The Metric Dimension of Sparse Random Graphs

Josep Díaz[a] Harrison Hartle[b] Cristopher Moore[b]

Submitted: Apr 27, 2025; Accepted: Jan 28, 2025; Published: Mar 27, 2026

Abstract

In 2013, Bollobás, Mitsche, and Prałat gave upper and lower bounds for the likely metric dimension of random Erdős-Rényi graphs $G(n, p)$ for a large range of expected degrees. However, their results only apply when $d = pn = \omega(\log^5 n)$, leaving open sparser random graphs with $d = O(\log^5 n)$ or $d = o(\log^5 n)$. Here we provide upper and lower bounds on the likely metric dimension of $G(n, p)$ in a range of d starting just above the connectivity transition, i.e., where $d = c \log n$ for some constant $c > 1$, up to $d = O(\log^5 n)$. Our lower bound technique is based on an entropic argument which is weaker but more general than the use of Suen's inequality by Bollobás, Mitsche, and Prałat, whereas our upper bound is similar to theirs.

Mathematics Subject Classifications: 05C80, 05C12

1 Introduction

Let $G = (V, E)$ be a graph with $|V| = n$ vertices. Let $d(u, v)$ denote the shortest path distance in G. If $u, v \in V$ are distinct and $w \in V$, we say that w *separates* u, v if $d(u, w) \neq d(v, w)$. We say that a set $W \subset V$ separates u, v if some $w \in W$ does so, and we call W a *separator* if it separates all distinct pairs. (We will sometimes call W the set of *landmark* vertices.) The *metric dimension* of G, denoted $\mathrm{MD}(G)$, is the size of the smallest separator W.

This problem was defined independently by Slater [25] and Harary-Melter [14] more than 50 years ago. From the complexity point of view, determining $\mathrm{MD}(G)$ is NP-complete for general graphs [13, 17], planar graphs (even with bounded degree) [9], split and bipartite graphs [10], and interval and permutation graphs (even with diameter 2) [12]. For general graphs, $\mathrm{MD}(G)$ can be approximated in polynomial time within a factor of $O(\log n)$ [17, 15], but not within an $o(\log n)$ factor unless $\mathsf{P} = \mathsf{NP}$ [1]. Even a $(1 - \epsilon) \log n$ approximation in polynomial time would imply that $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\log \log n})$ [15].

On the positive side, there are polynomial time algorithms for $\mathrm{MD}(G)$ for trees [17], chain graphs [11], co-graphs [10], and outer-planar graphs [9]. There also have been a

[a]Universitat Politècnica de Catalunya, Barcelona, Spain (diazcort@gmail.com).
[b]Santa Fe Institute, Santa Fe, NM, U.S.A. (hhartle@santafe.edu, moore@santafe.edu).

number of works dealing with the parameterized complexity of $\mathrm{MD}(G)$, producing some FPT algorithms under certain parametrizations; see [2] and the bibliography there. In the already cited [17] the authors characterized graphs with metric dimensions 1 and 2, and provided a general bound that appears as Equation (12) below. For further results on the combinatorial nature of the problem, see [6].

Given the difficult nature of the metric dimension problem on many types of graphs, a natural question is to find the typical value of $\mathrm{MD}(G)$ for various random graphs. For the Erdős-Rényi-Gilbert random graph model $G(n,p)$, Bollobás, Mitsche and Prałat [3] largely determined the likely value of $\mathrm{MD}(G)$ quite precisely for a wide range of p; we review their results as Theorem 2 in the next Section. However, their results do not apply unless the average degree is $\omega(\log^5 n)$, leaving sparser random graphs as an open problem.

In this paper we extend the results of [3] to sparser random graphs, and almost all the way down to the connectivity transition at $d = np = \log n$. Specifically, we give upper and lower bounds on $\mathrm{MD}(G)$ whenever $d \geqslant c \log n$ for any $c = c(n) > 1$. In exchange for this wider range of d, our lower bound is somewhat looser than those of [3].

We obtain our results with two techniques. Our lower bound is information-theoretic, using the entropy of the matrix of pairwise distances to show that a certain number of landmarks are needed to distinguish all pairs of vertices. This is significantly simpler than the lower bound method in [3] which uses Suen's inequality.

