# Diameter and Broadcast Time of Random Geometric Graphs in Arbitrary Dimensions

Tobias Friedrich[1], Thomas Sauerwald[1], and Alexandre Stauffer[2]

[1] Max-Planck-Institut für Informatik, Saarbrücken, Germany
[2] Computer Science Division, University of California, Berkeley, USA

**Abstract.** A random geometric graph (RGG) is defined by placing $n$ points uniformly at random in $[0, n^{1/d}]^d$, and joining two points by an edge whenever their Euclidean distance is at most some fixed $r$. We assume that $r$ is larger than the critical value for the emergence of a connected component with $\Omega(n)$ nodes. We show that, with high probability (w.h.p.), for any two connected nodes with a minimum Euclidean distance of $\omega(\log n)$, their graph distance is only a constant factor larger than their Euclidean distance. This implies that the diameter of the largest connected component is $\Theta(n^{1/d}/r)$ w.h.p.

We also analyze the following randomized broadcast algorithm on RGGs. At the beginning, only one node from the largest connected component of the RGG is informed. Then, in each round, each informed node chooses a neighbor independently and uniformly at random and informs it. We prove that w.h.p. this algorithm informs every node in the largest connected component of an RGG within $\Theta(n^{1/d}/r + \log n)$ rounds.

## 1 Introduction

We study Random Geometric Graphs (RGGs) in $d \geqslant 2$ dimensions. An RGG is a graph resulting from placing $n$ nodes independently and uniformly at random on $[0, n^{1/d}]^d$ and creating edges between pairs of nodes if and only if their Euclidean distance is at most $r$. These graphs have been studied intensively in relation to subjects such as cluster analysis, statistical physics, and wireless sensor networks [25]. Traditionally, most work on RGGs is restricted to two dimensions. However, wireless sensor networks also expand in three dimensions. Examples are sensors in water bodies [1] and sensor networks based on the use of flying anchors [19]. Another motivation for RGGs in arbitrary dimensions is multivariate statistics of high-dimensional data [23]. In this case the coordinates of the nodes of the RGG represent different attributes of the data. The metric imposed by the RGG then depicts the similarity between data elements in the high-dimensional space. Also in bioinformatics, RGGs in up to four dimensions have been observed to give an excellent fit for various global and local measures of protein-protein interaction networks [16].

Several algorithms and processes have been studied on RGGs. One prominent example is the cover time of random walks. [3] considered RGGs in two

dimensions when the coverage radius is a constant larger than the one that assures the RGG to be connected with probability $1 - o(1)$. They proved that in this regime, the cover time of an RGG is $\Theta(n \log n)$ with probability $1 - o(1)$, which is optimal up to constant factors. This has been improved by [5] who gave a more precise estimate of the cover time on RGGs that also extends to larger dimensions. However, all these works are restricted to the case where the probability that the RGG is connected goes to 1 as $n \to \infty$.

We are interested in a wider range for $r$. All the following results hold for the regime where the RGG is likely to contain a connected component with $\Omega(n)$ nodes. [4] proved for RGGs in $d = 2$ dimensions that with probability $1 - \mathcal{O}(n^{-1})$ any two connected nodes with a minimum Euclidean distance of $\Omega(\log^{3.5} n/r^2)$, their graph distance is only a constant factor larger than their Euclidean distance. We establish this result for all dimensions $d \geqslant 2$ under the weaker condition that the minimum Euclidean distance is $\omega(\log n)$. For this, we have to employ a different proof technique since the approach of [4] strongly depends on restrictions imposed by the geometry in two dimensions. Our result implies that the diameter of the largest connected component is $\mathcal{O}(n^{1/d}/r)$ with high probability[1], which was open for $d \geqslant 3$ and matches the corresponding bound for $d = 2$ [4, 7]. Our techniques are inspired by percolation theory and we believe them to be useful for other problems, like estimating the cover time for the largest connected component of RGGs.

## Broadcasting Information

We use the forementioned structural result of RGGs to study the problem of broadcasting information in RGGs. We study the well known randomized rumor spreading algorithm which is also known as push algorithm [10]. In this algorithm, in each round each informed node chooses a neighbor independently and uniformly at random and informs it. We are interested in the runtime, i.e., how long it takes to spread a piece of information from an arbitrary node of the largest connected component to all other connected nodes.

