', A, A') \in \mathcal{F}(\alpha)$, we have

$$\text{Cov} \left(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})\right) = \text{Cov} \left(\left[\Delta_j M(X) - \tilde{\Delta}_j M(X)\right] \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})\right) + \text{Cov} \left(\tilde{\Delta}_j M(X) [\Delta_j M(X^A) - \tilde{\Delta}_j M(X^A)], \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})\right)
+ \text{Cov} \left(\tilde{\Delta}_j M(X) \tilde{\Delta}_j M(X^A), [\Delta_{j'} M(X) - \tilde{\Delta}_{j'} M(X)] \Delta_{j'} M(X^{A'})\right)
+ \text{Cov} \left(\tilde{\Delta}_j M(X) \tilde{\Delta}_j M(X^A), \tilde{\Delta}_{j'} M(X) [\Delta_{j'} M(X^{A'}) - \tilde{\Delta}_{j'} M(X^{A'})]\right).
Let us give an upper bound for the first term on the right side of (6.29), the other terms can be dealt with in a similar fashion. Now,

\[(6.30) \quad \text{Cov} \left( [\Delta_j M(X) - \tilde{\Delta}_j M(X)] \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'}) \right) \leq \mathbb{E} \left( |[\Delta_j M(X) - \tilde{\Delta}_j M(X)] \Delta_j M(X^A) \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})| \right) + \mathbb{E} \left( |[\Delta_j M(X) - \tilde{\Delta}_j M(X)] \Delta_j M(X^A)| \mathbb{E} \left( |\Delta_{j'} M(X) \Delta_{j'} M(X^{A'})| \right) \right) =: T_1 + T_2.\]

Then for any $p$, $q > 1$ satisfying $p^{-1} + q^{-1} = 1$, we have from (6.24) that

\[(6.31) \quad T_2 \leq C_2^2 \left( \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))| \right)^{1/p} \cdot \left( \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))|^{q/2} \right)^{1/q} \leq c \left( \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))| \right)^{1/p}.\]

A similar bound holds for $T_1$. We plug all these estimates into (6.29) to get

\[(6.32) \quad \text{Cov} \left( [\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'}) \right) \leq c \left( \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))|^{1/2} + \mathbb{E} |(\Delta_{j'} M(X) - \tilde{\Delta}_{j'} M(X))|^{1/2} \right).\]

Let us write $\tilde{B}_j$ for $B(2s_j, n^\alpha)$ for every $j$ with $\|2s_j\|_\infty \leq n - n^\alpha$. For a sequence $a_n$ increasing to infinity in a way so that $a_n \leq \log \log n$, define the event $E$ as follows

\[E := \{ \mathcal{P} \text{ contains a } \tilde{B}_j-\text{wall around } B_j \text{ in } B(2s_j, a_n) \}.\]

Later, we will choose the sequence $a_n$ appropriately. Now, Lemma 6.11 and Lemma 6.13 give

\[\mathbb{E} \left( \mathbb{I}_{E^c} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))| \right) \leq \left( \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))|^2 \right)^{1/2} |\mathbb{P}(E^c)|^{1/2} \leq c a_n^{d-1} \exp(-c' a_n^d),\]

and hence

\[(6.33) \quad \mathbb{E} |(\Delta_j M(X) - \tilde{\Delta}_j M(X))| \leq c a_n^{d-1} \exp(-c' a_n^d) + \mathbb{E} \left( \mathbb{I}_E |(\Delta_j M(X) - \tilde{\Delta}_j M(X))| \right).\]

Note that the restrictions of the vectors $X$ and $X^j$ to $B(n) - B_j$ (resp. $\tilde{B}_j - B_j$) are the same, so we can write $M(B(n) - B_j)$ (resp. $M(\tilde{B}_j - B_j)$)
From now on we will work with the vector $E$ where

\[ \Delta_j M(X) - \tilde{\Delta}_j M(X) \]

\[ = \left[ (M(X) - M(B(n) - B_j)) - (M(X \cap \tilde{B}_j) - M(\tilde{B}_j - B_j)) \right] \]

\[ - \left[ (M(X^j) - M(B(n) - B_j)) - (M(X^j \cap \tilde{B}_j) - M(\tilde{B}_j - B_j)) \right], \]

and hence

\[ (6.34) \quad \mathbb{E} \left[ \mathbb{I}_E | \Delta_j M(X) - \tilde{\Delta}_j M(X) \right] \]

\[ \leq 2 \mathbb{E} \left[ \mathbb{I}_E | (M(X) - M(B(n) - B_j)) - (M(X \cap \tilde{B}_j) - M(\tilde{B}_j - B_j)) \right] \]

\[ = 2 \mathbb{E} \mathbb{E}_\eta \left[ \mathbb{I}_E | (M(X) - M(B(n) - B_j)) - (M(X \cap \tilde{B}_j) - M(\tilde{B}_j - B_j)) \right], \]

where $\mathbb{E}_\eta$ denotes expectation conditional on the event $\{P \cap B(2sj, a_n) = \eta\}$. From now on we will work with the vector $X$ only and write $M(\tilde{B}_j)$ to mean $M(X \cap \tilde{B}_j)$.

Fix realizations $\eta$, $\omega_1$ and $\omega_2$ of $P$ in $B(2sj, a_n)$, $\tilde{B}_j - B(2sj, a_n)$ and $B(n) - \tilde{B}_j$ respectively for which the event $E$ is true. If $|\eta \cap B_j| = 0$ then $M(X) - M(\tilde{B}_j - B_j)$ and $M(\tilde{B}_j) - M(\tilde{B}_j - B_j)$ are both zero. So let us assume $|\eta \cap B_j| > 0$, and write

\[ \eta \cap B_j = \{v_1, \ldots, v_m\} \quad \text{and} \quad \eta \cap [B(2sj, a_n) - B_j] = \{p_1, \ldots, p_r\}. \]

Let $\mathfrak{J}_0 = \emptyset$ and $\mathfrak{J}_i = \{v_1, \ldots, v_i\}$ for $1 \leq i \leq m$. Then

\[ (M(X) - M(B(n) - B_j)) - (M(\tilde{B}_j) - M(\tilde{B}_j - B_j)) = \sum_{i=1}^m \delta_i, \]

where

\[ (6.35) \quad \delta_i := \left[ M(\mathfrak{J}_i \cup (P \cap (B(n) - B_j))) - M(\mathfrak{J}_{i-1} \cup (P \cap (B(n) - B_j))) \right] \]

\[ - \left[ M(\mathfrak{J}_i \cup (\tilde{B}_j - B_j)) - M(\mathfrak{J}_{i-1} \cup (\tilde{B}_j - B_j)) \right]. \]

To keep notations simple, let us focus on $\delta_1$. An MST on $\{v_1\} \cup (P \cap (B(n) - B_j))$ (resp. $\{v_1\} \cup (P \cap (\tilde{B}_j - B_j))$) can be obtained from an MST on $P|_{B(n)-B_j}$ (resp. $P|_{\tilde{B}_j - B_j}$) by the add and delete algorithm (23), that is, by introducing the edges $\{v_1, p_j\}$ one by one and deleting the edge with maximum weight in the resulting cycle to make sure all paths in the new tree are minimax. We start with an MST $T_0$ (resp. $\tilde{T}_0$) on $P \cap (B(n) - B_j)$ (resp. $P \cap (\tilde{B}_j - B_j)$) with edge set $E$ (resp. $\tilde{E}$) and proceed in the following
manner.

Set \( E_0 = E \) (resp. \( \tilde{E}_0 = \tilde{E} \)), \( Y_0 = d(v_1, p_1) \) (resp. \( \tilde{Y}_0 = d(v_1, p_1) \)) and let \( w_0 \) (resp. \( \tilde{w}_0 \)) be the weight of \( T_0 \) (resp. \( \tilde{T}_0 \)). For \( k = 1, \ldots, r \)

(i) Introduce the edge \( \{v_1, p_k\} \). If \( k = 1 \), there will be no cycles

in \( E_0 \cup \{v_1, p_1\} \) (resp. \( \tilde{E}_0 \cup \{v_1, p_1\} \)). In this case, set

\( E_1 = E_0 \cup \{v_1, p_1\} \) (resp. \( \tilde{E}_1 = \tilde{E}_0 \cup \{v_1, p_1\} \)). Otherwise there will

be a unique cycle in \( E_{k-1} \cup \{v_1, p_k\} \) (resp. \( \tilde{E}_{k-1} \cup \{v_1, p_k\} \)) having

\( \{v_1, p_k\} \) as one of its edges. Delete the edge in this cycle with

maximum weight and set \( E_k \) (resp. \( \tilde{E}_k \)) to be the resulting set of

edges. If \( k \leq r - 1 \), let \( Y_k \) (resp. \( \tilde{Y}_k \)) be the maximum edge weight

in the path connecting \( v_1 \) and \( p_{k+1} \) in the resulting tree, \( T_k \)

(resp. \( \tilde{T}_k \)) and let \( w_k \) (resp. \( \tilde{w}_k \)) be the total weight of \( T_k \) (resp. \( \tilde{T}_k \)).

(ii) If \( k = r \), stop. Otherwise increase \( k \) by one and repeat step (i).

The tree that we get at the end of this process is an MST on the graph which

has \( \{v_1, p_1, \ldots, p_r\} \cup \{\omega_1 \cup \omega_2\} \) (resp. \( \{v_1, p_1, \ldots, p_r\} \cup \omega_1 \)) as its vertex set

and contains every possible edge between these vertices except the ones of

the form \( \{v_1, p\} \) with \( p \in \omega_1 \cup \omega_2 \) (resp. \( p \in \omega_1 \)) (Proposition 2 in [23]). It

is easy to see that the resulting tree is actually an MST for the complete

graph on \( \{v_1, p_1, \ldots, p_r\} \cup \{\omega_1 \cup \omega_2\} \) (resp. \( \{v_1, p_1, \ldots, p_r\} \cup \omega_1 \))

(respectively \( \{v_1, p_1, \ldots, p_r\} \cup \{\omega_1 \cup \omega_2\} \)) cannot contain an edge of the form \( \{v_1, p\} \) with

\( p \in \omega_1 \cup \omega_2 \) (resp. \( p \in \omega_1 \)) since \( \eta \) contains a wall around \( B_j \) in \( B(2sj, a_n) \).

