The diameter of a long-range percolation graph

Don Coppersmith * David Gamarnik † Maxim Sviridenko ‡

Abstract

We consider the following long-range percolation model: an undirected graph with the node set $\{0,1,\ldots,N\}^d$, has edges $(\mathbf{x},\mathbf{y})$ selected with probability $\approx \beta/||\mathbf{x}-\mathbf{y}||^s$ if $||\mathbf{x}-\mathbf{y}|| > 1$, and with probability 1 if $||\mathbf{x}-\mathbf{y}|| = 1$, for some parameters $\beta, s > 0$. This model was introduced by Benjamini and Berger [2], who obtained bounds on the diameter of this graph for the one-dimensional case $d = 1$ and for various values of s, but left cases $s = 1,2$ open. We show that, with high probability, the diameter of this graph is $\Theta(\log N/\log\log N)$ when $s = d$, and, for some constants $0 < \eta_1 < \eta_2 < 1$, it is at most N^{η_2} when $s = 2d$, and is at least N^{η_1} when $d = 1, s = 2, \beta < 1$ or when $s > 2d$. We also provide a simple proof that the diameter is at most $\log^{O(1)} N$ with high probability, when $d < s < 2d$, established previously in [2].

1 Introduction

Long-range percolation is a model in which any two elements x, y of some (finite or countable) metric space are connected by edges with some probability, inverse proportional to the distance between the points. The motivation for studying this model is dual. First, it naturally extends a classical percolation models on a lattice, by adding edges between non-adjacent nodes with some positive probability. The questions of existence of infinite components were considered specifically by Schulman [8], Aizenman and Newman [1] and Newman and Schulman [7], where the metric space is $\mathcal{Z}$ and edges $(i,j) \in \mathcal{Z}^2$ are selected with probability $\beta/|i-j|^s$ for some parameters β, s. Existence of such an infinite component with positive probability usually implies its existence with probability one, by appealing to Kolmogorov's $0-1$ law. It was shown in [7] and in [1] respectively, that percolation occurs if $s = 2, \beta > 1$ and (suitably defined) short range probability is high enough, and does not occur if $s = 2, \beta \le 1$, for any value of the short range probability.

The second motivation for studying long-range percolation is modelling social networks, initiated by Watts and Strogatz [9]. They considered a random graph model on integer points of a circle, in which neighboring nodes are always connected by an edge, and, in addition, each node is connected to a constant number of other nodes uniformly chosen from a circle. Their motivation was a famous experiment conducted by Milgram [6], which essentially studied the diameter of the "social acquaintances" network and introduced the notion of "six degrees of separation". Watts and Strogatz argued

*IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA. Email address: dcopper@us.ibm.com
†IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA. Email address: gamarnik@watson.ibm.com
‡IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA. Email address: sviri@us.ibm.com

that their graph provides a good model for different types of networks, not only social networks (world wide web, power grids), and showed that the diameter of their random graph is much smaller than the size of the graph. This model was elaborated later by Kleinberg [5], who considered a model similar to a long-range percolation model on a two-dimensional grid, although the work was concerned mostly with algorithmic questions of constructing simple decentralized algorithms for finding short paths between the nodes.

The present paper is motivated by a recent work by Benjamini and Berger [2]. They consider a one-dimensional long-range percolation model in which the nodes are elements of a finite circle $\{0, 1, \ldots, N\}$. An edge (i, j) exists with probability one if $|i-j| = 1$, and with probability $1-\exp(-\beta/|i-j|^s)$ otherwise, for some parameters β, s, here the distance $|\cdot|$ is taken with respect to a circle. Since for large $|i-j|$, $1-\exp(-\beta/|i-j|^s) \approx \beta/|i-j|^s$, this model is closely related to the infinite percolation model on $\mathcal{Z}$, with an important distinction, however. The graph is finite and, since neighboring nodes are connected with probability one, the graph is connected. Thus, the percolation question is irrelevant as such; rather, as in models of "social networks", the diameter of the graph is of interest. It is shown in [2] that the diameter of the circle graph above is, with high probability, a constant, when $s < 1$; is $O(\log^\delta N)$, for some $\delta > 1$, when $1 < s < 2$; and is linear $\Theta(N)$, when $s > 2$. These results apply immediately to a graph on an interval $\{0, 1, \ldots, N\}$. A multidimensional version of this problem with a graph on a node set $\{0, 1, \ldots, N\}^d$ was also considered by Benjamini, et al in [3], who showed that the diameter is $\lceil d/(d-s) \rceil$ when $s < d$. The critical cases $s = 1, 2$ were left open in [2] and the authors conjectured that the diameter is $\Theta(\log N)$ when $s = 1$, and $\Theta(N^\eta)$ for some $0 < \eta < 1$, when $s = 2$. In addition, the authors conjectured that, for the case $1 < s < 2$, $\Theta(\log^\delta N)$ is also a lower bound for some $\delta > 1$. In other words, the system experiences a phase transition at $s = 1$ and $s = 2$. Recently Biskup [4] proved that for the case $1 < s < 2$ the diameter is indeed $\Theta(\log^\delta N)$ for some constant δ which Biskup computes explicitly.

In this work we consider a multidimensional version of the problem. Our graph has a node set $\{0, 1, \ldots, N\}^d$ and edges are selected randomly using a long-range percolation $\beta/||\mathbf{x} - \mathbf{y}||^s$ law. We obtain upper and lower bounds on the diameter for the regimes $s = d, d < s < 2d, s = 2d$ and $s > 2d$. This corresponds to regimes $s = 1, 1 < s < 2, s = 2, s > 2$ for the one-dimensional case. We show that, with high probability, for $s = d$, the diameter of this graph is $\Theta(\log N / \log\log N)$; for $d < s < 2d$ the diameter is at most $\log^\delta N$ for some constant $\delta > 1$; and for $s = 2d$, the diameter is at most N^{η_2}, for some constant $0 < \eta_2 < 1$. We also prove a lower bound $N^{\eta_1}, \eta_1 < 1$ on the diameter, which holds with high probability but only when $d \geq 1, s > 2d$ or $d = 1, s = 2, \beta < 1$. We do not have lower bounds for other cases. Note that our lower bound for $s > 2d$ is weaker than known linear lower bound when $d = 1$. We conjecture that the linear lower bound holds for general dimensions. Our results, when applied to the one-dimensional case, support bounds conjectured in [2] for the case $s = 2$ and disprove it for the case $s = 1$. As we mentioned above, the upper bound $\log^\delta N$ for the case $d < s < 2d$ was proven in [2] for the one-dimensional case. It was pointed to the authors that the proof extends to a multidimensional case as well. We provide here an alternative proof which seems simpler. Summarizing the results of present paper and of [2], the diameter of the long-range percolation graph in one-dimensional case experiences a phase transition at $s = 1, 2$ and has a qualitatively different values for $s < 1$; $s = 1$; $1 < s < 2$; $s = 2$ and $\beta < 1$; $s > 2$. Whether the same holds true for general dimensions (whether $s = d, s = 2d$ are the only critical values) remains to be seen. Our results only partially support this conjecture.