The obvious lower bound of this process on an arbitrary graph $G$ is $\Omega(\text{diam}(G) + \log n)$, where $\text{diam}(G)$ denotes the diameter of the largest connected component. A matching upper bound of $\mathcal{O}(\text{diam}(G) + \log n)$ is known for complete graphs [13, 24], hypercubes [10], expander graphs [27, 12], several Cayley graphs [8], and RGGs in two dimensions [4]. In this paper we prove that also RGGs in $d \geqslant 3$ dimensions allow an optimal broadcast time of $\mathcal{O}(\text{diam}(G) + \log n) = \mathcal{O}(n^{1/d}/r + \log n)$ w.h.p. This generalizes the two-dimensional result of [4] and significantly improves upon the general bound of $\mathcal{O}(\Delta \cdot (\text{diam}(G) + \log n))$ [10] since for sparse RGGs (where $r = \Theta(1)$) the maximum degree is $\Delta = \Theta(\log n/ \log \log n)$. Note that also for connected RGGs our result implies that all nodes get informed after $\mathcal{O}(n^{1/d}/r + \log n)$ rounds.

---

[1] By with high probability (short: w.h.p.), we refer to an event that holds with probability at least $1 - \mathcal{O}(n^{-1})$.

## 2    Precise Model and Results

We consider the following random broadcast algorithm also known as the push algorithm (cf. [10]). We are given an undirected graph $G$. At the beginning, called round 0, a node $s$ of $G$ owns a piece of information, i.e., it is informed. In each subsequent round $1, 2, \ldots$, every informed node chooses a neighbor independently and uniformly at random and informs that neighbor. We are interested in the runtime of this algorithm, which is the time until every node in $G$ gets informed; in case of $G$ being disconnected, we require every node in the same connected component as $s$ to get informed. The runtime of this algorithm is a random variable denoted by $\mathcal{R}(s, G)$. Our aim is to prove bounds on $\mathcal{R}(s, G)$ that hold with high probability, i.e., with probability $1 - \mathcal{O}(n^{-1})$.

We study $\mathcal{R}(s, G)$ for the case of a random geometric graph $G$ in arbitrary dimension $d \geqslant 2$. We define the random geometric graph in the space $\Omega := [0, n^{1/d}]^d$ equipped with the Euclidean norm, which we denote by $\| \cdot \|_2$. The most natural definition of RGG is stated as follows.

**Definition 1 (cf. [23]).** *Let $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ be points in $\Omega$ chosen independently and uniformly at random. The random geometric graph $\mathcal{G}(\mathcal{X}_n; r)$ has node set $\mathcal{X}_n$ and edge set $\{(x, y) \colon x, y \in \mathcal{X}_n, \|x - y\|_2 \leqslant r\}$.*

In our analysis, it is more advantageous to use to the following definition.

**Definition 2 (cf. [23]).** *Let $N_n$ be a Poisson random variable with mean $n$ and let $\mathcal{P}_n = \{X_1, X_2, \ldots, X_{N_n}\}$ be points chosen independently and uniformly at random from $\Omega$; i.e., $\mathcal{P}_n$ is a Poisson Point Process over $\Omega$ with intensity 1. The random geometric graph $\mathcal{G}(\mathcal{P}_n; r)$ has node set $\mathcal{P}_n$ and edge set $\{(x, y) \colon x, y \in \mathcal{P}_n, \|x - y\|_2 \leqslant r\}$.*

The following basic lemma says that any result that holds in the setting of Definition 2 with sufficiently large probability holds with similar probability in the setting of Definition 1.

**Lemma 1.** *Let $\mathcal{A}$ be any event that holds with probability at least $1 - \alpha$ in $\mathcal{G}(\mathcal{P}_n; r)$. Then, $\mathcal{A}$ also holds in $\mathcal{G}(\mathcal{X}_n; r)$ with probability $1 - \mathcal{O}(\alpha\sqrt{n})$.*

Henceforth, we consider an RGG given by $G = \mathcal{G}(\mathcal{P}_n; r)$, and refer to $r$ as the coverage radius of $G$. It is known that, for $d \geqslant 2$, there exists a critical value $r_c = r_c(d) = \Theta(1)$ such that if $r > r_c$, then with high probability the largest connected component of $G$ has cardinality $\Omega(n)$. On the contrary, if $r < r_c$, each connected component of $G$ has $\mathcal{O}(\log n)$ nodes with probability $1 - o(1)$ [23]. The exact value of $r_c$ is not known, though some bounds have been derived in [18]. In addition, if $r^d \geqslant \frac{\log n + \omega(1)}{b_d}$, where $b_d$ is the volume of the $d$-dimensional ball of radius 1, then $G$ is connected with probability $1 - o(1)$ [21, 22].

Our main result is stated in the next theorem. It shows that if $r > r_c$, then for all $s$ inside the largest connected component of $G$, $\mathcal{R}(s, G) = \mathcal{O}(n^{1/d}/r + \log n)$ with probability $1 - \mathcal{O}(n^{-1})$. Note that $r_c$ does not depend on $n$, but if $r$ is regarded as a function of $n$, then here and in what follows, $r > r_c$ means that this strict inequality must hold in the limit as $n \to \infty$.