Hence

\[
\delta_1 = (w_r - w_0) - (\tilde{w}_r - \tilde{w}_0) = \sum_{k=1}^{r} [(w_k - w_{k-1}) - (\tilde{w}_k - \tilde{w}_{k-1})].
\]

Now,

\[
w_k - w_{k-1} = \begin{cases} 
d(v_1, p_1), & \text{if } k = 1, \\
d(v_1, p_k) - \max(Y_{k-1}, d(v_1, p_k)), & \text{if } 2 \leq k \leq r.
\end{cases}
\]

A similar statement holds for \( \tilde{w}_k \) with \( \tilde{Y}_{k-1} \) replacing \( Y_{k-1} \). Proposition 2

in [23] shows that \( T_{k-1} \) (resp. \( \tilde{T}_{k-1} \)) is an MST on the graph with vertex set

\( V = (P \cap (B(n) - B_j)) \cup \{v_1\} \) (resp. \( \tilde{V} = (P \cap (\tilde{B}_j - B_j)) \cup \{v_1\} \)) and edge

set \( E_{k-1} = \cup_{i=1}^{k-1} \{v_1, p_i\} \cup \{\text{edges in the complete graph on } P \cap (B(n) - B_j)\} \)

(resp. \( \tilde{E}_{k-1} = \cup_{i=1}^{k-1} \{v_1, p_i\} \cup \{\text{edges in the complete graph on } P \cap (\tilde{B}_j - B_j)\} \))

for \( k \geq 2 \). Hence \( Y_{k-1} \) (resp. \( \tilde{Y}_{k-1} \)) is the maximum edge-weight in a

minimax path connecting \( v_1 \) and \( p_k \) in \( (V, E_{k-1}) \) (resp. \( (\tilde{V}, \tilde{E}_{k-1}) \)). This
gives $Y_{k-1} \leq \tilde{Y}_{k-1}$. From (6.37),

\begin{equation}
0 \leq (w_k - w_{k-1}) - (\tilde{w}_k - \tilde{w}_{k-1}) \leq \tilde{Y}_{k-1} - Y_{k-1}.
\end{equation}

Consider a random variable $U$ uniformly distributed on $(0, 2\sqrt{d}a_n)$ which is independent of $\mathcal{P}$. We have

\begin{equation}
\mathbb{E}_\eta|(w_k - w_{k-1}) - (\tilde{w}_k - \tilde{w}_{k-1})| \leq \mathbb{E}_\eta(\tilde{Y}_{k-1} - Y_{k-1})
= 2\sqrt{d}a_n \mathbb{P}_\eta(Y_{k-1} < U < \tilde{Y}_{k-1})
= \int_0^{2\sqrt{d}a_n} \mathbb{P}_\eta(Y_{k-1} < u < \tilde{Y}_{k-1}) \, du
\leq \int_0^{2\sqrt{d}a_n} \mathbb{P}(B(2s_j, a_n) \xrightarrow{u/2} \tilde{B}_j) \, du.
\end{equation}

The last inequality holds because of the following reason. Assume that $Y_{k-1} < u < \tilde{Y}_{k-1}$ and let $(v_1 = z_0, z_1, \ldots, z_\ell = p_k)$ be a minimax path connecting $v_1$ and $p_k$ in $(\tilde{V}, \tilde{E}_{k-1})$. Since $Y_{k-1} < \tilde{Y}_{k-1}$, $z_i \in \tilde{B}_j^c$ for some $i \leq \ell$. Let $k_1 + 1 := \min\{i \leq \ell : z_i \in \tilde{B}_j^c\}$ and $k_2 - 1 := \max\{i \leq \ell : z_i \in \tilde{B}_j^c\}$.

Then the $u/2$-clusters in $\tilde{B}_j - B_j$ containing $\{z_1, \ldots, z_{k_1}\}$ and $\{z_{k_2}, \ldots, z_{\ell}\}$ are disjoint, since otherwise we could find a path $(z_i = y_0, y_1, \ldots, y_\ell = z_i')$ for some $i \leq k_1$, $i' \geq k_2$ such that $y_p \in \tilde{V} - \{v_1\}$ and $d(y_p, y_{p+1}) \leq u$ for every $p \leq t - 1$. But this would mean that $(z_0, \ldots, z_i, y_1, \ldots, y_{\ell-1}, z_{i'}, \ldots, z_{\ell})$ is a path in $(\tilde{V}, \tilde{E}_{k-1})$ connecting $v_1$ and $p_k$ with maximum edge-weight strictly smaller than $\tilde{Y}_{k-1}$, a contradiction. Then the restrictions of the (disjoint) $u/2$-clusters in $\tilde{B}_j - B_j$ containing $\{z_1, \ldots, z_{k_1}\}$ and $\{z_{k_2}, \ldots, z_{\ell}\}$ to $\tilde{B}_j - B(2s_j, a_n)$ will contain two disjoint $u/2$-clusters which will satisfy the criteria for $B(2s_j, a_n) \xrightarrow{u/2} \tilde{B}_j$ to hold.

A calculation of similar nature can be done for $\delta_i$, for $2 \leq i \leq m$. From (6.34), (6.35), (6.36) and (6.39) we get,

\begin{equation}
\mathbb{E}\left[\mathbb{I}_{\mathcal{E}}|\Delta_j M(X) - \tilde{\Delta}_j M(X)|\right] \leq 4\sqrt{d}a_n \sup_{0 < u < 2\sqrt{d}a_n} \mathbb{P}(B(2s_j, a_n) \xrightarrow{u/2} \tilde{B}_j) \mathbb{E}\left(|\mathcal{P} \cap B_j| \cdot |\mathcal{P} \cap B(2s_j, a_n)|\right)
\leq ca_n^{d+1} \sup_{0 < u < 2\sqrt{d}a_n} \mathbb{P}(B(2s_j, a_n) \xrightarrow{u/2} \tilde{B}_j).
\end{equation}

Hence from Lemma 6.5

\begin{equation}
\mathbb{E}\left[\mathbb{I}_{\mathcal{E}}|\Delta_j M(X) - \tilde{\Delta}_j M(X)|\right] \leq c_8 a_n^{5d/2} \exp(c_9 a_n^{d-1}) \frac{\exp(c_9 a_n^{d-1})}{(\log n)^{d/2}}.
\end{equation}
In view of (6.33), we choose $a_n$ so that $c'a_n^d = \frac{d}{2} \log \log n$ to get
\begin{equation}
\mathbb{E} |\Delta_j M(X) - \tilde{\Delta}_j M(X)| \leq c (\log \log n)^{\frac{5}{2}} \exp(c''(\log \log n)^{\frac{d-1}{d}})
\end{equation}

Hence
\begin{equation}
\sum_{2} \text{Cov}(\Delta_j M(X)\Delta_j M(X^A), \Delta_{j'} M(X)\Delta_{j'} M(X^{A'})) \leq c'' n^{2d}(\log \log n)^{\frac{5}{2}p} \exp(c''/p \cdot (\log \log n)^{\frac{d-1}{d}}).
\end{equation}

The last inequality is a consequence of (6.32) and (6.41).

To complete our proof, we need a lower bound on $\sigma^2$. This is implicit in the work of Kesten and Lee in [23]. Let us write $L = \{j_1, \ldots, j_l\}$ and define the sigma-fields $F_k := \sigma\{X_{i} : i \leq k\}$ for $k = 1, \ldots, \ell$ and let $F_0$ be the trivial sigma-field. Then we can express $M(B(n)) - \mathbb{E} M(B(n))$ as a sum of martingale differences
\begin{equation}
M(B(n)) - \mathbb{E} M(B(n)) = \sum_{k=1}^{\ell} Z_k \text{ where }
Z_k := \mathbb{E} (M(B(n))|F_k) - \mathbb{E} (M(B(n))|F_{k-1}).
\end{equation}

From (4.27) in [23], it will follow that
\begin{equation}
\frac{1}{\ell} \sum_{k=1}^{\ell} Z_k^2 \overset{P}{\rightarrow} \zeta,
\end{equation}

for a positive constant $\zeta$. Using this together with fact that $\ell = \Theta(n^d)$, we get
\begin{equation}
\liminf_{n \to \infty} \frac{1}{n^d} \mathbb{E} (M(B(n)) - \mathbb{E} M(B(n))^2 > 0.
\end{equation}

From (6.24) and (6.43), it follows that
\begin{equation}
\frac{1}{2\text{Var}(M(X))} \sum_{j=1}^{\ell} \mathbb{E} |\Delta_j M(X)|^3 \leq \frac{c}{n^{d/2}}.
\end{equation}

Combining (6.21), (6.25), (6.27), (6.42), (6.43) and (6.44), we see that there are positive constants $c$ and $c'$ depending on $p$ and $a$ such that
\begin{equation}
\mathcal{W}(\mu_n, G) \leq c (\log \log n)^{\frac{5}{2}p} \exp(c'(\log \log n)^{\frac{d-1}{d}}).
\end{equation}
Since (6.45) is true for every \( p > 1 \), we get the bound stated in (2.2) by an application of (6.22) for some constant \( c_2 \), again depending only on \( p \) and the dimension.