We note that Ódor and Thiran [20, 21] recently considered an adaptive version of the metric dimension, where there is a hidden vertex $v^\star$ and a player can ask a series of adaptive queries $v_1, \ldots, v_\ell$, each of which yields the distance $d(v_i, v^\star)$. They define the *sequential metric dimension* $\mathrm{SMD}(G)$ of a graph G as the smallest ℓ such that the player can always succeed in identifying $v^\star$. Clearly $\mathrm{SMD}(G) \leqslant \mathrm{MD}(G)$, since the usual metric dimension yields a non-adaptive strategy. They prove lower bounds on $\mathrm{SMD}(G(n, d/n))$ and find that for sufficiently large d the ratio between the MD and SMD is at most a constant. However, like [3], their results only apply when $d = \omega(\log^5 n)$ due to their use of Suen's inequality rather than entropy. In fact our entropic lower bound on $\mathrm{MD}(G(n,p))$ is also a lower bound on $\mathrm{SMD}(G(n,p))$.

Our upper bound technique is identical to that of [3]. Given u and v, consider the "shells" $V_t(u)$ and $V_t(v)$ of vertices whose shortest path distance to u or v respectively is exactly t. If the symmetric difference $V_t(u) \triangle V_t(v)$ is large for some t, then there are many vertices w that separate u and v. If this is true for all pairs u, v, we can construct a separator W simply by choosing landmarks uniformly at random.

Thus both our upper and lower bound work by inductively bounding the size of the shells V_t for $t = 1, 2, 3, \ldots$ It turns out that almost all of the fluctuations in $|V_t|$ occur at the first step of this process, namely in the degree $|V_1|$ of u or v. Proving our bounds for $d = c \log n$ for all constant $c > 1$, as opposed to some larger constant, requires a concentration inequality for the binomial distribution which is sharper than the standard Chernoff bound. Moreover, in the regime $d = O(\log n)$, the vertex degrees really do fluctuate multiplicatively, between αd and βd for some constants α, β: $0 < \alpha \leqslant 1 \leqslant \beta$. This introduces an additional multiplicative gap in our results. This may be unavoidable, since some pairs of vertices really do have fewer separators than other pairs.

Throughout the paper, we assume that the reader is familiar with the basic properties of $G(n,p)$ (see for example [16, 19]). In particular, since we have $d > \log n$, we condition throughout on the event that $G(n,p)$ is connected. We also assume basic knowledge of entropy manipulation (see for example [24, 22, 8]).

We say a series of events $E = (E_n)$ holds with high probability (w.h.p.) if $\Pr[E_n] \to 1$ as $n \to \infty$. We use asymptotic notation in these statements as follows: "w.h.p. $f(n) = O(g(n))$" means that there is a constant C such that w.h.p. $f(n) \leqslant Cg(n)$, and "w.h.p. $f(n) = o(g(n))$" means that w.h.p. $f(n) \leqslant Cg(n)$ for any constant $C > 0$. We use Ω and ω analogously, if $f(n) \geqslant Cg(n)$ w.h.p. for some, or all, constants $C > 0$.

2 Two cases

The behavior of the metric dimension of $G(n, d/n)$ turns out to be somewhat strange. Imagine a series of shells V_t surrounding a vertex, consisting of the vertices at distance $t = 1, 2, 3 \ldots$. Roughly speaking, we expect the number of vertices in V_t to grow as d^t where $d = np$. The question is what happens when t gets close to the diameter of G, so that the tth shell V_t contains a significant fraction of all n vertices. This depends on the nearest power of d to n.

Following [3] with small changes, we adopt the following notation, which we will use throughout.

Definition 1. Given n and $d = d(n) = \omega(1)$ and letting $p = d/n$, let $t^* = t^*(n)$ be the largest integer t such that $d^t = o(n)$. Let

$$\gamma = \gamma(n) = pd^{t^*} = d^{t^*+1}/n \,. \tag{1}$$

We then distinguish two cases:

- Case 1: $\gamma = \Theta(1)$.
- Case 2: $\gamma = \omega(1)$.

Note that this definition implies that $\gamma = o(d)$.

As we will see, the key difference between Case 1 and Case 2 is whether the $(t^* + 1)$st and (t^*+2)nd shells each contain $\Theta(n)$ vertices, with all others containing $o(n)$, or whether the $(t^* + 1)$st shell contains all but $o(n)$ vertices. This in turn determines whether the entropy of the distribution of distances from a given vertex is $\Theta(1)$ or $o(1)$, and whether a typical pair of vertices has $\Theta(n)$ or $o(n)$ separators. As a result, the metric dimension will be $\Theta(\log n)$ in Case 1 and $\omega(\log n)$ in Case 2.