**Theorem 2.** *For a random geometric graph $G = \mathcal{G}(\mathcal{P}_n; r)$ in $d \geqslant 2$ dimensions, if $r > r_c$, then $\mathcal{R}(s, G) = \mathcal{O}(n^{1/d}/r + \log n)$ with probability $1 - \mathcal{O}(n^{-1})$ for all nodes $s$ inside the largest connected component of $G$.*

The proof of Theorem 2, which is similar to the proof of [4, Theorem 2.2] and is given in [], requires an upper bound on the length of the shortest path between nodes of $G$. Our result on this matter, which is stated in the next theorem, gives that for any two nodes that are sufficiently distant in $\Omega$, the distance between them in the metric induced by $G$ is only a constant factor larger than the optimum with probability $1 - \mathcal{O}(n^{-1})$. In particular, this result implies that the diameter of the largest connected component of $G$ is $\mathcal{O}(n^{1/d}/r)$, a result previously known only for two dimensions and values of $r$ that give a connected $G$ with probability $1 - o(1)$.

For all $v_1, v_2 \in G$, we say that $v_1$ and $v_2$ are connected if there exists a path in $G$ from $v_1$ to $v_2$, and define $d_G(v_1, v_2)$ as the distance between $v_1$ and $v_2$ on $G$, that is, $d_G(v_1, v_2)$ is the length of the shortest path from $v_1$ to $v_2$ in $G$. Also, we denote the Euclidean distance between the locations of $v_1$ and $v_2$ by $\|v_1 - v_2\|_2$. Clearly, the length of the shortest path between two nodes $v_1$ and $v_2$ in $G$ satisfies $d_G(v_1, v_2) \geqslant \|v_1 - v_2\|_2/r$.

**Theorem 3.** *If $d \geqslant 2$ and $r > r_c$, for any two connected nodes $v_1$ and $v_2$ in $G = \mathcal{G}(\mathcal{P}_n; r)$ such that $\|v_1 - v_2\|_2 = \omega(\log n)$, we obtain $d_G(v_1, v_2) = \mathcal{O}(\|v_1 - v_2\|_2/r)$ with probability $1 - \mathcal{O}(n^{-1})$.*

**Corollary 4.** *If $r > r_c$, the diameter of the largest connected component of $G = \mathcal{G}(\mathcal{P}_n; r)$ is $\mathcal{O}(n^{1/d}/r)$ with probability $1 - \mathcal{O}(n^{-1})$.*

The statement of Theorem 3 generalizes and improves upon Theorem 2.3 of [4] which only holds for $d = 2$ and $\|v_1 - v_2\|_2 = \Omega(\log^{3.5} n/r^2)$. The current paper not only improves upon the previous results, but also employs different proof techniques which are necessary to tackle the geometrically more involved case $d \geqslant 3$.

## 3   The Diameter of the Largest Connected Component

We devote this section to prove Theorem 3. We consider $G = \mathcal{G}(\mathcal{P}_n; r)$. Recall that we assume $r > r_c$ and $r = \mathcal{O}(\log^{1/d} n)$. (When $r = \omega(\log^{1/d} n)$, $G$ is connected with probability $1 - o(1)$ and Theorem 3 becomes a slightly different version of [7, Theorem 8].) Note also that $r = \Omega(1)$ since $r_c = \Theta(1)$. We show that, for any two connected nodes $v_1$ and $v_2$ of $G$ such that $\|v_1 - v_2\|_2 = \omega(\log n)$, we have $d_G(v_1, v_2) = \mathcal{O}(\|v_1 - v_2\|_2/r)$ with probability $1 - \mathcal{O}(n^{-1})$.

Before going to the proof, we establish some notation and discuss results for discrete lattices that we will use later. For $m \geqslant 0$, whose value we will set later, let $\mathbb{S}_m$ be the elements of $\mathbb{Z}^d$ contained in the cube of side length $m$ centered at the origin (i.e., $\mathbb{S}_m = \{i \in \mathbb{Z}^d : \|i\|_\infty \leqslant m/2\}$). Let $L$ be the graph with vertex set $\mathbb{S}_m$ such that an edge between two vertices $i, j \in \mathbb{S}_m$ exists if and only if
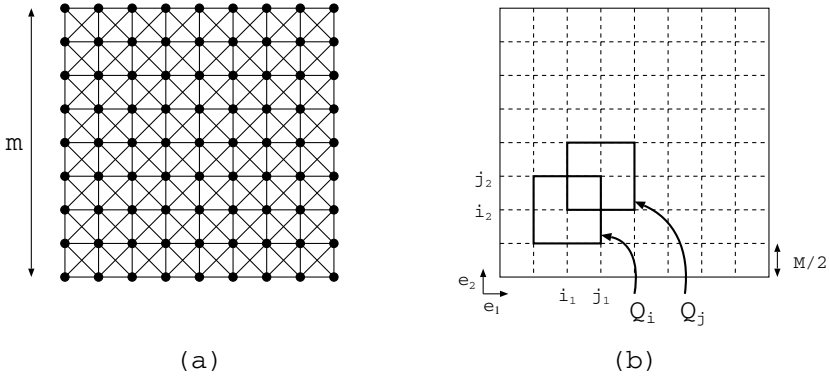
**Fig. 1.** (a) The graph $L$ over $\mathbb{S}_m$. (b) Illustration of the neighboring cubes $Q_i$ and $Q_j$.