2 Model and the main result

Our model is a random graph $G = G(N)$ on a node set $[N]_d \equiv \{0, 1, \ldots, N\}^d$ - integral points of the d-dimensional cube with side length N. Let $||\mathbf{x}||$ denote an $\mathbf{L_1}$ norm in the space $\mathcal{Z}^d$. That is $||\mathbf{x}|| = \sum_{i=1}^d |x_i|$. Nodes $\mathbf{x}, \mathbf{y} \in [N]_d$ are connected with probability 1 if $||\mathbf{x} - \mathbf{y}|| = 1$, and, otherwise, with probability $1 - \exp(-\frac{\beta}{||\mathbf{x}-\mathbf{y}||^s})$, where $\beta > 0, s > 0$ are some fixed parameters. Let $D(N)$ denote the (random) diameter of the graph $G(N)$, and let $P(N)$ denote the (random) length of a shortest path between nodes $\mathbf{0} \equiv (0, \ldots, 0)$ and $\mathbf{N} = (N, \ldots, N)$. For any $\mathbf{x}, \mathbf{y} \in [N]_d$ let also $P(\mathbf{x}, \mathbf{y})$ denote the length of a shortest path between nodes $\mathbf{x}, \mathbf{y}$ in the graph $G(N)$. Our main result is as follows.

Theorem 2.1 *There exist constants $C_1, C_2, C_s > 0, \delta > 1,\ 0 < \eta_1 < \eta_2 < 1$, which in general depend on s, β and on dimension d such that*

1. $\lim_{N\to\infty} \text{Prob}\{D(N) \geq N^{\psi}\} = 1$, *for any* $s > 2d, \psi < \frac{s-2d}{s-d-1}$.
2. $\lim_{N\to\infty} \text{Prob}\{D(N) \leq N^{\eta_2}\} = 1$, *for* $s = 2d$ *and*
   $\lim_{N\to\infty} \text{Prob}\{D(N) \geq N^{\eta_1}\} = 1$, *for* $d = 1, s = 2, \beta < 1$.
3. $\lim_{N\to\infty} \text{Prob}\{C_s \log N \leq D(N) \leq \log^{\delta} N\} = 1$, *for* $d < s < 2d$.
4. $\lim_{N\to\infty} \text{Prob}\{\frac{C_1 \log N}{\log\log N} \leq D(N) \leq \frac{C_2 \log N}{\log\log N}\} = 1$, *for* $s = d$.

As we mentioned above, it was shown in [3] that the diameter is, with high probability, $\lceil d/(d-s) \rceil$, when $s < d$. Also part 3 of the theorem above was proven by Benjamini and Berger in [2] for the one-dimensional case. They also pointed out to the authors that their proof holds for a multidimensional case as well. We provide here a simpler proof. Throughout the paper we use standard notations $f = O(g), f = \Omega(g), f = \Theta(g), f = o(g)$, which mean respectively that for two functions $f(N), g(N)$, $f(N) \leq C_1 g(N), f(N) \geq C_2 g(N), C_3 g(N) \leq f(N) \leq C_4 g(N), f(N)/g(N) \to 0$, for some constants $C_i, i = 1, 2, 3, 4$ which in general depend on β, s, but do not depend on N. Also, throughout the paper $[n]_d$ denotes an integral cube $\{0, 1, \ldots, n\}^d$ for any nonnegative integer n. The logarithmic function is always assumed to be with the base e.

3 Case $s > 2d$. Lower bound

In this section we show that,with high probability, the diameter of the graph $G(N)$ is at least essentially $N^{\frac{s-2d}{s-d-1}}$. As we noted, for the one dimensional case $d = 1$ this is weaker than the existing linear lower bound $\Omega(N)$ ([2]).

Proof of Theorem 2.1, Part 1: We fix a constant $\psi < \frac{s-2d}{s-d-1}$. For any $k > N^{1-\psi}$ let $L(k)$ be the total number of edges between pairs of points at distance exactly k. We will now show that if $\psi < (s-2d)/(s-d-1)$ then $\sum_{k>N^{1-\psi}} kL(k) \le dN/2$, with high probability. Since $||\mathbf{N}|| = dN$, then this would imply that, with high probability, any path between $\mathbf{0}$ and $\mathbf{N}$ would contain at least $dN/(2N^{1-\psi}) = (dN^{\psi})/2$ edges and the proof would be completed. For a fixed pair of nodes $\mathbf{x}, \mathbf{y}$ at a distance k, the probability that the edge between them exist is $1 - \exp(-\beta/k^s) \le \beta/k^s$, where we use $\exp(-\beta x) \ge 1 - \beta x$ for all $0 \le x \le 1$. For a fixed node $\mathbf{x}$ there are $\Theta(k^{d-1})$ nodes $\mathbf{y}$ which are at distance k from $\mathbf{x}$; also there are N^d choices for the node $\mathbf{x}$. Combining $E[L(k)] = O(N^d k^{d-1}(\beta/k^s))$. Then

$$\sum_{k>N^{1-\psi}} kE[L(k)] = O(\beta N^d \sum_{k>N^{1-\psi}} k^{d-s}) = O(N^d N^{(1-\psi)(d-s+1)}).$$

For the given choice of ψ, we have $d + (1-\psi)(d-s+1) < 1$ and the value above is $o(N)$. Using Markov's inequality, we obtain

$$\text{Prob}\{\sum_{k>N^{1-\psi}} kL(k) > N/2\} \le \frac{o(N)}{(N/2)} = o(1).$$

□

4 Case $s = 2d$.

4.1 Upper bound

In this subsection we prove that when $s = 2d$, there exists a constant $0 < \eta < 1$, which depends on β and d, such that with high probability $D(N) \le N^{\eta}$. To this end we first establish an upper bound on $\max_{\mathbf{x},\mathbf{y} \in [N]_d} E[P(\mathbf{x}, \mathbf{y})]$ and then use this bound to obtain a polynomially small bound on $\text{Prob}\{D(N) > N^{\eta}\}$ for some constant $\eta < 1$.