As pointed out in [3], if $d = c \log n$ or $d = \log^2 n$, say, this causes $\mathrm{MD}(G)$ to undergo a curious zig-zag behavior. As n increases we briefly visit Case 1 whenever n is close to an integer power of d. In between these integer powers, we have Case 2. Their main result, slightly rewritten to focus on the sparse case $d = o(n)$, is the following:

Theorem 2 ([3]). *Let $G = G(n,p)$ where $p = d/n$ and $d = d(n) = \omega(\log^5 n)$ and $d = o(n)$. Let t^* and γ be as in Definition 1. In Case 1 where $\gamma = \Theta(1)$, w.h.p.*

$$\mathrm{MD}(G) = (1 \pm o(1))\, \frac{2\log n}{-\log q} = \Theta(\log n)\,,$$

where

$$q = (\mathrm{e}^{-\gamma})^2 + (1-\mathrm{e}^{-\gamma})^2\,.$$

In Case 2 where $\gamma = \omega(1)$, w.h.p.

$$\mathrm{MD}(G) = \Theta\left(\frac{\log n}{\gamma/d + \mathrm{e}^{-\gamma}}\right) = \omega(\log n)\,.$$

3 Our results

We now state our bounds. They are somewhat verbose, especially when $d = O(\log n)$. Nevertheless, they determine the likely value of $\mathrm{MD}(G)$ to within a constant or slowly-growing function of n .

Let

$$H(y) = -y\log y - (1-y)\log(1-y)$$

be the binary entropy function and let

$$q(\alpha,\beta) = 1 - 2\,(1-\mathrm{e}^{-\alpha\gamma})\,\mathrm{e}^{-\beta\gamma}\,.$$

Theorem 3. *Let $G = G(n,p)$ where $p = d/n$ and $c\log n \leqslant d \leqslant \log^5 n$ for some constant $c > 1$. Then there are constants $0 < \alpha \leqslant 1 \leqslant \beta$ depending on c such that the following holds.*

In Case 1 where $\gamma = \Theta(1)$, w.h.p.

$$(1-o(1))\,\frac{\log n}{\max_{x\in[\alpha,\beta]} H(\mathrm{e}^{-x\gamma})} \leqslant \mathrm{MD}(G) \leqslant (1+o(1))\,\frac{2\log n}{-\log q(\alpha,\beta)}\,. \tag{2}$$

If $d = \omega(\log n)$ then

$$(1-o(1))\,\frac{\log n}{H(\mathrm{e}^{-\gamma})} \leqslant \mathrm{MD}(G) \leqslant (1+o(1))\,\frac{2\log n}{-\log q}\,, \tag{3}$$

where $q = q(1,1) = (\mathrm{e}^{-\gamma})^2 + (1-\mathrm{e}^{-\gamma})^2$ as in Theorem 2.

In Case 2 where $\gamma = \omega(1)$, w.h.p.

$$(1-o(1))\,\frac{\log n}{-(\beta\gamma/d)\log(\beta\gamma/d) + \gamma\mathrm{e}^{-\alpha\gamma}} \leqslant \mathrm{MD}(D) \leqslant (1+o(1))\,\frac{\log n}{\alpha\gamma/d + \mathrm{e}^{-\beta\gamma}}\,, \tag{4}$$

and if $d = \omega(\log n)$ we have

$$(1-o(1))\,\frac{\log n}{-(\gamma/d)\log(\gamma/d) + \gamma\mathrm{e}^{-\gamma}} \leqslant \mathrm{MD}(G) \leqslant (1+o(1))\,\frac{\log n}{\gamma/d + \mathrm{e}^{-\gamma}}\,. \tag{5}$$

We will prove Theorem 3 by breaking it into two parts, namely Theorem 17 and Theorem 21 for the lower and upper bounds respectively.

We can summarize these bounds as follows. First, our upper bounds on $\mathrm{MD}(G)$ match those in [3], with the complication that when $d = O(\log n)$ we have to introduce the constants α and β to handle fluctuations in vertex degrees. Our lower bounds are somewhat looser than those of [3], but have the advantage that they apply when $d \geqslant c \log n$ for any constant $c > 1$ as opposed to $d = \omega(\log^5 n)$.