$\|i - j\|_\infty = 1$ (see Figure 1(a)). It is easy to see that the maximum degree $\Delta$ of $L$ is $\Delta = 3^d - 1$. Let $X = (X_i)_{i \in \mathbb{S}_m}$ be a collection of binary random variables. For two vertices $i, j \in \mathbb{S}_m$, let $d_L(i,j)$ be their graph distance in $L$. Also, for any $i \in \mathbb{S}_m$ and $k \geqslant 0$, let $\mathcal{F}_k(i)$ be the $\sigma$-field generated by all $X_j$ with $d_L(i,j) > k$. Then, for $k \geqslant 0$ and $p \in (0,1)$, we say that $X$ is a $k$-*dependent site percolation process on $L$ with probability $p$* if, for any $i \in \mathbb{S}_m$, we have $\mathbf{Pr}\,[X_i = 1] \geqslant p$ and $\mathbf{Pr}\,[X_i = 1 \mid \mathcal{F}_k(i)] = \mathbf{Pr}\,[X_i = 1]$; i.e., $X_i$ is independent of any collection $(X_j)_j$ for which the distance between $i$ and $j$ in $L$ is larger than $k$ for all $j$ in the collection. Let $L(X)$ be the subgraph of $L$ induced by the vertices $i$ with $X_i = 1$. The following lemma is a direct application of a result by Liggett, Schonmann and Stacey [17, Theorem 1.3] that gives that $L(X)$ stochastically dominates an independent site percolation process.

**Lemma 5 ([17, Theorem 1.3]).** *For given constants $p \in (0,1)$ and $k \geqslant 0$, let $L(X)$ be the subgraph of $L$ obtained via a $k$-dependent site percolation process $X$ with probability $p$. If $p$ is large enough, then there exists a value $p' \in (0,p]$ depending only on $k$ and $p$ so that $L(X)$ stochastically dominates a collection of independent Bernoulli random variables with mean $p'$. Moreover, $p'$ can be made arbitrarily close to $1$ by increasing $p$ while $k$ is kept fixed.*

We will, from now on, assume that $p$ is so large that from Lemma 5 we can let $p'$ be arbitrarily large. In particular, $p'$ will be larger than the critical value for independent site percolation on the square lattice. Then if $(Y_i)_{i \in \mathbb{S}_m}$ is a collection of independent Bernoulli random variables with mean $p'$, we have that $L(Y)$ contains a giant component since the square lattice is contained in $L$ [15]. We will henceforth say that a vertex $i \in \mathbb{S}_m$ is *open* if $Y_i = 1$ and closed if $Y_i = 0$.

*Proof of Theorem 3.* We take two fixed nodes $v_1$ and $v_2$ satisfying the conditions of Theorem 3 and show that the probability that $v_1$ and $v_2$ are connected by a path and $d_G(v_1, v_2) = \omega(\|v_1 - v_2\|_2/r)$ is $\mathcal{O}(n^{-3})$. Then, we would like to take the union bound over all pairs of nodes $v_1$ and $v_2$ to conclude the proof of Theorem 3; however, the number of nodes in $G$ is a random variable and hence

the union bound cannot be employed directly. We employ the following lemma from [4] to extend the result to all pairs of nodes $v_1$ and $v_2$.

**Lemma 6 ([4, Lemma 3.1]).** *Let $\mathcal{E}(w_1, w_2)$ be an event associated to a pair of nodes $w_1, w_2 \in G = \mathcal{G}(\mathcal{P}_n, r)$. Assume that, for all pairs of nodes, $\mathbf{Pr}\left[\mathcal{E}(w_1, w_2)\right] \geqslant 1 - p$, with $p > 0$. Then,*

$$\mathbf{Pr}\left[\bigcap_{w_1, w_2 \in G} \mathcal{E}(w_1, w_2)\right] \geqslant 1 - 9n^2 p - e^{-\Omega(n)}.$$