Let us now turn to the case \( d = 2 \). Define

\[
\mathcal{E}^{(\alpha)} := \{(j, j', A, A') : A, A' \subseteq \mathcal{L}; j \in \mathcal{L} - A, j' \in \mathcal{L} - A' \text{ and } \|j - j'\|_{\infty} \leq n^\alpha \}
\]

and

\[
\tilde{\mathcal{E}}^{(\alpha)} := \{(j, j', A, A') : A, A' \subseteq \mathcal{L}; j \in \mathcal{L} - A, j' \in \mathcal{L} - A' \} - \mathcal{E}^{(\alpha)}
\]

for \( 0 < \alpha < 1 \). We break up the sum on the right side of (6.25) into two parts, \( \sum_1 \) and \( \sum_2 \), where \( \sum_1 \) denotes sum over all \( (j, j', A, A') \in \mathcal{E}^{(\alpha)} \) and \( \sum_2 \) denotes sum over the remaining terms. A calculation similar to (6.27) will show that

\[
(6.46) \quad \sum_1 \text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})) \leq cn^{2+2\alpha}.
\]

We now define \( \tilde{\Delta}_j M(X) \) as in (6.28) for every \( j \in \mathcal{L} \). Then for every \( (j, j', A, A') \in \tilde{\mathcal{E}}^{(\alpha)} \), (6.29), (6.30) and (6.32) hold. Denoting \( B(2s_j, n^\alpha) \) by \( \tilde{B}_j \) for every \( j \in \mathcal{L} \), define the event

\[
E := \{X \text{ contains a } B(n) - \text{wall around } \tilde{B}_j \text{ in } B(2s_j, a_n)\}
\]

where \( a_n \) is a sequence increasing to infinity in a way so that \( a_n \leq \log n \). Then we can proceed exactly as before to arrive at

\[
\mathbb{E} \left( \mathbb{I}_{E^c} |(\Delta_j M(X) - \tilde{\Delta}_j M(X)) | \right) \leq c\sqrt{a_n} \exp(-c' a^2_n),
\]

and

\[
\mathbb{E} \left[ \mathbb{I}_E |\Delta_j M(X) - \tilde{\Delta}_j M(X)| \right] \leq ca^3_n \sup_{0 < u < 2\sqrt{2a_n}} \mathbb{P} \left( B(2s_j, a_n) \xrightarrow{2/u} \tilde{B}_j \text{ in } B(n) \right).
\]

An application of Lemma 6.5 and arguments similar to the ones used previously for \( d \geq 3 \) will yield

\[
(6.47) \quad \sum_2 \text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})) \leq cn^{4/n^{\alpha/2}}.
\]

Combining (6.46) and (6.47), we get

\[
(6.48) \quad \text{Var}(\mathbb{E}(T|W)) \leq cn^{4/n^{\alpha/2}}.
\]

here we have chosen \( \alpha = 2p/(\beta_1 + 2p) \) in order to minimize the bound.

From (6.48), (6.43), (6.44) (with \( d = 2 \)), (6.21) and (6.22), we get the bound in (2.1).

This completes the proof of Theorem 2.1. \( \square \)
6.3. **Proof of Lemma 5.2.** The techniques used in this proof are similar to the ones used in the proof of Lemma 5.1. We will outline them briefly.

Fix \( p \in [p_1, p_2] \). Let \( u_1, \ldots, u_m \) be the edges of \( \mathbb{Z}^d \) both of whose endpoints lie in \( B(n) \) and let \( X_1, \ldots, X_m \) be i.i.d. Bernoulli(\( p \)) random variables (i.e. \( \mathbb{P}(X_1 = 1) = p = 1 - \mathbb{P}(X_1 = 0) \)) associated to them. Let \( X := (X_1, \ldots, X_m) \) and let \( X' := (X'_1, \ldots, X'_m) \) be an independent copy of \( X \). As earlier we define the event

\[
E := \{ \text{there is exactly one } p\text{-cluster in } B(n) \text{ which intersects both } B(a_n) \text{ and } \partial B(n) \}
\]

for some \( a_n \to \infty \) in a way so that \( a_n = o(n) \). Define the function \( f \) by

\[
f(X) := \mathbb{1}_E(X).
\]

Then an application of Lemma 6.1 yields

\[
\text{Var}(f(X)) \geq \sum_{i \in I} \text{Var}[E(f(X)|X_i)]
\]

where \( I := \{ i \leq m : \text{both endpoints of } u_i \text{ lie in } B(a_n/3) \} \). Fix \( i \in I \), denote the endpoints of \( u_i \) by \( v_1 \) and \( v_2 \). With the notations used in the proof of Lemma 5.1, we have

\[
\text{Var}[E(f(X)|X_i)] \geq \frac{1}{2} \mathbb{E} \left[ \left( \mathbb{E}(f(X)|X_i) - \mathbb{E}(f(X')|X'_i) \right)^2 \right] \\
\geq \frac{1}{2} \mathbb{E} \left[ \mathbb{P}(A_i)^2 \mathbb{I}\{X_i = 1, X'_i = 0\} \right]
\]

where

\[
A_i = \{ u_i \xleftarrow{p} B(n) - u_i, \text{ any } p\text{-cluster in } B(n) - u_i \text{ which intersects} \}
\]

\[
\text{both } \partial B(n) \text{ and } B(a_n) \text{ contains either } v_1 \text{ or } v_2 \\}
\]

Now,

\[
\mathbb{P}(A_i) \geq \mathbb{P}(u_i \xleftarrow{p} B(v_1, 2n) - u_i, \text{ if } C \text{ is a } p\text{-cluster in } B(v_1, 2n) - u_i
\]

\[
\text{then every connected component of } C \cap B(v_1, n/2) \text{ which intersects } \partial B(v_1, n/2) \text{ and } B(v_1, 2a_n)
\]

\[
\text{contains either } v_1 \text{ or } v_2 \}
\]

\[
= \mathbb{P}(F)/(1 - p),
\]

where

\[
F := \{ \{0, e_1\} \xleftarrow{p} B(2n), \text{ if } C \text{ is a } p\text{-cluster in } B(2n) - \{0, e_1\} \}
\]

\[
\text{then every connected component of } C \cap B(n/2) \text{ which intersects } \partial B(n/2) \text{ and } B(2a_n)
\]

\[
\text{contains either 0 or } e_1 \\}
\]
From (6.49) and (6.50), we conclude that $\mathbb{P}(F) \leq c/a_n^{d/2}$. As before, we have
\begin{equation}
\mathbb{P}(\{0, e_1\} \xrightarrow{2/p} B(2n)) \leq \mathbb{P}(F) + \mathbb{P}(B(2a_n) \xrightarrow{3/p} B(n/2)).
\end{equation}

We now define a cube $Q \subset B(n/2)$ to be a trifurcation box in $B(n/2)$ at level $p$, if

(i) there is a $p$-cluster $C$ in $B(n/2)$ with $C \cap Q \neq \emptyset$ and

(ii) the vertices of $C$ contained in $B(n/2) - Q$ contain at least three disjoint $p$-clusters in $B(n/2) - Q$ each of which intersects $\partial B(n/2)$.

We can then apply the arguments in the proof of Lemma 6.2 to show that
\begin{equation}
\mathbb{P}(B(2a_n)) \text{ is a trifurcation box in } B(n/2) \text{ at level } p \leq \frac{ca_n^d}{n},
\end{equation}
from which it will follow that
\begin{equation}
\mathbb{P}(B(2a_n) \xrightarrow{3/p} B(n/2)) \leq c \exp(c'a_n^d/a_n^d).
\end{equation}

Combining (6.51) and (6.52), we choose $c'a_n = \log n/2$ to get the desired bound.

6.4. Proof of Theorem 2.4. We start with an analogue of Lemma 6.5.

Lemma 6.14. For $p_1 \in (0, p_c(\mathbb{Z}^d))$ and $p_2 \in (p_c(\mathbb{Z}^d), 1)$, we have the following estimates.

(i) When $d = 2$,
\begin{equation}
\mathbb{P}(\{0, e_1\} \xrightarrow{2/p} B(n)) \leq \begin{cases} 
c_{22} \exp(-c_{23} n), & \text{if } p \leq p_1, 
c_{24}/n^{\beta_2}, & \text{if } p > p_1.
\end{cases}
\end{equation}
The same bound holds for $\mathbb{P}(B(1) \xrightarrow{2/p} B(n))$ and $\mathbb{P}(B(1) \xrightarrow{2/p} B(n) \text{ in } Q)$ whenever $Q$ is a cube containing the origin and $Q \cap \partial B(n) \neq \emptyset$. Here $\beta_2$ is a universal positive constant and the constants $c_{22}, c_{23}$ and $c_{24}$ depend only on $p_1$ and $p_2$.

(ii) When $d \geq 3$,
\begin{equation}
\mathbb{P}(\{0, e_1\} \xrightarrow{2/p} B(n)) \leq \begin{cases} 
c_{25} \exp(-c_{26} n) & \text{if } p \leq p_1, 
c_{27}/(\log n)^{d/2} & \text{if } p \in [p_1, p_2], 
c_{28} \exp(-c_{29} n) & \text{if } p \geq p_2.
\end{cases}
\end{equation}
The constants appearing here depend only on $p_1$, $p_2$ and $d$. When $p \leq p_1$ or $p \geq p_2$, the same bounds hold for $\mathbb{P}(B(1) \xrightarrow{2/p} B(n))$ and $\mathbb{P}(B(1) \xrightarrow{2/p} B(n) \text{ in } Q)$ for any cube $Q$ containing the origin for which $Q \cap \partial B(n) \neq \emptyset$.

Proof. The bounds in the subcritical regime follow from Menshikov’s Theorem (see e.g. [21]). When $d \geq 3$ and $p \geq p_2$, exponential decay will follow from the proof of Lemma 7.89 in [21]. The bound for $p > p_1$ in dimension two is standard and follows along the lines of the proof of Lemma 6.5. The
bound for \( p \in [p_1, p_2] \) in dimension bigger than two is just the content of Lemma 5.2.

Let \( u_1, \ldots, u_\ell \) be the edges of \( \mathbb{Z}^d \) having both endpoints in \( B(n) \) and let \( X_1, \ldots, X_\ell \) be the weights associated with them. Define \( X = (X_1, \ldots, X_\ell) \) and let \( X' = (X'_1, \ldots, X'_\ell) \) be an independent copy of \( X \). We will apply (6.21) with \( f(X) = M(X) \).

**Proof of (2.4), (2.5), (2.7) and (2.8).** Mimicking the proof of Theorem 2.1 we define the sets

\[ C^{(\alpha)} = \{(j, j', A, A'): j, j' \leq \ell, A, A' \subseteq \{1, \ldots, \ell\}; j \notin A, j' \notin A' \text{ and either some endpoint of } u_j \text{ or } u_{j'} \text{ does not lie in } B(n - n^\alpha) \text{ or } \|x_j - x_{j'}\|_\infty \leq 2n^\alpha \text{ for some endpoint } x_j \text{ (resp. } x_{j'}) \text{ of } u_j \text{ (resp. } u_{j'}) \} \]

\[ S^{(\alpha)} = \{(j, j', A, A'): j, j' \leq \ell, A, A' \subseteq \{1, \ldots, \ell\}; j \notin A, j' \notin A' \} - C^{(\alpha)} \]

for \( 0 < \alpha < 1 \). Note that under the assumption of finite \((4 + \delta)\)-th moment (or under the assumption of compact support) on \( \mu \),

\[
E|\Delta_j M(X)|^{(4+\delta)} \leq C, \text{ for every } j \leq \ell.
\]

Both (6.25) and (6.26) remain the same in our present setup with \( \mathcal{L} = \{1, \ldots, \ell\} \). If we split the sum appearing in (6.25) into two parts \( \Sigma_1 \) (the sum over \( C^{(\alpha)} \)) and \( \Sigma_2 \) (the sum over the rest of the terms), then (6.27) remains the same.