Case 1. When $\gamma = \Theta(1)$, our upper and lower bounds differ by a multiplicative constant depending on γ, and w.h.p.

$$\mathrm{MD}(G) = \Theta(\log n)\,.$$

When $d = c \log n$ for constant c this gap also depends on c, since it depends on α and β. But since $H(y) \leqslant \log 2$ for all y, we always have w.h.p.

$$\mathrm{MD}(G) \geqslant (1 - o(1)) \log_2 n\,.$$

Case 2. Since $\gamma = o(d)$ and $\gamma = \omega(1)$, the denominators in (4) and (5) are all $o(1)$, so w.h.p. $\mathrm{MD}(G) = \omega(\log n)$. However, in this case our upper and lower bounds are more than a constant apart. Moreover, the scaling of $\mathrm{MD}(G)$ with n depends on which term in the denominators of our bounds is larger, and this in turn depends on how γ grows with n. The crossover occurs roughly when $\gamma = \log d$.

First suppose $d = \omega(\log n)$, and recall that $d \leqslant \log^5 n$ so $\log d = O(\log \log n)$. Then (5) gives the following.

- If $d = \omega(\log n)$ and $\gamma \geqslant \log d$, then w.h.p.

  $$(1 - o(1)) \frac{\log n}{-(\gamma/d)\log(\gamma/d)} \leqslant \mathrm{MD}(G) \leqslant (1 + o(1)) \frac{\log n}{\gamma/d}\,,$$

  and so

  $$\Omega\left(\frac{d \log n}{\gamma \log \log n}\right) \leqslant \mathrm{MD}(G) \leqslant O\left(\frac{d \log n}{\gamma}\right)\,.$$

- If $d = \omega(\log n)$ and $\gamma \leqslant \delta \log d$ for some $\delta < 1$, then w.h.p.

  $$(1 - o(1)) \frac{\log n}{\gamma \mathrm{e}^{-\gamma}} \leqslant \mathrm{MD}(G) \leqslant (1 + o(1)) \frac{\log n}{\mathrm{e}^{-\gamma}}\,.$$

  In particular, if $\gamma = \delta \log d$ then

  $$\Omega\left(\frac{d^{\delta} \log n}{\log \log n}\right) \leqslant \mathrm{MD}(G) \leqslant O(d^{\delta} \log n)\,.$$

- Combining these, if $d = \omega(\log n)$ the multiplicative gap between our upper and lower bounds is $O(\min(\log \log n, \gamma))$.

Finally, consider Case 2 when $d = c \log n$ for some constant $c > 1$. Here the situation is complicated by the fact that $\alpha < 1 < \beta$, and that these appear in the exponents $\mathrm{e}^{-\alpha\gamma}$ and $\mathrm{e}^{-\beta\gamma}$. This creates a pair of crossovers with a range of γ in between. From (4) we have the following.

- If $d = c \log n$ and $\gamma \geqslant (1/\alpha) \log d$, then w.h.p.

$$(1 - o(1)) \frac{\log n}{-(\beta\gamma/d) \log(\beta\gamma/d)} \leqslant \mathrm{MD}(G) \leqslant (1 + o(1)) \frac{\log n}{\alpha\gamma/d}$$

  and so

$$\Omega\left(\frac{\log^2 n}{\gamma \log\log n}\right) \leqslant \mathrm{MD}(G) \leqslant O\left(\frac{\log^2 n}{\gamma}\right) .$$

- If $d = c \log n$ and $\gamma = \delta \log d$ for $1/\beta \leqslant \delta < 1/\alpha$, then w.h.p.

$$(1 - o(1)) \frac{\log n}{\gamma \mathrm{e}^{-\alpha\gamma}} \leqslant \mathrm{MD}(D) \leqslant (1 + o(1)) \frac{\log n}{\alpha\gamma/d} ,$$

  and so

$$\Omega\left(\frac{\log^{1+\alpha\delta} n}{\log\log n}\right) \leqslant \mathrm{MD}(G) \leqslant O\left(\frac{\log^2 n}{\log\log n}\right) . \tag{6}$$

- If $d = c \log n$ and $\gamma = \delta \log d$ for $\delta < 1/\beta$, then w.h.p.

$$(1 - o(1)) \frac{\log n}{\gamma \mathrm{e}^{-\alpha\gamma}} \leqslant \mathrm{MD}(D) \leqslant (1 + o(1)) \frac{\log n}{\mathrm{e}^{-\beta\gamma}} ,$$

  and so

$$\Omega\left(\frac{\log^{1+\alpha\delta} n}{\log\log n}\right) \leqslant \mathrm{MD}(G) \leqslant O\left(\frac{\log^{1+\beta\delta} n}{\log\log n}\right) . \tag{7}$$