Before establishing the result for two fixed nodes $v_1, v_2$, we describe our renormalization argument. Fix a sufficiently large constant $M > 0$. For each $i = (i_1, i_2, \ldots, i_d) \in \mathbb{Z}^d$, we define the cube $Q_i$ centered at $(i_1 M/2, i_2 M/2, \ldots, i_d M/2)$ whose sides have length $M$ and are parallel to the bases of $\mathbb{R}^d$ (see Figure 1(b)). Let $\mathcal{Q}$ be the set of $Q_i$ having center inside $\Omega$ and set $m$ so that $\Omega \cap \mathbb{Z}^d = \mathbb{S}_m$ (thus $Q_i \in \mathcal{Q}$ if and only if $i \in \mathbb{S}_m$). Note that $m = \Theta(n^{1/d})$ and the cubes in $\mathcal{Q}$ cover the whole of $\Omega$. We call two cubes $Q_i$ and $Q_j$ neighbors if $\|i - j\|_\infty \leqslant 1$. Note that in this case $i$ and $j$ are also neighbors in $L$. Therefore each cube has at most $\Delta = 3^d - 1$ neighbors, and there are at most $K = \lceil 2n^{1/d}/M \rceil^d = \Theta(n)$ cubes in $\mathcal{Q}$.

We say that a parallelogram $R$ in $\mathbb{R}^d$ has a *crossing component* if there exists a connected component inside $R$ such that, for each face of $R$, there exists at least one node of the component within distance $r$ of the face. Then, for each $i \in \mathbb{S}_m$, let $\mathcal{E}_i$ be defined as the event where all the following happen:

(i) For each neighbor $Q_j$ of $Q_i$, the parallelogram $Q_i \cap Q_j$ contains a crossing component.

(ii) $Q_i$ contains only one connected component with diameter larger than $M/5$.

Note that, when $\mathcal{E}_i$ happens for some $Q_i$, then (ii) above gives that the largest component of $Q_i$ intersects the crossing components of all parallelograms $Q_i \cap Q_j$, where $Q_j$ is a neighbor of $Q_i$. Moreover, for two $i$ and $j$ neighbors in $L$, we have that, if $\mathcal{E}_i$ and $\mathcal{E}_j$ happen, then the crossing components of $Q_i$ and $Q_j$ intersect. It is a direct consequence of a result of Penrose and Pisztora [20, Theorem 2] that, when $r > r_c$, for any $\varepsilon > 0$ there exists a $M_0$ depending only on $\varepsilon$ and $d$ such that, for any fixed $i$,

$$\mathbf{Pr}\left[\mathcal{E}_i\right] \geqslant 1 - \varepsilon \tag{1}$$

for all $M \geqslant M_0$. Since $M$ can be arbitrarily large, we set $M$ to be larger than $M_0$ and $2r$. So, in general, we can assume that $M/r$ is a constant.

Now we set $X_i = \mathbf{1}(\mathcal{E}_i)$ for all $i \in \mathbb{S}_m$. By construction, $\mathcal{E}_i$ does not depend on the events $\mathcal{E}_j$ for which $d_L(i, j) \geqslant 2$ since, in this case, the set of nodes in $Q_i$ and the set of nodes in $Q_j$ are disjoint. Therefore, $(X_i)_{i \in \mathbb{S}_m}$ is a 1-dependent site percolation process with probability $1 - \varepsilon$. Since $\varepsilon$ can be made arbitrarily small, we can apply Lemma 5 to find a collection of independent Bernoulli random variables $Y = (Y_i)_{i \in \mathbb{S}_m}$ with mean $p'$ so that $L(Y)$ is a subset of $L(X)$. Moreover,

$p'$ can be made arbitrarily close to 1 so that $L(Y)$ has a giant component with probability $1 - \exp\left(-\Theta\left(n^{1-1/d}\right)\right)$ [15].

We now show that, for any fixed pair of nodes $v_1, v_2$ of $G$ such that $\|v_1 - v_2\|_2 = \omega(\log n)$, we have that either $v_1$ and $v_2$ are in different connected components or $d_G(v_1, v_2) = \mathcal{O}(\|v_1 - v_2\|_2/r)$. Let $i_1$ be the closest vertex of $\mathbb{S}_m$ from $v_1$ and $i_2$ be the closest vertex of $\mathbb{S}_m$ from $v_2$. Clearly, $v_1 \in Q_{i_1}$ and $v_2 \in Q_{i_2}$. We use some ideas from Antal and Pisztora [2]. For any connected subset $H$ of $\mathbb{S}_m$, let $\partial H$ be the set of vertices of $\mathbb{S}_m \setminus H$ from which there exists an edge to a vertex in $H$; that is, $\partial H$ is the outer boundary of $H$. Note that $|\partial H| \leqslant \Delta |H|$. Let $L'(Y)$ be the graph induced by the *closed* vertices of $L$. For each $j \in \mathbb{S}_m$, if $j$ is closed, let $Z_j$ be the connected component of $L'(Y)$ containing $j$ and let $\widehat{Z}_j = \partial Z_j$. If $j$ is open, then set $Z_j = \emptyset$ and $\widehat{Z}_j = \{j\}$. Note that $Z_j$ only contains closed vertices and $\widehat{Z}_j$ only contains open vertices. Moreover, $\widehat{Z}_j$ *separates* $Z_j$ from $\mathbb{S}_m \setminus (Z_j \cup \widehat{Z}_j)$ in the sense that any path in $L$ from a vertex in $\{j\} \cup Z_j$ to a vertex in $\mathbb{S}_m \setminus (Z_j \cup \widehat{Z}_j)$ must contain a vertex of $\widehat{Z}_j$. Now, let $A_j = \cup_{k:\, \|k-j\|_\infty \leqslant 1} Z_k$. If $A_j \neq \emptyset$, set $\widehat{A}_j = \partial A_j$; otherwise set $\widehat{A}_j = \{k \in \mathbb{S}_m : \|k - j\|_\infty \leqslant 1\}$.