Proof of Theorem 2.1, Part 2:

We first assume that N is a power of $3 : N = 3^m$, for some integer $m > 0$, and then consider the general case. For any fixed integer n let $R(n) = \max_{\mathbf{x},\mathbf{y} \in [n]_d} E[P(\mathbf{x}, \mathbf{y})]$. That is, $R(n)$ is the maximum over expected lengths of shortest paths between all the pairs of points in the cube $[n]_d$. We obtain an upper bound on $R(N)$ by relating it to $R(N/3)$. Divide the cube $[N]_d$ into 3^d subcubes of the type $I_{i_1 \ldots i_d} \equiv \prod_{j=1}^d [i_j \frac{N}{3}, (i_j+1)\frac{N}{3}], 0 \le i_j \le 2$. Each cube has a side length $N/3$ (which is integer since N is a power of three). We say that two such cubes are neighboring if they have at least a common node. For example $[0, N/3]^d$ and $[N/3, 2N/3]^d$ are neighboring through a corner node $(N/3, \ldots, N/3)$. We now fix a pair of points $\mathbf{x}, \mathbf{y} \in [N]_d$ and estimate $P(\mathbf{x}, \mathbf{y})$ by considering two cases.

1. $\mathbf{x}, \mathbf{y}$ belong to the same subcube $I = I_{i_1 \ldots i_d}$. The length of a shortest path between these two points using edges of $[N]_d$ is not bigger than the length of the shortest path between same points but using only edges of the subcube I. Therefore $E[P(\mathbf{x}, \mathbf{y})] \le R(N/3)$.

2. $\mathbf{x}, \mathbf{y}$ belong to different subcubes I, I'. Let $\mathcal{E} = \mathcal{E}(I, I')$ be the event "there exists at least one edge between some nodes $\mathbf{v} \in I, \mathbf{v}' \in I'$". The probability that $\mathcal{E}$ occurs is at least $1 - \exp(-\beta(\frac{N}{3} + 1)^{2d}/(dN)^{2d})$ since there are $(\frac{N}{3}+1)^d$ nodes in each cube, and the largest possible distance between them is dN. In particular, $\text{Prob}\{\mathcal{E}\}$ is not smaller than a certain constant $\delta > 0$, independent of N. We now estimate $E[P(\mathbf{x}, \mathbf{y})]$ conditioned on $\mathcal{E}$ and $\bar{\mathcal{E}}$. Given that $\mathcal{E}$ occurs, select an edge $(\mathbf{v}, \mathbf{v}')$ between the cubes I, I'. Then

$$E[P(\mathbf{x}, \mathbf{y})|\mathcal{E}] \leq E[P(\mathbf{x}, \mathbf{v})|\mathcal{E}] + E[P(\mathbf{v}', \mathbf{y})|\mathcal{E}] + 1$$

Note, however, that edges within each cube I, I' are selected independently from edges between cubes and specifically are independent from the event $\mathcal{E}$. Therefore, since $\mathbf{x}, \mathbf{v}$ belong to the same cube, $E[P(\mathbf{x}, \mathbf{v})|\mathcal{E}] \leq R(N/3)$. Similarly, $E[P(\mathbf{v}', \mathbf{y})|\mathcal{E}] \leq R(N/3)$. We conclude $E[P(\mathbf{x}, \mathbf{y})|\mathcal{E}] \leq 2R(N/3) + 1$. Now, suppose $\mathcal{E}$ does not occur. Select a cube I'' which is a neighboring cube for cubes I, I' (it is easy to see that such a cube exists). Specifically, let $\mathbf{z}(\mathbf{z}')$ be the nodes shared by cubes I and I'' (I' and I''). Then arguing as above $E[P(\mathbf{x}, \mathbf{y})|\bar{\mathcal{E}}] \leq E[P(\mathbf{x}, \mathbf{z})|\bar{\mathcal{E}}] + E[P(\mathbf{z}, \mathbf{z}')|\bar{\mathcal{E}}] + E[P(\mathbf{z}', \mathbf{y})|\bar{\mathcal{E}}] \leq 3R(N/3)$. Combining, we obtain

$$E[P(\mathbf{x}, \mathbf{y})] \leq (2R(N/3)+1)\text{Prob}\{\mathcal{E}\}+3R(N/3)(1-\text{Prob}\{\mathcal{E}\}) = (3-\text{Prob}\{\mathcal{E}\})R(N/3)+\text{Prob}\{\mathcal{E}\} \leq$$

$$(3 - \delta)R(N/3) + 1.$$

We conclude, $R(N) = \max_{\mathbf{x}, \mathbf{y} \in [N]_d} E[P(\mathbf{x}, \mathbf{y})] \leq (3 - \delta)R(N/3) + 1$. Applying this bound $m - 1 = \log N / \log 3 - 1$ times, we obtain

$$R(N) \leq (3 - \delta)^{m-1}R(3) + \sum_{i=0}^{m-2} (3 - \delta)^i = O((3 - \delta)^m) = O(N^{\frac{\log(3-\delta)}{\log 3}}),$$

Note, $\alpha \equiv \log(3 - \delta)/\log 3 < 1$. We obtain $R(N) = O(N^\alpha)$ for some $\alpha < 1$.

In order to generalize the bound for all N, it is tempting to argue that $R(N) \leq R(3^m)$ as long as $N \leq 3^m$. This would require proving a seemingly obvious statement that $R(n)$ is a non-decreasing function of n. While this is most likely correct, proving it does not seem to be trivial. Instead, we proceed as follows. Let m be such that $3^m \leq N < 3^{m+1}$. We cover the cube $[N]_d$ with 3^d cubes $I_i, i = 1, \ldots, 3^d$ with side length 3^m, with a possible overlapping. Specifically, $I_i \subset [N]_d$ and $\cup_i I_i = [N]_d$. Let $\mathbf{x}, \mathbf{y} \in [N]_d$ be arbitrary. Find cubes $I_{i_1}, I_{i_2}, I_{i_3}$ such that $\mathbf{x} \in I_{i_1}, \mathbf{y} \in I_{i_3}$ and $I_{i_1} \cap I_{i_2} \neq \emptyset, I_{i_2} \cap I_{i_3} \neq \emptyset$. Let $\mathbf{z}_1, \mathbf{z}_2$ be some nodes lying in these intersections. Then $E[P(\mathbf{x}, \mathbf{y})] \leq E[P(\mathbf{x}, \mathbf{z}_1)] + E[P(\mathbf{z}_1, \mathbf{z}_2)] + E[P(\mathbf{z}_2, \mathbf{y})] = O((3^m)^\alpha)$, where the last equality follows since pairs $(\mathbf{x}, \mathbf{z}_1), (\mathbf{z}_1, \mathbf{z}_2), (\mathbf{z}_2, \mathbf{y})$ lie within cubes $I_{i_1}, I_{i_2}, I_{i_3}$ respectively and each of them has a side length 3^m. But $3^m \leq N$. We conclude $E[P(\mathbf{x}, \mathbf{y})] = O(N^\alpha)$ and $R(N) = \max_{\mathbf{x}, \mathbf{y}} E[P(\mathbf{x}, \mathbf{y})] = O(N^\alpha)$.