The multiplicative gap between our bounds is most severe in (6) and (7), where we remain uncertain about the power of $\log n$. The source of this gap is that if $d = c \log n$ for constant c, then there typically exist vertices with degrees ranging from αd to βd for some $\alpha < 1 < \beta$ depending on c. As a result, different pairs of vertices really do have different numbers of separators that distinguish them, and different vertices have different distributions of distances to the rest of the graph. This doesn't necessarily mean that $\mathrm{MD}(G)$ fluctuates from one realization of $G(n, d/n)$ to another, but it does make it harder to prove that it is concentrated around some typical value.

4 An entropic lower bound on metric dimension

We begin by proving a general lower bound on the width of a matrix whose rows are distinct and whose columns have bounded entropy. Note that this lemma applies to arbitrary deterministic matrices, although we will apply it below to a random matrix.

Definition 4. Let A be a matrix with n rows and m columns with entries in a finite set S. For each $1 \leqslant j \leqslant m$ and each $s \in S$, let $p_j(s) = |\{i : A_{ij} = s\}|/n$ denote the frequency of s in the jth column, and define the entropy

$$H_j(A) = -\sum_{s\in S} p_j(s) \log p_j(s) \, ,$$

where $0 \log 0 := 0$. When the base matters, our logarithms are natural.

Remark 5. Even though we will consider the case where each A_{ij} is a random variable, namely the shortest path distance in a random graph, the entropy H_j is not the entropy of that random process. Rather, H_j is the entropy of the frequency distribution of the jth column of a fixed matrix. In statistical language, it is the entropy of the empirical distribution of that column.

Lemma 6. *Let A be an $n \times m$ matrix and let $H_j = H_j(A)$. Suppse that $H_j \leqslant \widehat{H}$ for all $1 \leqslant j \leqslant m$ for some $\widehat{H}$. Then if the n rows of A are distinct, we have*

$$m \geqslant \frac{\log n}{\widehat{H}} \, .$$

Proof. This is a classic information-theoretic proof. While we will phrase it in terms of random variables, the entries A_{ij} of the matrix A are fixed; the only randomness will come from picking a random row i.

For each fixed column $j \in \{1, \ldots, m\}$, consider the random variable A_{ij} where i is uniformly random in $\{1, \ldots, n\}$. Then $p_j(s)$ is the probability distribution of this variable, i.e., the empirical distribution of s in column j, and H_j is its entropy.

Now consider the random variable $\mathbf{r}$ consisting of a randomly chosen row of A, namely $\mathbf{r} = (A_{i,1}, A_{i,2}, \ldots, A_{i,m})$ where i is uniformly random in $\{1, \ldots, n\}$. Let $H_{\mathbf{r}}$ be the entropy of the distribution of $\mathbf{r}$. For each $1 \leqslant j \leqslant m$ the marginal distribution of the jth component $\mathbf{r}_j = A_{ij}$ is distributed according to p_j. Thus by the sub-additivity property of entropy we have

$$H_{\mathbf{r}} \leqslant \sum_{j=1}^{m} H_j \leqslant m\widehat{H} \, . \tag{8}$$

since $\max_j H_j \leqslant \widehat{H}$ by assumption. If the n rows are distinct, then each row appears with probability $1/n$ and $H_{\mathbf{r}} = \log n$. Combining this with (8) completes the proof. □

Remark 7. This entropic argument has the advantage that it does not require any independence in the entries of A. A stronger lower bound on the number of columns is proved in [20] in the case where the entries of A are independent Bernoulli variables.

We use this entropic argument to lower bound the metric dimension as follows. Consider a graph $G = (V, E)$, and let $W \subseteq V$ be a set of vertices. Let $n = |V|$ and $m = |W|$. Then let A be the matrix of shortest path distances $A_{v,w} = d(v, w)$, where each row corresponds to some $v \in V$ and each column corresponds to some $w \in W$. If W is a separator, then the n rows of A are distinct. Applying Lemma 6 then yields a lower bound on m, and thus on $\mathrm{MD}(G)$.