Now we give an upper bound for the tails of $|Z_j|$ and $|A_j|$.

**Lemma 7.** *Let $j \in \mathbb{S}_m$. Then, if $p' > 1 - 3^{-\Delta}$, there exists a positive constant $c$ such that, for all large enough $z > 0$, we have*

$$\mathbf{Pr}\left[|Z_j| \geqslant z\right] \leqslant \exp(-cz) \quad and \quad \mathbf{Pr}\left[|A_j| \geqslant z\right] \leqslant \exp(-cz).$$

Therefore, for some sufficiently large constant $c_1$, we obtain $\mathbf{Pr}\left[|Z_j| \geqslant c_1 \log m\right] = \mathcal{O}(m^{-3d})$ and $\mathbf{Pr}\left[|A_j| \geqslant c_1 \log m\right] = \mathcal{O}(m^{-3d})$, and using the union bound over the $m^d$ choices for $j$, we conclude that, for all $j \in \mathbb{S}_m$, we have $|Z_j| \leqslant c_1 \log m$ and $|A_j| \leqslant c_1 \log m$ with probability $1 - o(1/m)$.

Now we take an arbitrary path $j_1, j_2, \ldots, j_\ell$ in $L$ such that $j_1 = i_1$, $j_\ell = i_2$ and $\ell \leqslant \|i_1 - i_2\|_1$. For $2 \leqslant k \leqslant \ell - 1$ we consider the set $\widehat{Z}_{j_k}$. Note that, since $\widehat{Z}_j$ separates $Z_j$ from $\mathbb{S}_m \setminus (Z_j \cup \widehat{Z}_j)$, we have that $\bigcup_{k \in [2, \ell-1]} \widehat{Z}_{j_k}$ contains a connected component with at least one vertex from each $\widehat{Z}_{j_k}$, $2 \leqslant k \leqslant \ell - 1$. We call this component the *bridging* component and denote it by $B(i_1, i_2)$. For $i_1$ and $i_2$ we consider the sets $\widehat{A}_{i_1}$ and $\widehat{A}_{i_2}$.

We will show how to find a path from $v_1$ to $v_2$ in $G$ in three parts. We will bound the length of these parts by $F_1$, $F_2$, and $F_3$ so that this path from $v_1$ to $v_2$ in $G$ contains $F_1 + F_2 + F_3$ edges. Note that, since $v_1$ and $v_2$ are such that $\|v_1 - v_2\|_2 = \omega(\log n)$ and $|A_j| \leqslant c_2 \log m = \mathcal{O}(\log n)$ for all $j$, we have that $A_{i_1}$ and $A_{i_2}$ are disjoint. Note that, by definition, $i_1$ is not a neighbor of any vertex that is not in $A_{i_1}$, which gives that, intuitively, $\widehat{A}_{i_1}$ envelops the region $Q_{i_1}$. This property gives that, if there exists a path from $v_1$ to $v_2$ in $G$, this path must cross the region $\bigcup_{k \in \widehat{A}_{i_1}} Q_k$. Now, since $\widehat{A}_{i_1}$ is a set of open vertices, we have that, for each $j \in \widehat{A}_{i_1}$, the cube $Q_j$ has a crossing component. For any connected set $V \subseteq \mathbb{S}_m$ of open vertices, where connectivity is defined with respect to $L$, let $\mathcal{C}(V)$ be the set of vertices of $G$ that belong to the crossing component of at least one $Q_j$ with $j \in V$. With this definition, we have that the path from $v_1$

to $v_2$ must have a node in $\mathcal{C}(\widehat{A}_{i_1})$. Let $F_1$ be the length of the shortest path between $v_1$ and a node of $\mathcal{C}(\widehat{A}_{i_1}) \cap \mathcal{C}(\widehat{Z}_{j_2})$. Note that this node must exist since $\widehat{A}_{i_1} \cap \widehat{Z}_{j_2} \neq \emptyset$ by construction. If we denote by $R$ the set $A_{i_1} \cup \widehat{A}_{i_1}$, we have that this path is completely contained inside $\cup_{k \in R} Q_k$. Therefore, we can bound $F_1$ using the following geometric lemma.