We now finish the proof of part 2, upper bound, by obtaining a similar bound on the diameter $D(N)$. Fix an arbitrary $0 < \epsilon, \gamma < 1$ such that $\alpha + \epsilon < 1$ and $\epsilon - d(1 - \gamma) > 0$. Divide the cube $[N]_d$ into equal subcubes $I_{i_1 \ldots i_d} = \prod_{j=1}^d [i_j N^\gamma, (i_j + 1)N^\gamma],\ 0 \leq i_j \leq N^{1-\gamma}$, each with side length N^γ. The total number of subcubes is $N^{d(1-\gamma)}$. Fix any such cube I and let $\mathbf{x}(I)$ be its lower corner (the node with smallest possible coordinates). We showed above $E[P(\mathbf{0},\mathbf{x}(I))] \le O(N^\alpha)$, from which, using Markov inequality,

$$\text{Prob}\{P(\mathbf{0},\mathbf{x}(I)) \ge N^{\alpha+\epsilon}\} = O(\frac{N^\alpha}{N^{\alpha+\epsilon}}) = O(\frac{1}{N^\epsilon}).$$

Then

$$\text{Prob}\{\max_I P(\mathbf{0},\mathbf{x}(I)) \ge N^{\alpha+\epsilon}\} = O(\frac{N^{d(1-\gamma)}}{N^\epsilon}) = O(\frac{1}{N^{\epsilon-d(1-\gamma)}}).$$

On the other hand for every cube I and every $\mathbf{x} \in I$ we have trivially, $P(\mathbf{x},\mathbf{x}(I)) \le dN^\gamma$. Since $D(N) \le 2\sup_{\mathbf{x}\in[N]_d} P(0,\mathbf{x})$, then

$$\text{Prob}\{D(N) \ge 2(dN^\gamma + N^{\alpha+\epsilon})\} = O(\frac{1}{N^{\epsilon-d(1-\gamma)}}) = o(1).$$

We take $\eta = \max\{\gamma, \alpha+\epsilon\} < 1$ and obtain $\text{Prob}\{D(N) \ge 4dN^\eta\} = o(1)$. This completes the proof of the upper bound. □

4.2 Lower bound

The proof of the lower bound for the one-dimensional case $d = 1, s = 2, \beta < 1$ is similar to the proof for the case $s > 2$, from [2] and uses the notion of a cut point. We first show that $E[D(N)] \ge N^\eta$ for a certain constant $0 < \eta < 1$, for large N. Then we show that this bound holds with high probability. Given a node $1 \le i \le N-1$, we call it a cut node if there are no edges which go across i. Namely, i is a cut point if edges (j,k) do not exist for all $j < i < k$. The probability that i is a cut node is $\exp(-\beta\sum_{j<i<k}\frac{1}{|j-k|^2}) \ge \exp(-\beta\sum_{1\le n\le N}\frac{n-1}{n^2}) = \Theta(\frac{1}{N^\beta})$. Then the expected number of cuts is $\Omega(N^{1-\beta})$ (which will be helpful to us only if $\beta < 1$). But the shortest path $P(N)$ and as a result the diameter $D(N)$ are not smaller than the number of cuts. Taking $\eta < 1-\beta$, we obtain the bound $E[D(N)] \ge N^\eta$ for large N.

We now complete the proof, by showing that the lower bound holds with high probability. Divide the interval $[N]$ into $N^{\frac{2}{3}}$ intervals $I_1, I_2, \ldots, I_{N^{\frac{2}{3}}}$ each of length $N^{\frac{1}{3}}$. For each interval I_i and each $x \in I_i$, we say that x is a local cut point if it is a cut point with respect to just the graph induced by vertices from I_i. We showed above that the expected number of local cut points in the interval I_i is at least $|I_i|^\eta = N^{\frac{\eta}{3}}$, for any $\eta < 1-\beta$ and for all i. Let $C(I_i)$ be the number of local cut points in the interval I_i. We now show that, with high probability, at least one of the intervals has at least $(1/2)N^{\frac{\eta}{3}}$ local cut points. Note $\{C(I_i)\}_{1\le i\le N^{\frac{2}{3}}}$ are independent from each other. We have $E[C(I_i)] \ge N^{\frac{\eta}{3}}$. Also $\text{Var}(C(I_i)) \le |I_i|^2 = N^{\frac{2}{3}}$. Applying Chebyshev's inequality, we have

$$\text{Prob}\{\sum_i C(I_i)/N^{\frac{2}{3}} < \frac{1}{2}N^{\frac{\eta}{3}}\} \le \frac{\text{Var}(\text{I}_\text{i})}{\frac{1}{2}N^{\frac{\eta}{3}}N^{\frac{2}{3}}} = O(\frac{1}{N^{\frac{\eta}{3}}}),$$

Therefore, with high probability, at least one of the intervals contains at least $(1/2)N^{\frac{\eta}{3}}$ local cut points. We denote this interval by I_{i^*}. Let us estimate the number of edges between I_{i^*} and $[N]\setminus I_{i^*}$. Note that in defining interval I_{i^*} with many local cut points, we only considered edges within intervals I_i.

Note also, that for each $k \geq 1$ there are at most $2k$ edges of length k between I_{i^*} and its complement. Then, the expected number of edges between I_{i^*} and $[N] \setminus I_{i^*}$ is at most

$$\sum_{k=1}^{N} 2k(1 - \exp(-\frac{\beta}{k^2})) + O(1) = O(\sum_{k=1}^{N} \frac{\beta}{k}) = O(\log N),$$

where we use $\exp(-\beta x) \geq 1 - \beta x$ for all $x \in [0, 1]$. Using Markov's inequality, the probability that the number of edges between I_{i^*} and its complement is bigger than $\log^2 N$ is at most $O(1/\log N)$. We conclude that with high probability there are at most $\log^2 N$ edges between I_{i^*} and its complement. Since the number of local cuts in I_{i^*} is $\Omega(N^{\frac{\eta}{3}})$ then there are two local cuts i_1, i_2, such that the interval $[i_1, i_2]$ contains at least $\Omega(N^{\frac{\eta}{3}}/\log^2 N) = \Omega(N^{\frac{\eta}{4}})$ local cuts and no outside edges are connected to nodes in interval $[i_1, i_2]$. Let the number of local cuts in $[i_1, i_2]$ be L. We take the $(1/3)L$-th and the $(2/3)L$-th local cut in this interval. By construction, the shortest path between these local cuts is at least $(1/3)L = \Omega(N^{\frac{\eta}{4}})$. We conclude, $D(N) = \Omega(N^{\frac{\eta}{4}})$, with high probability. □

5 Case $d < s < 2d$.

The lower bound $D(N) \geq C_s \log N$ was proven to hold with high probability in [2] for the case $d = 1$, using branching theory and the fact that for each node, the expected number of its neighbors is a constant. The proof extends easily to all dimensions d. We now focus on an upper bound. Our proof is similar to the one in [2] and is based on renormalization technique, although our analysis is simpler.