Letting \( C^{(\alpha)}_{1,2} := \{(j, j'): j, j' \leq \ell, (j, j', 0, 0) \in C^{(\alpha)} \}, \) \( B(x_j, n^\alpha) \) and \( B(x_{j'}, n^\alpha) \) are disjoint whenever \( (j, j') \notin C^{(\alpha)}_{1,2} \) and \( x_j \) (resp. \( x_{j'} \)) is an endpoint of \( u_j \) (resp. \( u_{j'} \)). For each \( u_j \) having both endpoints in \( B(n - n^\alpha) \), fix an endpoint \( x_j \) of \( u_j \) and let \( B_j \) (resp. \( \tilde{B}_j \)) denote the weighted graph having \( B(x_j, n^\alpha) \) as the underlying graph whose edge weights come from the appropriate subvector of \( X \) (resp. \( X^j \)).

Let \( \tilde{\Delta}_j M(X) := M(B_j) - M(\tilde{B}_j) \). Then (6.29) and (6.30) apply whenever \( (j, j', A, A') \in S^{(\alpha)} \). If \( \mu \) has bounded support, then

\[
T_1 + T_2 \leq cE|\Delta_j M(X) - \tilde{\Delta}_j M(X)|,
\]

with \( T_1 \) and \( T_2 \) as in (6.30). If \( \mu \) has unbounded support and finite \((4 + \delta)\)-th moment, then

\[
T_1 \leq \left( E|\Delta_j M(X) - \tilde{\Delta}_j M(X)| \right)^{1/q'} \cdot \left[ E \left( |\Delta_j M(X) - \tilde{\Delta}_j M(X)| \left( |\Delta_j M(X^A)\Delta_j' M(X)\Delta_j' M(X^A')| \right)^q \right] \right)^{1/q}
\]

\[
= C' \left( E|\Delta_j M(X) - \tilde{\Delta}_j M(X)| \right)^{1/q'},
\]
where \( q = 1 + \delta/3 \) and \( q' = 1 + 3/\delta \). That \( C' \) is finite is ensured by \eqref{6.54}. An application of Hölder’s inequality will give a similar bound for \( T_2 \). So the task now is to get an upper bound for \( \mathbb{E}|\Delta_j M(X) - \tilde{\Delta}_j M(X)| \) for each \( u_j \) having both endpoints in \( B(n - n^\alpha) \).

To this end, define \( Y_j \) (resp. \( \tilde{Y}_j \)) to be the maximum edge-weight in the path connecting the two endpoints of \( u_j \) in an MST of \( B(n) - u_j \) (resp. \( B(x_j, n^\alpha) - u_j \)) when the edge-weights are given by the appropriate sub-vector of \( X \). Then

\[
|\Delta_j M(X) - \tilde{\Delta}_j M(X)| = |\max(X_j, Y_j) - \max(X_j, \tilde{Y}_j) - \max(X'_j, Y_j) + \max(X'_j, \tilde{Y}_j)| \\
\leq 2|Y_j - \tilde{Y}_j|.
\]

The first equality holds since \( M(X) = M(B(n) - u_j) + X_j - \max(X_j, Y_j) \) and similar assertions are true when \( X \) is replaced by \( X' \) and \( B(n) \) is replaced by \( B(x_j, n^\alpha) \).

Let \( Z_j \) be the maximum of the weights associated with the edges of \( B(x_j, 1) - u_j \). Clearly \( Y_j \leq \tilde{Y}_j \leq Z_j \). Then

\[
\mathbb{E}|Y_j - \tilde{Y}_j| = \mathbb{E}\left(Z_j \mathbb{E}_1 \frac{|Y_j - \tilde{Y}_j|}{Z_j}\right),
\]

where \( \mathbb{E}_1 \) denotes expectation conditional on the weights associated with the edges of \( B(x_j, 1) - u_j \).

For a random variable \( U \) following Uniform\([0, 1]\) distribution that is independent of \( X, X' \),

\[
\mathbb{E}_1 \frac{|Y_j - \tilde{Y}_j|}{Z_j} = \mathbb{P}_1\left(Y_j < U Z_j < \tilde{Y}_j\right) \\
\leq \mathbb{P}_1\left(B(x_j, 1) \xleftarrow{2} F_{\mu(U Z_j)} B(x_j, n^\alpha)\right),
\]

where \( F_{\mu} \) is the distribution function associated with \( \mu \). That the last inequality holds can be seen by following the argument given right after \eqref{6.39}.

From Lemma 6.14

\[
\sup_{0 < u < 1} \mathbb{P}_1\left(B(x_j, 1) \xleftarrow{2} F_{\mu(u Z_j)} B(x_j, n^\alpha)\right) \leq c/(\log n)^d
\]

when \( d \geq 3 \). From the estimates in \eqref{6.55}, \eqref{6.56}, \eqref{6.57}, \eqref{6.58}, \eqref{6.59} and \eqref{6.60}, we get

\[
\sum_{2} \text{Cov}(\Delta_j M(X) \Delta_j M(X'), \Delta_j M(X) \Delta_j M(X')) \left(\frac{\ell}{|A|} \left|\ell - |A|\right| \left(\frac{\ell}{|A'|} \left|\ell - |A'|\right|\right)\right) \leq \frac{c \cdot n^{2d}}{\left(\log n\right)^{d/2q}},
\]

where \( q = 1 \) if \( \mu \) has compact support and \( q = 1 + 3/\delta \) if \( \mu \) has unbounded support and finite \((4 + \delta)\)-th moment.
Once we prove that $\text{Var}(M(X)) = \Theta(n^d)$, it will follow that \eqref{eq:6.44} remains true in the present setting. Hence \eqref{eq:6.21}, \eqref{eq:6.25}, \eqref{eq:6.27}, \eqref{eq:6.61}, \eqref{eq:6.44} and \eqref{eq:6.22} will yield \eqref{eq:2.4} and \eqref{eq:2.7}. The proof of $\text{Var}(M(X)) = \Theta(n^d)$ will be given in the more general setup of Theorem 2.5.

If $d \geq 3$ and there exists a unique $x \in \mathbb{R}$ such that $\mu[0, x] = p_c(\mathbb{Z}^d)$ then there are two possibilities. The first is that, $F'_\mu$ is continuous at $x$. In this case choose a small enough positive $\epsilon_0$ so that $x - \epsilon_0 > 0$. Set $p_1(\eta) = F'_\mu(x - \eta)$ and $p_2(\eta) = F'_\mu(x + \eta)$.

For $\epsilon \in (0, \epsilon_0)$, we have from \eqref{eq:6.59}

$$
\mathbb{E}_1 |Y_j - \tilde{Y}_j| \leq \int_0^1 \mathbb{P}_1 \left( B(x_j, 1) \xrightarrow{\mathcal{F}_n} B(x_j, n^\alpha) \right) du.
$$

Breaking up the integral into

$$
\int_0^{\min((x-\epsilon)/Z_j, 1)}, \int_0^{\min((x+\epsilon)/Z_j, 1)} \text{ and } \int_0^1,
$$

we get

$$
\mathbb{E}_1 |Y_j - \tilde{Y}_j| \leq c_{30} \exp(-c_{31} n) + \frac{2\epsilon}{Z_j} \cdot c_{27} (\log n)^{-\frac{d}{2}}
$$

by an application of Lemma 6.14. The constants $c_{30}$ and $c_{31}$ depend on $c_{25}$, $c_{26}$, $c_{28}$ and $c_{29}$ as in Lemma 6.14 corresponding to the choices $p_i = p_i(\epsilon)$, $i = 1, 2$ and the constant $c_{27}$ is the one from Lemma 6.14 corresponding to the choices $p_i = p_i(\epsilon_0)$, $i = 1, 2$.

From \eqref{eq:6.57}, \eqref{eq:6.58} and \eqref{eq:6.62}, we get

$$
\mathbb{E} |\Delta_j M(X) - \tilde{\Delta}_j M(X)| \leq 2c_{30} \mathbb{E}(Z_j) \exp(-c_{31} n) + \frac{4c_{27}\epsilon}{(\log n)^{\frac{d}{2}}}
$$

for $(j, j') \notin \xi^{(\alpha)}_{1, 2}$. The second possibility, namely $F'_\mu(x-) < F'_\mu(x)$ can be handled similarly. Proceeding exactly as before, we arrive at

$$
\sum_2 \frac{\text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'}))}{\left( \frac{\ell}{|A|} \right)^{\ell - |A|} \left( \frac{\ell}{|A'|} \right)^{\ell - |A'|}} \leq c \cdot n^{2d} \left( 2c_{30} \mathbb{E}(Z_j) \exp(-c_{31} n) + \frac{4c_{27}\epsilon}{(\log n)^{\frac{d}{2}}} \right)^{\frac{1}{\mathcal{G}}},
$$

where $\mathcal{G}$ has the same meaning as in \eqref{eq:6.61}. The last inequality together with \eqref{eq:6.21}, \eqref{eq:6.25}, \eqref{eq:6.27} and \eqref{eq:6.44} yields

$$
\mathcal{W}(\nu_n, G) \leq c' \left[ \frac{1}{n^{\frac{1+\alpha}{2}}} + \left( 2c_{30} \mathbb{E}(Z_j) \exp(-c_{31} n) + \frac{4c_{27}\epsilon}{(\log n)^{\frac{d}{2}}} \right)^{\frac{1}{\mathcal{G}}} + \frac{1}{n^{\frac{d}{2}}} \right],
$$

the constant $c'$ is free of $\epsilon$. 

Hence
\[ \limsup_{n} (\log n)^{\frac{d}{d+1}} \mathcal{W}(\nu_n, G) \leq c'(4c_2\epsilon)^{\frac{1}{d+1}}. \]

Since \( \epsilon \) was arbitrary, we get (2.5) and (2.8) by an application of (6.22). The argument is similar if \( \mu[0, x] = p_c(Z^d) \) for some unique \( x \in \mathbb{R} \).