We use the following notation for the "shells" of vertices at a given distance from w:

Definition 8. Let $G = (V, E)$ with $|V| = n$ and diameter Diam. For each $w \in V$ and each $0 \leqslant t \leqslant \mathrm{Diam}$, let

$$V_t(w) = \{v \in V : d(v, w) = t\}$$

be the set of vertices at distance exactly t from w. We will also write

$$V_{\leqslant t} = \bigcup_{t'=0}^{t} V_{t'} \quad \text{and} \quad V_{>t} = V \setminus V_{\leqslant t} = \bigcup_{t'=t+1}^{\mathrm{Diam}} V_{t'} \,.$$

Applying Lemma 6 to the matrix of distances, in Definition 4 we have $p_w(t) = |V_t(w)|/n$. This yields the following lower bound on $\mathrm{MD}(G)$ for any graph:

Corollary 9. *For any graph $G = (V, E)$ with $|V| = n$ and diameter* Diam*, and for a given $w \in V$, let*

$$H_w = -\sum_{t=0}^{\mathrm{Diam}} \frac{|V_t(w)|}{n} \log \frac{|V_t(w)|}{n} \,. \tag{9}$$

Then if $H_w \leqslant \widehat{H}$ for all $w \in V$, we have

$$\mathrm{MD}(G) \geqslant \frac{\log n}{\widehat{H}} \,.$$

Remark 10. This is also a lower bound on the sequential metric dimension $\mathrm{SMD}(G)$ considered in [20], since the player can gain at most $\widehat{H}$ bits of information with each query, and they need $\log n$ bits to identify the hidden vertex.

As a simple application, for any graph G with diameter Diam we have $\widehat{H} \leqslant \log(\mathrm{Diam}+1)$ since there are $\mathrm{Diam}+1$ distinct possible distance values. Thus, if G has n vertices we have

$$\mathrm{MD}(G) \geqslant \frac{\log n}{\log(\mathrm{Diam}+1)} \,. \tag{10}$$

We can improve this slightly by focusing on the $(n-m) \times m$ matrix of distances between $V \setminus W$ and W. Then $S = \{1, \ldots, \mathrm{Diam}\}$ and

$$\mathrm{MD}(G) \geqslant \frac{\log(n - \mathrm{MD}(G))}{\log \mathrm{Diam}} \,. \tag{11}$$

This is equivalent to a bound from [17],

$$n \leqslant \mathrm{Diam}^{\mathrm{MD}(G)} + \mathrm{MD}(G) \,. \tag{12}$$

However, the lower bounds (10) and (11) are quite pessimistic. We can only have $\widehat{H} = \log(\mathrm{Diam}+1)$ if, for some w, the distribution of distances $d(v, w)$ is uniform between

0 and Diam. This holds for a path graph on n vertices, where $\mathrm{Diam} = n - 1$ and these bounds give $\mathrm{MD}(G) \geqslant 1$. Indeed, for path graphs we have $\mathrm{MD}(G) = 1$ since a single vertex at either end is a separator.

But in an expander, in particular w.h.p. in $G(n, p)$, this is far from true: only a small fraction of vertices have distances much less than the diameter. In fact w.h.p. for any w the distribution of distances $d(v, w)$ is concentrated on one or two values. This was already known [3] for $d = \omega(\log n)$; we give a proof here that includes the case $d = O(\log n)$. In that case we have $\widehat{H} = O(1)$, improving the lower bound from $\mathrm{MD}(G) \geqslant (\log n)/(\log \mathrm{Diam}) = \Omega(\log n/ \log\log n)$ to $\mathrm{MD}(G) = \Omega(\log n)$.

The following two lemmas overlap with classic results about expansion in $G(n, p)$ (e.g. [4, 7]) but we tune them to our needs and give simple self-contained proofs. First we show that w.h.p. all vertices have degree within a multiplicative factor of d.

Lemma 11. *Let $G = G(n, p)$ where $p = d/n$ and $d \geqslant c \log n$ for some constant $c > 1$. Then there are constants α, β depending only on c with $0 < \alpha \leqslant 1 \leqslant \beta < \mathrm{e}$ such that the following holds w.h.p.: for all $w \in V$,*

$$(1 - o(1))\, \alpha d \leqslant \deg w \leqslant (1 + o(1))\, \beta d\,, \tag{13}$$

and this holds with probability $1 - o(1/n)$ for any fixed w. As c increases α and β tend to 1 from below and above respectively. If $d = \omega(\log n)$, (13) holds with $\alpha = \beta = 1$.