Proof of Theorem 2.1, Part 3: We have $d < s < 2d$. Let us fix $\alpha < 1$ such that $2d\alpha > s$. Split the cube $[N]_d$ into equal subcubes $I_{i_1 \ldots i_d} \equiv \prod_{j=1}^{d} [i_j \lceil N^\alpha \rceil, (i_j + 1)\lceil N^\alpha \rceil - 1]$ with side length $\lceil N^\alpha \rceil$. If $N/\lceil N^\alpha \rceil$ is not an integer then we make the cubes containing nodes $(\ldots, N, \ldots)$ overlap partially with some other cubes. In the following we drop the rounding $\lceil \cdot \rceil$ for simplicity, the argument still holds. Consider the following event $\mathcal{E}_1$: "there exist two cubes I, I' such that no edge exists between points $\mathbf{x} \in I$ and $\mathbf{y} \in I'$". Each resulting cube $I = I_{i_1 \ldots i_d}$ we split further into subcubes with side length N^{α^2}. We consider the event $\mathcal{E}_2$: "there exist a cube I with side length N^α and its two subcubes I_1, I_2 with side length N^{α^2}, such that no edge exists between points in I_1 and I_2". We continue this process m times, obtaining in the end cubes with side length N^{α^m}. Assume that none of the events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m$ occurs. We claim that then the diameter of our original graph is at most $2^{m+1} N^{\alpha^m}$. In fact, since event $\mathcal{E}_1$ does not occur any two points $\mathbf{x}, \mathbf{y} \in [N]_d$ are connected by a path with length at most $2\bar{D}(N^\alpha) + 1$, where $\bar{D}(N^\alpha)$ is the (random) largest diameter of the cubes $I_{i_1 \ldots i_d}$ with side length N^α. Similarly, since event $\mathcal{E}_2$ does not occur, $\bar{D}(N^\alpha) \leq 2\bar{D}(N^{\alpha^2}) + 1$, where $\bar{D}(N^{\alpha^2})$ is the largest diameter of the subcubes with side length N^{α^2}, obtained in second stage. In the end we obtain that the diameter of our graph satisfies $D(N) \leq 2^m D(N^{\alpha^m}) + 2^m \leq 2^{m+1} d N^{\alpha^m}$, since trivially, $D(N^{\alpha^m}) \leq d N^{\alpha^m}$. We now show that for a certain value of m, which depends on N, this upper bound on the diameter $D(N)$ is at most $\log^\delta N$ for some constant $\delta > 1$ and simultaneously, the probability $\text{Prob}\{\wedge_{r=1}^{m} \bar{\mathcal{E}}_r\} \to 1$, as $N \to \infty$. For a given cube with side length $N^{\alpha^{r-1}}$ and its two given subcubes with side length N^{α^r}, the probability that no edges exist between these two subcubes is at most $\exp(-\beta N^{2d\alpha^r}/(dN)^{s\alpha^{r-1}}) = \exp(-\Theta(N^{\alpha^{r-1}(2d\alpha - s)}))$, since there are $N^{2d\alpha^r}$ pairs of points considered and the largest distance among any two of them is $dN^{\alpha^{r-1}}$. Since there are at most N^{2d} pairs of such subcubes, then the probability of the event $\mathcal{E}_r$ is bounded above by $N^{2d}\exp(-\Theta(N^{\alpha^{r-1}(2d\alpha-s)}))$. We conclude

$$\text{Prob}\{\vee_{r=1}^m \mathcal{E}_r\} \le \sum_{r=1}^m N^{2d}e^{-\Theta(N^{\alpha^{r-1}(2d\alpha-s)})} \le mN^{2d}e^{-\Theta(N^{\alpha^m(2d\alpha-s)})}.$$

Let us fix a large constant C and take

$$m = \frac{\log\log N - \log\log\log N + \log(2d\alpha - s) - \log C}{\log\frac{1}{\alpha}} = O(\log\log N).$$

A straightforward computation shows that for this value of m,

$$\text{Prob}\{\vee_{r=1}^m \bar{\mathcal{E}}_r\} = O(e^{-\Theta(\log^C N)}). \tag{1}$$

On the other hand, we showed above that, conditioned on event $\wedge_r \bar{\mathcal{E}}_r$, we have $D(N) = O(2^m dN^{\alpha^m})$. For our choice of m a simple calculation shows that $\alpha^m \log N = O(\log\log N)$ or $N^{\alpha^m} = \log^{O(1)} N$. Also, since $m = O(\log\log N)$, then $2^m = O(\log^{O(1)} N)$. This completes the proof. □

In the course of the proof we established the following bound which follows immediately from (1).

Corollary 5.1 *For any constant C, there exists a constant $\delta > 1$ such that*

$$\text{Prob}\{D(N) > \log^\delta N\} \le O(e^{-\Theta(\log^C N)}).$$

6 Case $s = d$.

Proof of Theorem 2.1, Part 4: We first prove a lower bound. We show that $D(N) \ge (d-\epsilon)\log N/\log\log N$ with high probability, for any constant $0 < \epsilon < 1$. Observe, that, for any $1 < k \le N$ and for each node $\mathbf{x} \in [N]_d$, there are $\Theta(k^{d-1})$ nodes at distance k from $\mathbf{x}$. Each such node is connected to $\mathbf{x}$ with probability $1 - \exp(-\beta/k^d) \le \beta/k^d$. (We used $\exp(-\beta x) \ge 1 - \beta x$ for all $x \in [0, 1]$). Then the expected number of nodes connected to $\mathbf{x}$ by an edge is at most $O(1) + O(\sum_{1\le k\le dN}(k^{d-1}/k^d)) = O(\log N)$. Then, the total expected number of nodes which are reachable from $\mathbf{x}$ by paths with length $\le m$ is at most $c^m \log^m N$, for some constant c. We denote the number of such nodes $B(m)$. Using Markov's inequality

$$\text{Prob}\{B(m) \ge N^d\} \le \frac{E[B(m)]}{N^d} \le \frac{c^m \log^m N}{N^d} \to 0$$

if $m = (d - \epsilon)\log N/\log\log N$. Therefore, with probability tending to one, the diameter $D(N)$ is $\Omega(\log N/\log\log N)$.