**Lemma 8.** *Let $I$ be a set of vertices of $\mathbb{S}_m$ and $Q = \cup_{i \in I} Q_i$. Let $w_1$ and $w_2$ be two nodes of $G$ inside $Q$. If there exists a path between $w_1$ and $w_2$ entirely contained in $Q$, then there exists a constant $c > 0$ depending only on $d$ such that*

$$d_G(w_1, w_2) \leqslant \frac{c|I|M^d}{r^d}.$$

**Remark 9.** *The result in Lemma 8 also holds when $I$ is replaced by any bounded subset of $\mathbb{R}^d$ composed of the union of parallelograms; in this case, the term $|I|M^d$ gets replaced by the volume of this set.*

Lemma 8 then establishes that there exists a constant $c_2$ such that

$$F_1 \leqslant \frac{c_2|A_{i_1} \cup \widehat{A}_{i_1}|M^d}{r^d} \leqslant \frac{c_2(1+\Delta)(|A_{i_1}|+1)M^d}{r^d} = \mathcal{O}(\log m),$$

since $|\widehat{A}_j| \leqslant \Delta|A_j|+\Delta = \mathcal{O}(\log m)$ for all $j$, and $M/r$ is constant. Similarly, there is a path from $v_2$ to a node inside $\mathcal{C}(\widehat{A}_{i_2}) \cap \mathcal{C}(\widehat{A}_{j_{\ell-1}})$, whose length we denote by $F_2$. An analogous derivation then gives that $F_2 = \mathcal{O}(\log m)$. These paths must intersect $\mathcal{C}(B(i_1, i_2))$ since they intersect $\mathcal{C}(\widehat{A}_{j_2})$ and $\mathcal{C}(\widehat{A}_{j_{\ell-1}})$, respectively. Denote the length of the path in $\mathcal{C}(B(i_1, i_2))$ that connects the two paths we found above by $F_3$. Using Lemma 8 we obtain a constant $c_3$ such that

$$F_3 \leqslant \frac{c_3|B(i_1, i_2)|M^d}{r^d}. \tag{2}$$

In order to bound $|B(i_1, i_2)|$, we use a coupling argument by Fontes and Newman [11] and a result of, Deuschel and Pisztora [6, Lemma 2.3], which gives $\mathbf{Pr}\left[\sum_{k=2}^{\ell-1} |\widehat{Z}_{j_k}| \geqslant \ell\alpha\right] \leqslant \mathbf{Pr}\left[\sum_{k=2}^{\ell-1} |Z_{j_k}| \geqslant (\ell\alpha-1)/\Delta\right] \leqslant \mathbf{Pr}\left[\sum_{k=2}^{\ell-1} |\tilde{Z}_{j_k}| \geqslant (\ell\alpha-1)/\Delta\right]$, where the first inequality follows since $|\widehat{Z}_j| \leqslant 1 + \Delta|Z_j|$ for all $j$, and the $\tilde{Z}$'s are defined to be independent random variables such that $\tilde{Z}_{j_k}$ has the same distribution as $Z_{j_k}$. From Lemma 7 we know that $\tilde{Z}_{j_k}$ is stochastically dominated by an exponential random variable with mean $\mu = \Theta(1)$. Then, applying a Chernoff bound for exponential random variables in the equation above, we obtain a constant $c_4$ such that, for any large enough $\alpha$, we have $\mathbf{Pr}\left[\sum_{k=2}^{\ell-1} |\widehat{Z}_{j_k}| \geqslant \alpha\ell\right] \leqslant \exp\left(-\frac{c_4\alpha\ell}{\mu}\right)$. Since $\ell = \omega(\log n)$ we have that $\mathbf{Pr}\left[\sum_{k=2}^{\ell-1} |\widehat{Z}_{j_k}| \geqslant \alpha\ell\right] = \mathcal{O}(m^{-3d})$ for some large enough $\alpha$. Then, using (2), we have that, with probability $1 - \mathcal{O}(m^{-3d})$,

$$F_3 \leqslant \frac{c_3\alpha\ell M^d}{r^d} = \mathcal{O}(\|i_1 - i_2\|_1) = \mathcal{O}\left(\frac{\|v_1 - v_2\|_2}{M}\right) = \mathcal{O}\left(\frac{\|v_1 - v_2\|_2}{r}\right).$$

Putting everything together, with probability $1 - \mathcal{O}(m^{-3d})$, we obtain a path from $v_1$ to $v_2$ with length at most

$$F_1 + F_2 + F_3 = \mathcal{O}\left(\log n + \frac{\|v_1 - v_2\|_2}{r}\right). \tag{3}$$

By Lemma 6, the result above holds for all connected pairs of nodes $v_1, v_2$ such that $\|v_1 - v_2\|_2 = \omega(\log n)$. Then using $m = \Theta(n^{1/d})$ completes the proof. □