**Proof of (2.3), (2.6) and (2.9).** We introduce
\[ \tilde{\mathcal{E}}^{(\alpha)} := \{(j, j', A, A') : j, j' \leq \ell, A, A' \subseteq \{1, \ldots, \ell\}; j \notin A, j' \notin A' \}
\]
and \( \tilde{\mathcal{S}}^{(\alpha)} := \{(j, j', A, A') : j, j' \leq \ell, A, A' \subseteq \{1, \ldots, \ell\}; j \notin A, j' \notin A' \} - \tilde{\mathcal{E}}^{(\alpha)} \)
for \( 0 < \alpha < 1 \). We split the sum appearing in (6.25) into \( \Sigma_1 \), the sum over \((j, j', A, A') \in \tilde{\mathcal{E}}^{(\alpha)} \) and \( \Sigma_2 \), the sum over \((j, j', A, A') \in \tilde{\mathcal{S}}^{(\alpha)} \).

For each \( j \leq \ell \), fix an endpoint \( u_j \) of \( u_j \) and define \( B_j \) (resp. \( \tilde{B}_j \)) to be the weighted graph having \( B(n) \cap B(x_j, n^\alpha) \) as the underlying graph whose edge weights are elements of the appropriate subvector of \( X \) (resp. \( X^2 \)). Define \( \Delta_1 M(X) = M(B_j) - M(\tilde{B}_j) \) for each \( j \leq \ell \).

For every \( j \leq \ell \), define \( Y_j \) (resp. \( \tilde{Y}_j \)) to be the maximum edge weight in the path connecting the endpoints of \( u_j \) in an MST of \( B(n) - u_j \) (resp. \( B_j - u_j \)) and let \( Z_j \) be the maximum of the weights associated with the edges of \( B(n) \cap B(x_j, 1) - u_j \). We emphasize that these weights are elements of the random vector \( X \).

Then (6.29), (6.30), (6.55) (if \( \mu \) has bounded support) or (6.56) (if \( \mu \) has unbounded support but finite \((4 + \delta)\)-th moment), (6.57), (6.58) hold whenever \((j, j', A, A') \in \tilde{\mathcal{S}}^{(\alpha)} \). Similar to our previous calculations, we get the estimate
\[ (6.63) \quad \mathbb{E}_1 \left| \frac{Y_j - \tilde{Y}_j}{Z_j} \right| \leq \sup_{u \in [0, 1]} \mathbb{P}_1 \left( B(x_j, 1) \xrightarrow{2} F_{\mu(u Z_j)} B(x_j, n^\alpha) \text{ in } B(n) \right). \]

When \( d = 2 \), it follows from Lemma 6.14 that the bound in (6.47) holds with \( \beta_1 \) replaced by \( \beta_2 \) and \( p \) replaced by \( \bar{q} / (\bar{q} \text{ having the same meaning as in } (6.61)) \). Also (6.46) remains true if \( d = 2 \). Since (6.44) holds (with \( d = 2 \)), we get the bounds in (2.3) and (2.6) upon choosing \( \alpha = 2\bar{q}/(\beta_2 + 2\bar{q}) \).

When \( d \geq 3 \) and \( \mu[0, x] < p_c(Z^d) < \mu[0, x] \) for some \( x \in \mathbb{R} \), we have
\[ (6.64) \quad \sum_1^n \text{Cov}(\Delta_1 M(X) \Delta_1 M(X^A), \Delta_1 M(X) \Delta_1 M(X^A)) \leq cn^{d+\alpha d}. \]
Since $\text{Range}(F_\mu) \subset (p_c - \epsilon, p_c + \epsilon)^c$ for some $\epsilon > 0$, it follows from (6.63) and Lemma 6.14 that

$$\mathbb{E}_1 \frac{|Y_j - \tilde{Y}_j|}{Z_j} \leq \sup_{p \in (p_c - \epsilon, p_c + \epsilon)} \mathbb{P} \left( B(x_j, 1) \overset{\rightarrow}{\rightarrow} B(x_j, n^\alpha) \text{ in } B(n) \right)$$

$$\leq c \exp(-c'n^\alpha).$$

Hence

(6.65)

$$\sum_{2} \text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'})) \leq cn^{2d} \exp(-c'n^\alpha / \eta).$$

From (6.64), (6.65), (6.44) and the fact $\text{Var}(M(X)) = \Theta(n^d)$, we conclude that

$$\mathcal{D}(\nu_n, G) \leq \frac{c}{n^{(1 - \alpha)/4}},$$

by an application of (6.22). We get the bound in (2.9) once we replace $d(1 - \alpha)/4$ by $\gamma$. □

This concludes the proof of Theorem 2.4.

6.5. Proof of Theorem 2.5. To fix ideas, we first assume that $G$ is symmetric.

If $G$ is symmetric and deletion of an edge of $G$ creates two components, then $G$ is a regular tree. Hence all our claims follow trivially. So we can assume that this is not the case.

Let $E_n = \{u_1, \ldots, u_{e_n}\}$ and let $X_1, \ldots, X_{e_n}$ be the associated edge weights. Note that $M_n = M(X_1, \ldots, X_{e_n})$. Let $X = (X_1, \ldots, X_{e_n})$ and let $X' = (X'_1, \ldots, X'_{e_n})$ be an independent copy of $X$.

Define $\mathcal{I}_{r}^n := \{i \leq e_n : S(v, r) \subset G_n \text{ for each endpoint } v \text{ of } u_i\}$. For large $r$ and $i \in \mathcal{I}_{r}^n$, fix an endpoint $v_i$ of $u_i$ and let $Y^n_i$ (resp. $Y^n_i(r)$) be the maximum edge weight in a path connecting the endpoints of $u_i$ in an MST of $G_n - u_i$ (resp. $S(v_i, r) - u_i$) with the edge weights being the appropriate subvector of $X$. We will suppress the dependence on $n$ and simply write $\mathcal{I}_r$, $Y_i$ and $Y_i(r)$.

An application of Lemma 6.1 yields, with our usual notations,

(6.66) \hspace{1cm} \text{Var}(M_n) \geq \sum_{i=1}^{\ell_n} \text{Var}(E(M_n|X_i))$

$$= \frac{1}{2} \sum_{i=1}^{\ell_n} \mathbb{E} \left[ \text{Var}(M(X)|X_i) - \text{Var}(M(X^i)|X'_i) \right]^2$$

$$\geq \frac{1}{2} \sum_{i \in \mathcal{I}_r} \mathbb{E} \left[ \text{Var}(M(X)|X_i) - \text{Var}(M(X^i)|X'_i) \right]^2$$
For $i \in I_r$, $M(X) = M(G_n - u_i) + X_i - \max(X_i, Y_i)$ and hence

$$M(X) - M(X_i) = \min(X_i, Y_i) - \min(X_i', Y_i').$$

Since $\mu$ is non-degenerate we can find real numbers $b > a$ such that $\mu[0, a] > 0$ and $\mu[b, \infty] > 0$. Going back to (6.66)

$$\text{Var}(M_n) \geq \frac{1}{2} \sum_{i \in I_r} \mathbb{E} \left[ \mathbb{E}(\min(X_i, Y_i) | X_i) - \mathbb{E}(\min(X_i', Y_i') | X_i') \right]^2$$

(6.67)  

$$\geq \frac{1}{2} \sum_{i \in I_r} \mathbb{E} \left[ \{ (b - a) \mathbb{P}(Y_i \geq b) \}^2 \mathbb{I}\{X_i \leq a, X_i' \geq b\} \right]$$

$$\geq \frac{1}{2} |I_r| (b - a)^2 \cdot p^2 \mu[0, a] \mu[b, \infty]$$

where

$$p := \mathbb{P}(\text{The weight associated with each edge sharing one vertex with } u_i \text{ is at least } b).$$

(Since $G$ is symmetric, $p$ does not depend on the edge $u_i$.) By assumption (III)

$$|I_r| = \Theta(|V_n|).$$

From (6.67) and (6.68), it follows that

$$\text{Var}(M_n) \geq c|V_n|.$$  

The upper bound is a simple consequence of the Efron-Stein inequality

$$\text{Var}(M_n) \leq \frac{1}{2} \sum_{j=1}^{\ell_n} \mathbb{E}(\Delta_j M_n)^2.$$ 

Thus we have proven that $\text{Var}(M_n) = \Theta(|V_n|)$.

Turning toward the proof of the central limit theorem, define, for large $r$

$$\mathcal{E}_n(r) := \{(j, j', A, A') : j, j' \leq \ell_n, A, A' \subseteq \{1, \ldots, \ell_n\}, j \notin A, j' \notin A'$$

and either $d_G(x_j, x_{j'}) \leq 2r$ or $S(x_j, r) \not\subseteq G_n$ or $S(x_{j'}, r) \not\subseteq G_n$

for some endpoints $x_j, x_{j'}$ of $u_j$ and $u_{j'}$ respectively} and

$$\mathcal{F}_n(r) = \{(j, j', A, A') : j, j' \leq \ell_n, A, A' \subseteq \{1, \ldots, \ell_n\},$$

$$j \notin A, j' \notin A' \} - \mathcal{E}_n(r).$$

We will apply the approximation Theorem from [18] with $f(X) = M(G_n)$.

Proceeding as before, we split the sum in (6.25) into $\Sigma_1$, the sum over all $(j, j', A, A') \in \mathcal{E}_n(r)$ and $\Sigma_2$, the sum over the rest of the terms. Since $|\Delta_j M(X)| \leq |X_j - X_j'|$ and $\mathbb{E}(X_j^4) < \infty$, a computation similar to (6.27),
will yield

\[
\sum_1 \frac{\text{Cov}(\Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'}))}{(|A|)(\ell_n - |A|)(|A'|)(\ell_n - |A'|)} \leq c |V_n| (a_r + |\{v \in V_n : S(v, r) \not\subset G_n\}|),
\]

where \(a_r\) is the number of vertices in a sphere of radius 2\(r\) around a vertex in \(G\), i.e. \(a_r = |\{v' \in V : d_G(v, v') \leq 2r\}|\) for some (and hence all, by symmetry) \(v \in V\).