We now focus on a more difficult part – the upper bound. The proof is fairly technical, but is based on a simple observation which we present now. We have already noted that any fixed node $\mathbf{z}$, in particular, node $\mathbf{N} = (N, N, \ldots, N)$, has in expectation $\Theta(\log N)$ neighbors. We will show later in the formal proof that this actually holds with high probability. Consider a subcube $I = [0, N/\log^c N]^d$ for a certain constant c. Let $\mathbf{y}$ be a neighbor of $\mathbf{x}$. The probability that $\mathbf{y}$ has no neighbors in I is at most $\exp(-\beta N^d/(d^d N^d \log^{cd} N))$, since the largest possible distance is dN and the number of nodes in I is $N^d/\log^{cd} N$. Then probability that none of the $\Theta(\log N)$ neighbors of $\mathbf{N}$ is connected to some node of I by a path of length $\leq$ two is at most $\exp(-\beta N^d \log N/(d^d N^d \log^{cd} N)) = \exp(-\Theta(\log^{1-cd} N))$. If $c < 1/d$ then this quantity converges to zero. Therefore, with high probability $\mathbf{N}$ is connected to some node $\mathbf{X}_1 \in I$ by a path of length 2. Applying this argument for $\mathbf{X}_1$ we find a node $\mathbf{X}_2$ which is connected to $\mathbf{X}_1$ by a path of length two and such that all the coordinates of $\mathbf{X}_2$ are at most $N/\log^{2c} N$. Continuing m times we will obtain that $\mathbf{N}$ is connected by a path of length $O(m)$ to some node $\mathbf{X}_m$ with all the coordinates $\leq N/\log^{cm} N$. Taking $m = O(\log N/\log\log N)$ we will obtain that, with high probability, $\mathbf{N}$ is connected to $\mathbf{0}$ by a path of length $\leq O(m)$. We now formalize this intuitive argument.

We fix an arbitrary node $\mathbf{z}_0 \in [N]_d$. Consider all the paths $(\mathbf{x}, \mathbf{y}, \mathbf{z}_0)$ with length two, which end in node $\mathbf{z}_0$. That is edges $(\mathbf{x}, \mathbf{y}), (\mathbf{y}, \mathbf{z}_0)$ exist. Let $\mathbf{X}_1 = \text{argmin}||\mathbf{x}||$, where the minimum is taken over all such paths. In other words, $\mathbf{X}_1$ is the smallest, in norm, node connected to z_0 via a path of length at most 2. Note, $\mathbf{X}_1$ is random and $||\mathbf{X}_1|| \leq ||\mathbf{z}_0||$, as $\mathbf{z}_0$ is connected to itself by a path of length two. Similarly, let $\mathbf{X}_2 < \mathbf{X}_1$ be the smallest, in norm, node, connected to $\mathbf{X}_1$ via a path of length 2. We continue this procedure for m (to be defined later) steps and obtain a (random) node $\mathbf{X}_m$.

Lemma 6.1 *For any constantly large integer c, if $m = (2d+2) \cdot 2^{c+1} \log N/\log\log N$, then the bound $||\mathbf{X}_m|| \leq \exp((\log N)^{d/2^c})$ holds with probability at least $1 - 1/N^{2d}$.*

Before we prove the lemma, let us show how it is used to prove the result. We invoke part 3 of Theorem 2.1, which we proved in the previous section. Choose a constant integer c such that $2^c/d \geq 2\delta$, where $\delta > 1$ is a constant from part 3 of Theorem 2.1. Applying part 3 of Theorem 2.1, the diameter of the cube $[\exp((\log N)^{d/2^c})]_d$ is at most $((\log N)^{d/2^c})^\delta \leq \log^{\frac{1}{2}} N = o(\log N/\log\log N)$ with high probability. In particular $\sup_{\mathbf{x}:||\mathbf{x}|| \leq \exp((\log N)^{d/2^c})} P(0, \mathbf{x}) = o(\log N/\log\log N)$ with high probability. By the conclusion of the lemma, with probability at least $1 - O(1/N^{2d})$, each fixed node $z_0 \in [N]_d$ is connected to some node $\mathbf{X}_m$ with $||\mathbf{X}_m|| \leq \exp((\log N)^{d/2^c})$ by a path of length $m = O(\log N/\log\log N)$. Then, with probability at least $1 - O(1/N^d)$, all the nodes $z_0 \in [N]_d$ are connected to some corresponding nodes $\mathbf{X}_m \in [\exp((\log N)^{d/2^c})]_d$ by a path of length $O(\log N/\log\log N)$. Combining, we obtain that $\sup_{\mathbf{z}_0 \in [N]_d} P(0, \mathbf{z}_0) = O(\log N/\log\log N)$ with probability at least $1 - o(1)$. But $D(N) \leq 2\sup_{\mathbf{z}_0 \in [N]_d} P(0, \mathbf{z}_0)$. □

Proof of Lemma 6.1: We fix a node $\mathbf{x}$ with $||\mathbf{x}|| \leq ||\mathbf{z}_0||$, fix $1 \leq r \leq m$ and consider $\mathbf{X}_r$ conditioned on event $\mathbf{X}_{r-1} = \mathbf{x}$ (assume $\mathbf{X}_0 = \mathbf{z}_0$). Our goal for the remaining part is the following

Lemma 6.2 *If $||\mathbf{x}|| > \exp((\log N)^{\frac{d}{2^c}})$, then*

$$E\Big[||\mathbf{X}_r||\Big|\mathbf{X}_{r-1} = \mathbf{x}\Big] \leq O(\frac{||\mathbf{x}||}{(\log N)^{1/2^{c+1}}}). \tag{2}$$

In other words, at each step $r = 1, 2, \ldots, m$, the expected value of $||\mathbf{X}_r||$ decreases by a factor of $O(\frac{1}{(\log N)^{1/2^{c+1}}})$, provided that $||\mathbf{X}_{r-1}||$ is still bigger than $\exp((\log N)^{\frac{d}{2^c}})$.