# References

[1] Akyildiz, I., Pompili, D., Melodia, T.: Underwater acoustic sensor networks: research challenges. Ad Hoc Networks 3, 257–279 (2005)

[2] Antal, P., Pisztora, A.: On the chemical distance for supercritical Bernoulli percolation. The Annals of Probability 24, 1036–1048 (1996)

[3] Avin, C., Ercal, G.: On the Cover Time and Mixing Time of Random Geometric Graphs. Theoretical Computer Science 380(1-2), 2–22 (2007)

[4] Bradonjić, M., Elsässer, R., Friedrich, T., Sauerwald, T., Stauffer, A.: Efficient broadcast on random geometric graphs. In: 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2010), pp. 1412–1421 (2010)

[5] Cooper, C., Frieze, A.: The cover time of random geometric graphs. In: 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2009), pp. 48–57 (2009)

[6] Deuschel, J.-D., Pisztora, A.: Surface order large deviations for high-density percolation. Probability Theory and Related Fields 104, 467–482 (1996)

[7] Ellis, R.B., Martin, J.L., Yan, C.: Random geometric graph diameter in the unit ball. Algorithmica 47(4), 421–438 (2007)

[8] Elsässer, R., Sauerwald, T.: Broadcasting vs. Mixing and Information Dissemination on Cayley Graphs. In: Thomas, W., Weil, P. (eds.) STACS 2007. LNCS, vol. 4393, pp. 163–174. Springer, Heidelberg (2007)

[9] Elsässer, R., Sauerwald, T.: On the runtime and robustness of randomized broadcasting. Theoretical Computer Science 410(36), 3414–3427 (2009)

[10] Feige, U., Peleg, D., Raghavan, P., Upfal, E.: Randomized Broadcast in Networks. Random Structures and Algorithms 1(4), 447–460 (1990)

[11] Fontes, L., Newman, C.: First passage percolation for random colorings of $\mathbb{Z}^d$. The Annals of Applied Probability 30(3), 746–762 (1993)

[12] Fountoulakis, N., Panagiotou, K.: Rumor Spreading on Random Regular Graphs and Expanders. In: Serna, M., Shaltiel, R., Jansen, K., Rolim, J. (eds.) APPROX 2010, LNCS, vol. 6302, pp. 560–573. Springer, Heidelberg (2010)

[13] Frieze, A., Grimmett, G.: The shortest-path problem for graphs with random-arclengths. Discrete Applied Mathematics 10, 57–77 (1985)

[14] Frieze, A., Molloy, M.: Broadcasting in random graphs. Discrete Applied Mathematics 54, 77–79 (1994)

[15] Grimmett, G.: Percolation, 2nd edn. Springer, Heidelberg (1999)

[16] Higham, D., Rašajski, M., Pržulj, N.: Fitting a geometric graph to a protein–protein interaction network. Bioinformatics 24(8), 1093 (2008)

[17] Liggett, T., Schonmann, R., Stacey, A.: Domination by product measures. The Annals of Probability 25(1), 71–95 (1997)

[18] Meester, R., Roy, R.: Continuum percolation. Cambridge University Press (1996)

[19] Ou, C.-H., Ssu, K.-F.: Sensor position determination with flying anchors in three-dimensional wireless sensor networks. IEEE Transactions on Mobile Computing 7, 1084–1097 (2008)

[20] Penrose, M., Pisztora, A.: Large deviations for discrete and continuous percolation. Advances in Applied Probability 28, 29–52 (1996)

[21] Penrose, M.D.: The longest edge of the random minimal spanning tree. The Annals of Applied Probability 7(2), 340–361 (1997)

[22] Penrose, M.D.: On $k$-connectivity for a geometric random graph. Random Struct. Algorithms 15(2), 145–164 (1999)

[23] Penrose, M.D.: Random Geometric Graphs. Oxford University Press (2003)

[24] Pittel, B.: On spreading rumor. SIAM Journal on Applied Mathematics 47(1), 213–223 (1987)

[25] Pottie, G.J., Kaiser, W.J.: Wireless integrated network sensors. Commun. ACM 43(5), 51–58 (2000)

[26] Ravelomanana, V.: Extremal properties of three-dimensional sensor networks with applications. IEEE Transactions on Mobile Computing 3, 246–257 (2004)

[27] Sauerwald, T.: On mixing and edge expansion properties in randomized broadcasting. Algorithmica 56, 51–88 (2010)

[28] Tolle, G., Polastre, J., Szewczyk, R., Culler, D., Turner, N., Tu, K., Burgess, S., Dawson, T., Buonadonna, P., Gay, D., et al.: A macroscope in the redwoods. In: 3rd ACM International Conference on Embedded Networked Sensor Systems (SenSys), pp. 51–63 (2005)