For \(j \in I_r\), define

\[
\tilde{\Delta}_j M(X) = M(B_j) - M(\tilde{B}_j)
\]

where the underlying graph for both \(B_j\) and \(\tilde{B}_j\) is \(S(v_j, r)\) and the associated edge weights are appropriate subvectors of \(X\) and \(X^j\) respectively. With this definition of \(\tilde{\Delta}_j M(X)\), (6.29) and (6.30) hold for \((j, j', A, A') \in \mathcal{F}_n(r)\). As in (6.56), we get

\[
T_1 \leq c \left( \mathbb{E}|\Delta_j M(X) - \tilde{\Delta}_j M(X)| \right)^{1/3/\delta}
\]

and a similar bound holds for \(T_2\). A calculation similar to (6.57) yields

\[
|\Delta_j M(X) - \tilde{\Delta}_j M(X)| \leq 2(Y_j(r) - Y_j)
\]

for \(j \in I_r\).

Fix a vertex \(v\) of \(G\) and let \(e\) be an edge incident to \(v\). Let \(Y(v, e, r)\) be the maximum edge weight in the path connecting the endpoints of \(e\) in an MST of \(S(v, r) - e\), clearly \(Y(v, e, r)\) is decreasing in \(r\). Define

\[
Y(v, e) := \lim_{r \to \infty} Y(v, e, r).
\]

The above convergence also holds in \(L^1\) as a consequence of dominated convergence theorem. Since \(G\) is symmetric, \(Y^n_i(r)\) has the same distribution as \(Y(v, e, r)\) and \(Y^n_i\) dominates \(Y(v, e)\) stochastically for every \(i \in I^n_r\). Hence

\[
\lim_{r \to \infty} \limsup_{n \to \infty} \max_{i \in I^n_r} \mathbb{E}(Y^n_i(r) - Y^n_i) \leq \lim_{r \to \infty} \mathbb{E}(Y(v, e, r) - Y(v, e)) = 0.
\]

Thus we have

\[
\lim_{r \to \infty} \limsup_{n \to \infty} \max_{(j, j', A, A') \in \mathcal{F}_n(r)} \text{Cov} \left( \Delta_j M(X) \Delta_j M(X^A), \Delta_{j'} M(X) \Delta_{j'} M(X^{A'}) \right) = 0,
\]

which gives us control over \(\sum_2\). Further

\[
\frac{1}{2 \text{Var}(M_n)^{3/2}} \sum_{j=1}^{\ell_n} |\mathbb{E}\Delta_j M(X)|^3 \leq \frac{c}{|V_n|^{1/2}}.
\]
The last inequality together with \( 6.69 \), \( 6.71 \), \( 6.21 \) and the fact that \( \text{Var}(M_n) = \Theta(|V_n|) \) yields
\[
\limsup_n W(\mu_n, G) = 0,
\]
\( \mu_n \) being the law of \( (M_n - \mathbb{E}M_n)/\sqrt{\text{Var}(M_n)} \).

Assume now that \( G \) is vertex-transitive so that there are two kinds of edges. Call an edge \( e \in E \) of type A if deletion of \( e \) results in the creation of two disjoint components. We say \( e \) is of type B if it is not of type A. Define \( \tilde{I}_r^n := \{ i \in I_r : i \text{ is of type B} \} \). Define \( Y_i^n \) and \( X_i^n(\mathbb{Y}) \) by the previous definitions for each \( i \in \tilde{I}_r^n \). Then
\[
\text{Var}(M_n) \geq \frac{1}{2} \sum_{i \in \tilde{I}_r^n} \mathbb{E} \left[ \mathbb{E}(M(X)|X_i) - \mathbb{E}(M(X)|X_i') \right]^2
\]
and \( |\tilde{I}_r| = \Theta(|V_n|) \) if \( G \) is not a tree. So we can argue as before to conclude that \( \text{Var}(M_n) = \Theta(|V_n|) \).

Next, note that if \( j \in I_r^n \) and \( u_j \) is of type A, then \( \Delta_j M(X) - \Delta_j M(X) = 0 \). Further, our previous arguments show that
\[
\lim_{r \to \infty} \limsup_{n \to \infty} \max_{i \in \tilde{I}_r^n} \mathbb{E}(Y_i^n(r) - Y_i^n) \leq \lim_{r \to \infty} \sum_{v \in \mathbb{Y}} \mathbb{E}(Y(v,e,r) - Y(v,e)) = 0,
\]
\( \sum_{v \in \mathbb{Y}} \) being the sum over all type B edges \( e \) incident to \( v \). The rest of the arguments remain the same.

This finishes the proof of the central limit theorem.

**Appendix A.**

In this section we fill in the gaps in the proof of Lemma 6.5.

**Proposition.** Assume the setup of Lemma 6.5. Then
(i) When \( d = 2 \), \( a \in (1/2, \log n) \) and \( r_c \leq r \leq (\log n)^2 \),
\[
\mathbb{P}(B(a) \xrightarrow{2 \to r} B(n)) \leq c_{14}/n^{\beta_1}.
\]
The same bound holds for \( \mathbb{P}(B(a) \xrightarrow{2 \to r} B(n) \text{ in } K) \) for a square \( K \) which contains the origin and \( K \cap \partial B(n) \neq \emptyset \).
(ii) When \( d \geq 3 \), \( a \in (1/2, \log \log n) \) and \( r_2 \leq r \leq n/8 \),
\[
\mathbb{P}(B(a) \xrightarrow{2 \to r} B(n)) \leq c_{20} \exp(-c_{21}n).
\]
The same bound holds for \( \mathbb{P}(B(a) \xrightarrow{2 \to r} B(n) \text{ in } K) \) for a cube \( K \) which contains the origin and \( K \cap \partial B(n) \neq \emptyset \).
Proof. As usual $\mathcal{P}$ will denote a Poisson process of intensity one. Starting with $d = 2$, let $\sigma((a,b);r,j)$ denote the probability of an occupied crossing of the rectangle $[0,a] \times [0,b]$ at level $r$ in the $j$-th direction, $j = 1, 2$; that is

$$\sigma((a,b);r,1) = \mathbb{P}(\mathcal{P}^{(r)} \text{ contains a curve } \gamma \subset [0,a] \times [0,b] \text{ such that } \gamma \text{ intersects both } S_1 \text{ and } S_2)$$

where $S_1 = \{0\} \times [0,b]$ and $S_2 = \{a\} \times [0,b]$ and define $\sigma((a,b);r,2)$ similarly. First we note that $\sigma^2((m,3m);r,c_1) \geq \kappa_0$ whenever $m > r_c$. The reason is as follows. Since $\sigma^2((m,3m);r,1)$ is a continuous function of $r$, $\sigma((m,3m);r,c_1) < \kappa_0$ would imply that $\sigma((m,3m);r,1) < \kappa_0$ for $r \in (r_c, r_c + \delta)$ for a small $\delta > 0$. Then Lemma 3.3 of [27] would imply exponential decay of $\mathbb{P}(\text{diameter}(\mathcal{C}_0) \geq n \text{ at level } r)$, a contradiction since $r > r_c$.

Now, the proof of Lemma 4.4 of [27] applies to occupied crossings as well. Since $\sigma((m,3m);r,c_1) \geq \kappa_0$ for $m > r_c$, the arguments of Lemma 4.4 of [27] would furnish positive constants $f(t)$ for each $t > 0$ such that

$$\sigma((m,(1+t)m);r,c_1) \geq f(t).$$

Applying Theorem 2.1 of [7] with the parameters $h = \ell/(1+t)$ and $b = \ell/(1+t)^2$ with $t$ small enough so that $2/(1+t)^2 - 1/2 > 1 + \epsilon$ (for some positive $\epsilon$) and $(1+t)^2 < 4/3$ and $\ell$ large so that $h > 4r_c$ and $b > \ell/2 + 2r_c$, we get

$$\sigma\left(\left(\ell \left[\left(\frac{2}{(1+t)^2} - \frac{1}{2}\right) + r_c, \frac{\ell}{1+t} - 2r_c\right) ; r_c, 1\right)\right) \geq c\sigma\left(\left(\frac{\ell}{1+t} + r_c, \frac{\ell}{1+t} - 4r_c\right) ; r_c, 1\right) \times \sigma\left(\left(\frac{\ell}{1+t} + 3r_c\right) ; r_c, 2\right)^2$$

for large $\ell$. Hence

$$\sigma\left((\ell(1+\epsilon), \ell) ; r_c, 1\right) \geq c\sigma\left(\left(\frac{\ell}{1+3t/4} + r_c, \frac{\ell}{1+5t/4}\right) ; r_c, 1\right) \times \sigma\left(\left(\frac{\ell}{1+t/2}\right) ; r_c, 2\right)^2 \geq cf \left(\frac{(1+3t/4)^2}{1+5t/4} - 1\right)^4 \times f\left(\frac{t}{2}\right)^2$$

for every $\ell$ bigger than a fixed threshold $\ell_0$. Hence Lemma 3.1 of [7] yields

$$\sigma((3\ell, \ell) ; r_c, 1) \geq \kappa_1$$

for a positive constant $\kappa_1$ and $\ell \geq \ell_0$.

Let $A_k$ be the event that there is an occupied circuit at level $r_c$ in the annulus $B(3\ell_k/2) - B(\ell_k/2)$, where $\ell_k = 3\ell_{k-1} + 4r_c$ and $\ell_1 = \max(2a +
2r, \ell_0). FKG inequality and (A.1) gives $P(A_k) \geq \kappa_4^4$. Hence
\[
P(B(a) \xrightarrow{2/\ell} B(n)) \leq P(A_1^c \cap \ldots \cap A_t^c)
= \prod_{k=1}^t P(A_k^c) \leq (1 - \kappa_4^4)^t
\]
where $3\ell_t/2 + r_c \leq n - r < 3\ell_{t+1}/2 + r_c$. This yields the desired bound.

When dealing with $P(B(a) \xrightarrow{2/\ell} B(n))$ in $K$, one needs to construct occupied crossings with Poisson points in a quadrant (Figure 4).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.png}
\caption{Figure 4.}
\end{figure}

For this purpose we define the quantity
\[
\sigma^+((a,b); r_c, 1) = P((P \cap \mathbb{R}_+^2)(r_c) \cap \mathbb{R}_+^2 \text{ contains a curve } \gamma \subset R(a,b) \text{ such that } \gamma \text{ intersects both } L_1 \text{ and } L_2)
\]
where $R(a,b) := [0,a] \times [r_c, r_c + b]$, $L_1 = \{0\} \times [r_c, r_c + b]$ and $L_2 = \{a\} \times [r_c, r_c + b]$.