Proof. Let $B(\mathbf{x})$ be the total number of nodes which are connected to $\mathbf{X}_{r-1} = \mathbf{x}$ and which have a norm smaller than $||\mathbf{x}||$. Note, that for each such node $\mathbf{y}$, $||\mathbf{y} - \mathbf{x}|| \le ||\mathbf{y}|| + ||\mathbf{x}|| < 2||\mathbf{x}||$. We first show that with probability at least $1 - O(\frac{1}{(\log N)^{d/2^c}})$, the equality $B(\mathbf{x}) = \Omega(\log ||\mathbf{x}||)$ holds. For any fixed $k \le ||\mathbf{x}||$ there are $\Theta(k^{d-1})$ nodes $\mathbf{y}$ which for which $||\mathbf{y} - \mathbf{x}|| = k$ and $||\mathbf{y}|| < ||\mathbf{x}||$. Each such node is connected by an edge to $\mathbf{x}$ with probability $1 - \exp(-\beta/k^d)$. Then

$$E[B(\mathbf{x})] = \sum_{0 \le k \le ||\mathbf{x}||} (1 - \exp(-\frac{\beta\Theta(k^{d-1})}{k^d})) = \Theta(\log ||\mathbf{x}||).$$

Let $c_1 < c_2$ be constants, such that $c_1 \log ||\mathbf{x}|| \le E[B(\mathbf{x})] \le c_2 \log ||\mathbf{x}||$. We now estimate the second moment

$$E[B^2(\mathbf{x})] = E[B(\mathbf{x})] + \sum_{\mathbf{y}_1 \ne \mathbf{y}_2, ||\mathbf{y}_1||, ||\mathbf{y}_2|| < ||\mathbf{x}||} (1 - \exp(-\frac{\beta}{||\mathbf{y}_1 - \mathbf{x}||}))(1 - \exp(-\frac{\beta}{||\mathbf{y}_2 - \mathbf{x}||})) \le$$

$$E[B(\mathbf{x})] + \sum_{||\mathbf{y}_1||, ||\mathbf{y}_2|| < ||\mathbf{x}||} (1 - \exp(-\frac{\beta}{||\mathbf{y}_1 - \mathbf{x}||}))(1 - \exp(-\frac{\beta}{||\mathbf{y}_2 - \mathbf{x}||})) =$$

$$E[B(\mathbf{x})] + (E[B(\mathbf{x})])^2.$$

It follows, $\mathrm{Var}(B(\mathbf{x})) \le E[B(\mathbf{x})]$. Using Chebyshev's inequality,

$$\mathrm{Prob}\{B(\mathbf{x}) \le (1/2)c_1 \log ||\mathbf{x}||\} \le \mathrm{Prob}\{|B(\mathbf{x}) - E[B(\mathbf{x})]| \ge (1/2)c_1 \log ||\mathbf{x}||\} \le$$

$$\frac{\mathrm{Var}(B(\mathbf{x}))}{(1/4)c_1^2 \log^2 ||\mathbf{x}||} \le \frac{c_2 \log ||\mathbf{x}||}{(1/4)c_1^2 \log^2 ||\mathbf{x}||} = O(\frac{1}{\log ||\mathbf{x}||}) \le O(\frac{1}{(\log N)^{d/2^c}}), \tag{3}$$

where the last inequality follows from the assumption $||\mathbf{x}|| > \exp((\log N)^{d/2^c})$ of the lemma. Let

$$V(\mathbf{x}) = \{\mathbf{z} : ||\mathbf{z}|| \le \frac{||\mathbf{x}||}{(\log N)^{1/2^{c+1}}}\}.$$

In particular, $|V(\mathbf{x})| = \Theta(||\mathbf{x}||^d/(\log N)^{d/2^{c+1}})$. Suppose $\mathbf{y}, ||\mathbf{y}|| < ||\mathbf{x}||$ is any node connected by an edge to $\mathbf{x}$ (if any exist). Note that the distance between $\mathbf{y}$ and any node in $V(\mathbf{x})$ is smaller than $3||\mathbf{x}||$. Then, the probability that $\mathbf{y}$ has no nodes in $V(\mathbf{x})$ connected to it by an edge is at most

$$\exp(-\frac{\beta\Theta(||\mathbf{x}||^d)}{(\log N)^{d/2^{c+1}}||\mathbf{x}||^d}) = \exp(-\frac{\Theta(1)}{(\log N)^{d/2^{c+1}}}).$$

By (3), with probability at least $1 - O(\frac{d}{(\log N)^{1/2^c}})$, $\mathbf{x}$ has $\Omega(\log ||\mathbf{x}||)$ nodes $\mathbf{y}, ||\mathbf{y}|| < ||\mathbf{x}||$ connected to it. Conditioned on this event, the probability that no node in $V(\mathbf{x})$ is connected to $\mathbf{x}$ by a path of length two is at most $\exp(-\frac{\Omega(\log ||\mathbf{x}||)}{(\log N)^{d/2^{c+1}}})$. By assumption, $||\mathbf{x}|| > \exp((\log N)^{\frac{d}{2^c}})$ or $\log ||\mathbf{x}|| > (\log N)^{\frac{d}{2^c}}$, using which, $\exp(-\frac{\Omega(\log ||\mathbf{x}||)}{(\log N)^{d/2^{c+1}}}) \le \exp(-\Omega((\log N)^{d/2^{c+1}}))$. It follows, that the probability that no node in $V(\mathbf{x})$ is connected to $\mathbf{x}$ by a path of length two, is at most

$$O(\frac{1}{(\log N)^{d/2^c}}) + \exp(-\Omega((\log N)^{d/2^{c+1}})) = O(\frac{1}{(\log N)^{d/2^c}}).$$

Summarizing, conditioned on $\mathbf{X}_{r-1} = \mathbf{x}$, the bound $||\mathbf{X}_r|| \le \frac{||\mathbf{x}||}{(\log N)^{1/2^{c+1}}}$ holds with probability at least $1 - O(\frac{d}{(\log N)^{1/2^c}})$. On the other hand, with probability one $||\mathbf{X}_r|| \le ||\mathbf{X}_{r-1}||$. We conclude

$$E\Big[||\mathbf{X}_r||\Big|\mathbf{X}_{r-1} = \mathbf{x}\Big] \le \frac{||\mathbf{x}||}{(\log N)^{1/2^{c+1}}} + O(\frac{||\mathbf{x}||}{(\log N)^{d/2^c}}) = O(\frac{||\mathbf{x}||}{(\log N)^{1/2^{c+1}}}).$$