We prove Lemma 11 in Section 6. While it ensures that every vertex has degree $\Theta(d)$, it also means that the size $|V_1(w)| = \deg w$ of the first shell can fluctuate significantly when $d = c \log n$. On the other hand, the following lemma shows that subsequent shells grow by almost exactly a factor of d at each step until $t = t^*$.

Lemma 12. *Let $G = G(n, p)$ where $p = d/n$ and $d \geqslant c \log n$ for some constant $c > 1$. With t^* as in Definition 1 and $V_t(w)$ as in Definition 8, w.h.p. for all $w \in V$ and all $t \leqslant t^*$,*

$$|V_t(w)| = (1 \pm o(1))\, d^{t-1} \deg w\,. \tag{14}$$

In particular,

$$(1 - o(1))\, \alpha d^t \leqslant |V_t(w)| \leqslant (1 + o(1))\, \beta d^t \tag{15}$$

with α and β as in Lemma 11.

We prove Lemma 12 in Section 7.

Now that we know how the shells V_t grow exponentially with t, the question is how this process ends in the last few shells. First we use a standard argument to show that the diameter of G is concentrated on a few values.

Lemma 13. *Let $G = G(n, p)$ where $p = d/n$ and $d \geqslant c \log n$ for some constant $c > 1$. Let t^* and γ be defined as in Definition 1. Then w.h.p.*

$$\mathrm{Diam}(G) \leqslant t^* + 3 \quad \textit{in Case 1 where } \gamma = \Theta(1),$$
$$\mathrm{Diam}(G) \leqslant t^* + 2 \quad \textit{in Case 2 where } \gamma = \omega(1).$$

Proof. Focusing on Case 2 first, we have $d^{t^*+1} = \gamma n = \omega(n)$. Lemma 12 then implies that w.h.p. for any pair of vertices u, v and any $1 \leqslant s, t \leqslant t^*$ with $s + t = t^* + 1$,

$$|V_s(u)||V_t(v)| \geqslant (1 - o(1))\, \alpha^2 d^{s+t} = \Theta(d^{t^*+1}) = \omega(n)\,.$$

If $V_s(u)$ and $V_t(v)$ intersect, then $d(u, v) \leqslant s + t = t^* + 1$. If they are disjoint but there is an edge between them, then $d(u, v) \leqslant s + t + 1 = t^* + 2$. The probability there is no such edge is

$$(1 - p)^{|V_s(u)||V_t(v)|} = (1 - d/n)^{\omega(n)} \leqslant \mathrm{e}^{-\omega(d)}\,.$$

Since $d \geqslant c \log n$, this is smaller than n^{-C} for any constant C. Taking the union bound over all n^2 pairs u, v thus shows that w.h.p. $\mathrm{Diam}(G) \leqslant s + t + 1 \leqslant t^* + 2$, completing the proof for Case 2.

For Case 1, we use the same argument but with $s + t = t^* + 2$, in which case from $d^{s+t} = d^{t^*+2} = \omega(n)$. □

Remark 14. Since w.h.p.$\mathrm{Diam}(G) \geqslant t^*$, Lemma 13 shows that the diameter is concentrated on at most four values, which was shown by [7]. In fact the diameter is two-point concentrated whenever $d = \omega(1)$ [23], but these bounds are sufficient for our purposes.

Next, the following lemma shows that in Case 1, two shells occupy almost the entire graph, namely V_{t^*+1} and V_{t^*+2}. That is, for all w, the distance $d(w, v)$ takes just two different values for almost all v. In Case 2, just one shell V_{t^*+1} dominates the graph, so $d(w, v)$ takes just one value for almost all v. As a result, the entropy of distances is $O(1)$ and $o(1)$ in Case 1 and Case 2 respectively.

Lemma 15. *Let $G = G(n, p)$ where $p = d/n$ and let $c \log n \leqslant d \leqslant \log^5 n$ for some constant $c > 1$. Let $V_t(w)$, t^*, α, β, and γ be as above. In Case 1 where $\gamma = \Theta(1)$, w.h.p. the following holds for all $w \in V$:*

$$1 - \mathrm{e}^{-\alpha\gamma} - o(1) \leqslant \frac{|V_{t^*+1}(w)|}{n} \leqslant 1 - \mathrm{e}^{-\beta\gamma} + o(1)\,, \tag{16}$$

$$\mathrm{e}^{-\beta\gamma} - o(1) \leqslant \frac{|V_{t^*+2}(w)|}{n} \leqslant \mathrm{e}^{-\alpha\gamma} + o(1)\,, \tag{17}$$

$$\frac{|V_{t^*+1}(w)|}{n} + \frac{|V_{t^*+2}(w)|}{n} = 1 - o(1)\,. \tag{18}$$

In Case 2 where $\gamma = \omega(1)$, w.h.p. for all $w \in V$:

$$\frac{|V_{t^*+1}(w)|}{n} \geqslant 1 - (1 + o(1)) \left(\frac{\beta\gamma}{d} + \mathrm{e}^{-\alpha\gamma} \right)\,. \tag{19}$$

We prove Lemma 15 in Section 7.