Then it is easy to see that for $\ell$ bigger than a fixed threshold,
\[
\kappa_1 \leq \sigma((3\ell, \ell); r_c, 1)
\leq \sigma^+((3\ell, \ell); r_c, 1) + \sigma^+((3\ell - r_c, \ell); r_c, 1)
\]
and
\[
\sigma^+((3\ell, \ell); r_c, 1) \geq c' \sigma^+((3\ell - r_c, \ell); r_c, 1)
\]
for some (positive) universal constant $c'$. Hence
\[
\sigma^+((3\ell, \ell); r_c, 1) \geq \kappa_1/(1 + 1/c').
\]
The rest will follow from standard arguments.

Let us now turn to the case $d \geq 3$. Since $r_2 > r'_c(= r'_c(1))$ and $r'_c(\lambda)$ is a continuous function of $\lambda$, we can choose a $\lambda_0 < 1$ so that $r'_c(\lambda_0) < r_2$ and then choose $L$ large enough so that $r_c(\lambda_0, \mathbb{R}_+^2 \times [0, L]^{d-2}) < r_2$. 
Proof. For $\exists$ Lemma A.1.

Following the notation in [21], we define

$$S_t := [-t, t]^2 \times [0, L]^{d-2}$$

and

$$T_t := [-t, t]^{d-1} \times [0, L],$$

for $t \geq 1$. Given $\omega \in \mathcal{X}(\mathbb{R}^d)$ and two points $x, y \in A \subset \mathbb{R}^d$, we will write $x \overset{A}{\sim} y$ if $x, y \in (\omega \cap A)^{(r_2)}$ and are in the same connected component of $A \cap (\omega \cap A)^{(r_2)}$. The following Lemma is the analogue of Lemma 7.78 in [21].

**Lemma A.1.** There exists a $\delta > 0$ such that for every $t \geq 1$

\begin{equation}
\mathbb{P} \left( x \overset{\omega}{\sim}_{S_t} y \right) \geq \delta \text{ for every } x, y \in S_t
\end{equation}

and

\begin{equation}
\mathbb{P} \left( x \overset{\omega}{\sim}_{T_t} y \right) \geq \delta \text{ for every } x, y \in T_t.
\end{equation}

**Proof.** For $s \geq 1$ define

$$H_1(s) = \{0\} \times [0,s] \times [0,L]^{d-2}, \quad H_2(s) = \{s\} \times [0,s] \times [0,L]^{d-2}$$

$$H_3(s) = [0,s] \times \{0\} \times [0,L]^{d-2}, \quad H_4(s) = [0,s] \times \{s\} \times [0,L]^{d-2}$$

and let $U_s := [0,s]^2 \times [0,L]^{d-2}$. Let

$$x_1 = (s,s,0,\ldots,0), \quad x_2 = (0,\ldots,0)$$

$$x_3 = (0,s,0,\ldots,0), \quad x_4 = (s,0,\ldots,0).$$

For a configuration $\omega$ in $\mathbb{R}^d$, we write $x_j \overset{U_s}{\sim} H_i$, if $x_j \in (\omega \cap U_s)^{(r_2)}$ and the connected component of $U_s \cap (\omega \cap U_s)^{(r_2)}$ containing $x_j$ intersects the set

$$\{y \in U_s : d(y, H_i(s)) \leq r_2\}.$$

Since $r_2 > r_c(\lambda_0, \mathbb{R}^2_+ \times [0,L]^{d-2})$, we have

$$\mathbb{P} \left( x_2 \overset{P_i}{\sim}_{\mathbb{R}^2_+ \times [0,L]^{d-2}} \infty \right) = \theta > 0,$$

the event in the parentheses defined in the obvious way. Hence,

$$\mathbb{P} \left( x_2 \overset{P_i}{\sim}_{U_s} H_2(s) \right) + \mathbb{P} \left( x_2 \overset{P_i}{\sim}_{U_s} H_4(s) \right) \geq \theta$$

for $s > r_2$. From this and symmetry we conclude that

$$\mathbb{P} \left( x_j \overset{P_i}{\sim}_{U_s} H_j(s) \right) \geq \theta/2, \quad j = 1, \ldots, 4$$

whenever $s > r_2$. Noting that a similar assertion is trivial for $1 < s \leq r_2$, we conclude (by an application of FKG) that

\begin{equation}
\mathbb{P} \left( x_j \overset{P_i}{\sim}_{U_s} H_j(s), \quad j = 1, \ldots, 4 \right) \geq \delta_1 \text{ for every } s \geq 1
\end{equation}
for a positive constant $\delta_1$. Let us denote the event on the left side of (A.4) \(E\) by $E$. If $x_j \in (P_1 \cap U_s)^{(r_2)}$, let $C(x_j)$ be the connected component of $U_s \cap (P_1 \cap U_s)^{(r_2)}$ containing $x_j$, otherwise set $C(x_j) = \{x_j\}$. Define

$$F := \{\text{there is a connected component of } U_s \cap (P_2 \cap U_s)^{(r_2)}\}$$

which intersects both $C(x_1)$ and $C(x_2)$. Then

$$P \left( x \sim_{S_t} 0 \right) \geq P \left( x \sim_{S_t} x' \right) \cdot P \left( x' \sim_U y_1 \right) \cdot P \left( y_1 \sim_{U''} 0 \right)$$

for a positive constant $\delta_2$ depending on $L$ and $\lambda_0$. By a similar argument, the same lower bound continues to hold for $P \left( x \sim_{U_s^j} x \right)$ for $i, j \leq 4$.

It is enough to prove (A.2) for $y = 0$ and $x = (x_1, \ldots, x_d) \in S_t$ for which $0 \leq x_1 \leq x_2$. If $1 \leq x_1 \leq x_2 - 1$, define $x' = (x_1, x_2, 0, \ldots, 0)$, $y_1 = (0, x_2 - x_1, 0, \ldots, 0)$, $U = [0, x_1] \times [x_2 - x_1, x_2] \times [0, L]^{d-2}$ and $U' = [0, x_2 - x_1]^2 \times [0, L]^{d-2}$. Then

$$P \left( x \sim_{S_t} 0 \right) \geq P \left( x \sim_{S_t} x' \right) \cdot P \left( x' \sim_{U'} y_1 \right) \cdot P \left( y_1 \sim_{U''} 0 \right)$$

If $x_1 < 1 \leq x_2$, define $y_2 = (0, x_2, 0, \ldots, 0)$ and $U'' = [0, x_2] \times [0, x_2] \times [0, L]^{d-2}$. Then

$$P \left( x \sim_{S_t} 0 \right) \geq P \left( x \sim_{S_t} x' \right) \cdot P \left( x' \sim_{U''} y_2 \right) \cdot P \left( y_2 \sim_{U''} 0 \right)$$

for a universal constant $\delta_3 > 0$. The other possibilities, namely $x_1 \leq x_2 < 1$ and $x_2 - x_1 < 1 \leq x_1$ can be handled similarly. This proves (A.2).

The inequality (A.3) follows from (A.2) and the argument for this is similar to the one used in the proof of Lemma 7.78 in \[21\], so we omit it. \(\square\)

Going back to the proof of the Proposition, we note that

(A.5)

$$P(B(a) \sim_{\frac{2}{r}} B(n)) = E \left[ P \left( B(a) \sim_{\frac{2}{r}} B(n) \mid P \cap (B(a)^{(2r)} - B(a)) \right) \right].$$

Given a realization $\eta = \{\eta_1, \ldots, \eta_s\}$ of $P \cap (B(a)^{(2r)} - B(a))$,

$$\{B(a) \sim_{\frac{2}{r}} B(n)\} \subset \bigcup_{1 \leq i \neq j \leq s} \{C(\eta_i) \cap B(n)^{(2r)} \neq \emptyset; C(\eta_j) \cap B(n)^{(2r)} \neq \emptyset; C(\eta_i) \cap C(\eta_j) = \emptyset\}$$
where $C(\eta_i)$ denotes the $r$-cluster in $(B(n) - B(a))$ containing $\eta_i$. Hence

\begin{equation}
\P(\eta) (B(a) \nrightarrow B(n)) \leq \sum_{1 \leq i \neq j \leq k} \P(\eta_i \cap (B(n)(2r) \neq \emptyset) ; C(\eta_i) \cap C(\eta_j) = \emptyset)
\end{equation}

where $\P(\eta)$ denotes probability conditional on

$$\mathcal{P}(\emptyset) \cap (B(a)(2r) - B(a)) = \{\eta_1, \ldots, \eta_k\}.$$

Fix an $\epsilon \in (0, r_2/2)$ and let $D_k = B(a + 2r + (L + \epsilon)k)$ and let $C_k(\eta_i)$ be the $r$-cluster in $(D_k - B(a))$ containing $\eta_i$. Then

\begin{equation}
\P(\eta_i) (\eta_i \cap B(n)(2r) \neq \emptyset) ; C(\eta_i) \cap (B(n)(2r) \neq \emptyset) ; C(\eta_i) \cap C(\eta_j) = \emptyset)
\end{equation}

where $A_k := \{C_k(\eta_i) \cap D_k(2r) \neq \emptyset ; C_k(\eta_j) \cap D_k(2r) \neq \emptyset ; C_k(\eta_i) \cap C_k(\eta_j) = \emptyset\}$ and $t$ is the integer such that $a + 2r + (L + \epsilon)t \leq n < a + 2r + (L + \epsilon)(t + 1)$. Then it suffices to get an upper bound on $\P(A_k|A_k)$ which is uniform in $k$.

Let $\delta$ be as in Lemma A.1. Define $S_k := \{D_k - B(a + 2r + (L + \epsilon)(k - 1) + \epsilon)\}$.

Let us work conditional on the event $\{\mathcal{P} \cap (D_k - B(a)) = \emptyset\}$ where $\xi_k$ is a configuration for which $A_k$ is true and $\xi_k \cap (B(a)(2r) - B(a)) = \emptyset$. Let us denote the corresponding conditional probability measure by $\P(\xi_k)$.

Consider the following possibilities:

(I) both $C_k(\eta_i)(2r)$ and $C_k(\eta_j)(2r)$ intersect $S_{k+1}$,

(II) neither of $C_k(\eta_i)(2r)$ and $C_k(\eta_j)(2r)$ intersects $S_{k+1}$ and

(III) exactly one of $C_k(\eta_i)(2r)$ and $C_k(\eta_j)(2r)$ intersect $S_{k+1}$.