This completes the proof of Lemma 6.2. □

We now complete the proof of Lemma 6.1. Note, that for any $2 \le r \le m$, $E[\mathbf{X}_r|\mathbf{X}_{r-1}, \mathbf{X}_{r-2}, \ldots, \mathbf{X}_1] = E[\mathbf{X}_r|\mathbf{X}_{r-1}]$. We denote $\exp((\log N)^{d/2^c})$ by $\alpha(N)$. We have,

$$\text{Prob}\{||\mathbf{X}_m|| > \alpha(N)\} = \sum_{\alpha(N)<||\mathbf{x}_m||\le||\mathbf{x}_{m-1}||<||\mathbf{z}_0||} \text{Prob}\{\mathbf{X}_m = \mathbf{x}_m|\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\}\text{Prob}\{\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\} \le$$

$$\sum_{\alpha(N)<||\mathbf{x}_m||\le||\mathbf{x}_{m-1}||<||\mathbf{z}_0||} ||\mathbf{x}_m||\text{Prob}\{\mathbf{X}_m = \mathbf{x}_m|\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\}\text{Prob}\{\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\} \le$$

$$\sum_{\alpha(N)<||\mathbf{x}_{m-1}||<||\mathbf{z}_0||} E\Big[||\mathbf{X}_m||\Big|\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\Big]\text{Prob}\{\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\}.$$

But, using bound (2) of Lemma 6.2, we have $E\Big[||\mathbf{X}_m||\Big|\mathbf{X}_{m-1} = \mathbf{x}_{m-1}] \le O(||\mathbf{x}_{m-1}||/(\log N)^{1/2^{c+1}})$, as long as $||\mathbf{x}_{m-1}|| > \alpha(N)$. We obtain

$$\text{Prob}\{||\mathbf{X}_m|| > \alpha(N)\} \le O(\frac{1}{(\log N)^{1/2^{c+1}}}) \sum_{\alpha(N)<||\mathbf{x}_{m-1}||<||\mathbf{z}_0||} ||\mathbf{x}_{m-1}||\text{Prob}\{\mathbf{X}_{m-1} = \mathbf{x}_{m-1}\} =$$

$$O(\frac{1}{(\log N)^{1/2^{c+1}}}) \sum_{\alpha(N)<||\mathbf{x}_{m-1}||\le||\mathbf{x}_{m-2}||<||\mathbf{z}_0||} ||\mathbf{x}_{m-1}||\text{Prob}\{\mathbf{X}_{m-1} = \mathbf{x}_{m-1}|\mathbf{X}_{m-2} = \mathbf{x}_{m-2}\}\text{Prob}\{\mathbf{X}_{m-2} = \mathbf{x}_{m-2}\} \le$$

$$O(\frac{1}{(\log N)^{1/2^{c+1}}}) \sum_{\alpha(N)<||\mathbf{x}_{m-2}||<||\mathbf{z}_0||} E\Big[||\mathbf{X}_{m-1}||\Big|\mathbf{X}_{m-2} = \mathbf{x}_{m-2}\Big]\text{Prob}\{\mathbf{X}_{m-2} = \mathbf{x}_{m-2}\} \le$$

$$\Big(O(\frac{1}{(\log N)^{1/2^{c+1}}})\Big)^2 \sum_{\alpha(N)<||\mathbf{x}_{m-2}||<||\mathbf{z}_0||} ||\mathbf{x}_{m-2}||\text{Prob}\{X_{m-2} = \mathbf{x}_{m-2}\},$$

where in the last inequality we used bound (2) of Lemma 6.2 again. Continuing this conditioning argument $m-1$ times, we obtain that for some constant C

$$\text{Prob}\{||\mathbf{X}_m|| > \alpha(N)\} \le \frac{C^{m-1}}{(\log N)^{\frac{m-1}{2^{c+1}}}}||\mathbf{z}_0|| \le \frac{(\log N)^{\frac{1}{2^{c+1}}}C^m}{(\log N)^{\frac{m}{2^{c+1}}}}dN.$$

But, by assumption of the lemma, $m = (2d+2)\cdot 2^{c+1}\log N/\log\log N$, from which $(\log N)^{\frac{1}{2^{c+1}}}C^m = o(N)$ and $\text{Prob}\{||\mathbf{X}_m|| > \alpha(N)\} \le 1/N^{2d}$ for large N. □

7 Concluding remarks and open questions

We considered a long-range percolation model on an graph with a node set $\{0, 1, \ldots, N\}^d$. Answering some open questions raised by Benjamini and Berger in [2], we showed that if two nodes at a distance r are connected by an edge with probability $\approx \beta/r^s$, then, with high probability, the diameter of this graph is $\Theta(\frac{\log N}{\log\log N})$ when $s = d$, and is at most N^η for some value $\eta < 1$, when $s = 2d$. We also proved a lower bound $N^{\eta'}, \eta' < 1$ on the diameter for the cases $d = 1, s = 2, \beta < 1$ and $s > 2d, d \geq 1$. Note that for the case $d = 1, s > 2$ our bound is weaker than known linear lower bound $\Omega(N)$ established in [2]. We conjecture that this linear lower bound holds for all dimensions d as long as $s > 2d$. Other unanswered regimes are lower bounds for $s = 2d$ and $d = 1, s = 2, \beta > 1$. It would also be interesting to compute the limits $\frac{D(N)}{(\log N/\log\log N)} \to C$ and $\log D(N)/\log N \to \eta$ or even show that these limits actually exist when $s = d, 2d$ respectively.

Acknowledgments. We wish to thank I. Benjamini and N. Berger for clarifying their work and identifying several errors in an earlier version of this paper.

References

[1] M. Aizenman and C. M. Newman. Discontinuity of the percolation density in one dimensional $1/|x-y|^2$ percolation models. *Commun. Math. Phys.*, 107:611–647, 1986.

[2] I. Benjamini and N. Berger. The diameter of long-range percolation clusters on finite cycles. *Random Structures and Algorithms*, 19(2):102–111, 2001.

[3] I. Benjamini, H. Kesten, Y. Peres, and O. Schramm. The geometry of the uniform spanning forest: transitions in dimensions $4, 8, 12, \ldots$. *In preparation*, 2000.

[4] M. Biskup. *Work in Preparation*, 2001.

[5] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. Cornell Computer Science Technical Report 99-1776, 1999.

[6] S. Milgram. The small world problem. *Psychology Today*, 1(61), 1967.

[7] C. M. Newman and L. S. Schulman. One dimensional $1/|j-i|^s$ percolation models: the existence of a transition for $s \leq 2$. *Commun. Math. Phys.*, 180:483–504, 1986.

[8] L. S. Schulman. Long-range percolation in one dimension. *J.Phys. A*, 16(17):L639–L641, 1983.

[9] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(440), 1998.