If (I) is true for $\xi_k$, then there exists a $\zeta > 0$ depending only on $\epsilon$ and $r_2$ such that we can place two cubes $Q_1$ and $Q_2$ each of side length $\zeta$ in $C_k(\eta_i)(2r) \cap \{D_k + (S_{k+1} + D_k)\}$ and $C_k(\eta_j)(2r) \cap \{D_k + (S_{k+1} + D_k)\}$ respectively. Let $A_i = \{\mathcal{P} \cap Q_i = \emptyset\}$ for $i = 1, 2$. Conditioning further on $A_1$ and $A_2$, we can choose two points $x_1$ and $x_2$ in $S_{k+1}$ such that $d(x_m, y_m) \leq \epsilon$ for some $y_m \in S_{k+1}$, $i = 1, 2$. If $E = \{x_1 \nrightarrow x_2\}$, then $A_k$ does not happen whenever the event $A_1 \cap A_2 \cap E$ is true. An application of Lemma A.1 together with arguments similar to the ones in the proof of Lemma 7.89 in [21] will yield

$$\inf_{x_1, x_2 \in S_{k+1}} \P(E) \geq \delta^{d+2}.$$ 

Hence

\begin{equation}
\P(\eta_i) (A_k|A_k) \leq \sup_{\xi_k} \P(\xi_k) (A_k|A_k)
\leq 1 - \P(A_1)\P(A_2)\delta^{d+2} \leq 1 - \alpha,
\end{equation}
the constant $\alpha$ equals $\delta d^2 (1 - \exp(-\zeta d))^2$.

If (II) is true for the configuration $\xi_k$, then

$$\mathbb{P}_{\xi_k}(A_{k+1}) \leq \mathbb{P}_{\xi_k}(A_{k+1}|B_{k+1})$$

where

$$B_{k+1} = \{P \cap C_k(\eta_i)^{(2r)} \cap \{D_{k+1} - (S_{k+1} \cup D_k)\} \neq \emptyset\} \cap \{P \cap C_k(\eta_j)^{(2r)} \cap \{D_{k+1} - (S_{k+1} \cup D_k)\} \neq \emptyset\}.$$  

Conditioning further on $P \cap (D_{k+1} - S_{k+1})$ being in $B_{k+1}$, we can choose a point $x_1$ (resp. $x_2$) in $S_{k+1}$ such that $d(x_1, y_1) \leq \epsilon$ (resp. $d(x_2, y_2) \leq \epsilon$) for some $y_1$ (resp. $y_2$) in $P \cap \{D_{k+1} - (S_{k+1} \cup D_k)\} \cap C_k(\eta_i)^{(2r)}$ (resp. $P \cap \{D_{k+1} - (S_{k+1} \cup D_k)\} \cap C_k(\eta_j)^{(2r)}$). As before, $A_{k+1}$ does not happen if $x_1 \xrightarrow{P} S_{k+1} x_2$. Hence

$$(\text{A.9}) \quad \mathbb{P}_{\eta}(A_{k+1}|A_k) \leq (1 - \delta d^2).$$

The possibility (III) can be handled in a similar way, in this case we will get

$$(\text{A.10}) \quad \mathbb{P}_{\eta}(A_{k+1}|A_k) \leq 1 - \delta d^2 (1 - \exp(-\zeta d)).$$

From (A.5) through (A.10), we conclude that

$$\mathbb{P}(B(a) \xrightarrow{2_r} B(n)) \leq \exp(-cn)\mathbb{E} \left[|P \cap (B(a)^{(2r)} - B(a))|^2\right] \leq c_{20} \exp(-c_{21}n),$$

as desired.

Again the bound for $\mathbb{P}(B(a) \xrightarrow{2_r} B(n)$ in $K)$ follows from similar arguments and we do not repeat them.

\textbf{Acknowledgments.} The authors are indebted to Larry Goldstein and Ümit Islak for their valuable help in improving the manuscript and checking the proof.

\textbf{References}

[1] Aizenman, M., Burchard, A., Newman, C.M. and Wilson, D.B. (1999). Scaling limits for minimal and random spanning trees in two dimensions. \textit{Random Struct. Alg.} 15 no. 3-4, 319-367.

[2] Aizenman, M., Kesten, H. and Newman, C. M. (1987). Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. \textit{Commun. Math. Phys.}, 111 no. 4, 505–531.

[3] Aldous, D. and Steele, J.M. (1992). Asymptotics for Euclidean minimal spanning trees on random points. \textit{Probab. Theory Related Fields} 92 no. 2, 247-258.

[4] Alexander, K.S. (1994). Rates of Convergence of Means for Distance-Minimizing Subadditive Euclidean Functionals. \textit{Ann. Appl. Probab.} 4 no. 3, 902-922.

[5] Alexander, K.S. (1995). Percolation and minimal spanning forests in infinite graphs. \textit{Ann. Probab.} 23 no. 1, 87-104.
[6] **Alexander, K.S. and Molchanov, S.A.** (1994). Percolation of level sets for two-dimensional random fields with lattice symmetry. *J. Statist. Phys.* **77** no. 3-4, 627-643.

[7] **Alexander, K.S.** (1996). The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees. *Ann. Appl. Probab.* **6** no. 2, 466-494.

[8] **Avram, F. and Bertsimas. D.** (1992). The minimum spanning tree constant in geometrical probability and under the independent model: A unified approach. *Ann. Appl. Probab.* **2** no. 1, 113-130.

[9] **Avram, F. and Bertsimas. D.** (1993). On central limit theorems in geometrical probability. *Ann. Appl. Probab.* **3** no. 1, 113-130.

[10] **Bai, Z.D., Lee, S. and Penrose, M.D.** (2006). Rooted edges of a minimal directed spanning tree on random points. *Adv. Appl. Probab.* **38** no. 1, 1-30.

[11] **Baldi, P. and Rinott, Y.** (1989). On Normal Approximations of Distributions in Terms of Dependency Graphs. *Ann. Probab.* **17** no. 4, 1646-1650.

[12] **Beardwood, J., Halton, J.H. and Hammersley, J.M.** (1959). The shortest path through many points. *Proc. Cambridge Philos. Soc.* **55** no. 4, 299-327.

[13] **Bhattacharya, A.G., and Roy, R.** (2004). On a random directed spanning tree. *Adv. Appl. Probab.* **36** no. 1, 19-42.

[14] **Bollobás, B. and Riordan, O.** (2006). *Percolation*. Cambridge University Press.

[15] **Burton, R. and Keane, M.** (1989). Density and uniqueness in percolation. *Commun. Math. Phys.* **121** no. 3, 501-505.

[16] **Camia, F., Fontes, L.R. and Newman, C.M.** (2006). Two-dimensional scaling limits via marked nonsimple loops. *Bull. Braz. Math. Soc.* **37** no. 4, 537-559.

[17] **Cerf, R.** (2013). A lower bound on the two-arms exponent for critical percolation on the lattice. Preprint. Available at http://arxiv.org/abs/1306.3105

[18] **Chatterjee, S.** (2008). A new method of normal approximation. *Ann. Probab.* **36** no. 4, 1584-1610.

[19] **Chatterjee, S. and Soundararajan, K.** (2010). Random multiplicative functions in short intervals. *To appear in Int. Math. Res. Notices.*

[20] **Gandolfi, A., Grimmett, G. and Russo, L.** (1988). On the uniqueness of the infinite cluster in the percolation model. *Comm. Math. Phys.* **114** no. 4, 549-552.

[21] **Grimmett, G.** (1999). *Percolation*. Springer-Verlag, New York.

[22] **Häggström, O.** (1995). Random-cluster measures and uniform spanning trees. *Stochastic Process. Appl.* **59** no. 2, 267-275.

[23] **Kesten, H. and Lee. S.** (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **6** no. 2, 495-527.

[24] **Lee, S.** (1997). The central limit theorem for Euclidean minimal spanning trees I. *Ann. Appl. Probab.* **7** no. 4, 996-1020.

[25] **Lee, S.** (1999). The central limit theorem for Euclidean minimal spanning trees II. *Ann. Appl. Probab.* **31** no. 4, 969-984.

[26] **Lyons, R., Peres, Y. and Schramm, O.** (2006). Minimal spanning forests. *Ann. Probab.* **34** no. 5, 1665-1692.

[27] **Meester, R. and Roy, R.** (1996). *Continuum percolation*. Cambridge University Press.

[28] **Penrose, M.D.** (1996). The random spanning tree in high dimensions. *Ann. Probab.* **24** no. 4, 1903-1925.

[29] **Penrose, M.D.** (1997). The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.* **7** no. 2, 340-361.

[30] **Penrose, M.D. and Wade, A.R.** (2004). Random minimal directed spanning trees and dickman-type distributions. *Adv. Appl. Probab.* **36** no. 3, 691-714.

[31] **Penrose, M.D. and Yukich, J.E.** (2005). Normal approximation in geometric probability. In *Steins Method and Applications*, eds. Barbour, A.D. and Chen, L.H.Y., Lecture Notes Series, Institute for Mathematical Sciences, **5**, pp. 37-58. World Scientific, Singapore.
[32] **Pete, G., Garban, C. and Schramm, O.** (2009). The scaling limit of the minimal spanning tree - a preliminary report. *Preprint*. Available at http://arxiv.org/pdf/0909.3138v1.pdf

[33] **Roy, R.** (1990). The Russo-Seymour-Welsh theorem and the equality of critical densities and the ”dual” critical densities for continuum percolation on $\mathbb{R}^2$. *Ann. Probab.* **18** no. 4, 1563-1575.

[34] **Steele, J.M.** (1981). Subadditive Euclidean functionals and nonlinear growth in geometric probability. *Ann. Probab.* **9** no. 3, 365-376.

[35] **Steele, J.M.** (1988). Growth rates of Euclidean minimal spanning trees with power weighted edges. *Ann. Probab.* **16** no. 4, 1767-1787.

[36] **Tanemura, H.** (1993). Behavior of the supercritical phase of a continuum percolation model on $\mathbb{R}^d$. *J. Appl. Probab.* **30** no. 2, 382-396.

**Courant Institute of Mathematical Sciences**
**New York University**
**251 Mercer Street**
**New York, NY-10012**
**United States of America**

*E-mail address:* sourav@cims.nyu.edu, sen@cims.nyu.edu