Remark 16. We restrict our statement to $d \leqslant \log^5 n$ for convenience and because when $d = \omega(\log^5 n)$ the results of [3] apply. We claim that in fact Lemma 15 holds for any $d = O(n/\log n)$.

Finally, we combine Lemma 15 with the entropic bound of Lemma 6 to prove our lower bounds:

Theorem 17. *In Case 1 where $\gamma = \Theta(1)$, w.h.p.*

$$\text{if } d = O(\log n), \quad \mathrm{MD}(G) \geqslant (1 - o(1)) \frac{\log n}{\max_{x\in[\alpha,\beta]} H(\mathrm{e}^{-x\gamma})} \tag{20}$$

$$\text{if } d = \omega(\log n), \quad \mathrm{MD}(G) \geqslant (1 - o(1)) \frac{\log n}{H(\mathrm{e}^{-\gamma})}\,, \tag{21}$$

where $H(y) = -y \log y - (1-y)\log(1-y)$ is the binary entropy function. In Case 2 where $\gamma = \omega(1)$, w.h.p.

$$\text{if } d = O(\log n), \quad \mathrm{MD}(G) \geqslant (1 - o(1)) \frac{\log n}{-(\beta\gamma/d)\log(\beta\gamma/d) + \gamma \mathrm{e}^{-\alpha\gamma}} \tag{22}$$

$$\text{if } d = \omega(\log n), \quad \mathrm{MD}(G) \geqslant (1 - o(1)) \frac{\log n}{-(\gamma/d)\log(\gamma/d) + \gamma \mathrm{e}^{-\gamma}}\,. \tag{23}$$

Proof. Given Lemma 6, we simply have to show that w.h.p. the entropy H_w of distances for each w is less than some $\widehat{H}$ for all w. The denominators in (20)–(23) are then essentially $\widehat{H}$ in each case.

Recall that

$$H_w = -\sum_{t=0}^{\mathrm{Diam}} \frac{|V_t(w)|}{n} \log \frac{|V_t(w)|}{n}\,.$$

We first bound the contribution to H_w from the shells V_t for $t \leqslant t^*$, each of which occupies $o(n)$ vertices. In both cases this contribution is $o(1)$, but in Case 2 it contributes importantly to the denominator. Since $-x \log x$ is monotonically increasing for $x < 1/\mathrm{e}$, Lemma 12 gives

$$\begin{aligned} -\sum_{t=0}^{t^*} \frac{|V_t(w)|}{n} \log \frac{|V_t(w)|}{n} &\leqslant -(1+o(1)) \sum_{t=0}^{t^*} \frac{\beta d^t}{n} \log \frac{\beta d^t}{n} \\ &\leqslant -(1+o(1)) \frac{\beta d^{t^*}}{n} \log \frac{\beta d^{t^*}}{n} \\ &= -(1+o(1)) \frac{\beta\gamma}{d} \log \frac{\beta\gamma}{d}\,, \end{aligned} \tag{24}$$

where we used the fact that, analogous to (49), the sum over t is dominated by its largest term, namely the contribution from V_{t^*}. This contribution is $o(1)$ since $\gamma/d = o(1)$.

In Case 1, Eqs. (16), (17), and (18) in Lemma 15 imply that the contribution from the two large shells is

$$\begin{aligned} &-\frac{|V_{t^*+1}(w)|}{n} \log \frac{|V_{t^*+1}(w)|}{n} - \frac{|V_{t^*+2}(w)|}{n} \log \frac{|V_{t^*+2}(w)|}{n} \\ &\leqslant H\left(\frac{|V_{t^*+1}|}{n}\right) + o(1) = H(\mathrm{e}^{-x\gamma}) + o(1) \quad \text{for some } x \in [\alpha, \beta]. \end{aligned}$$

By Lemma 13 we have at most one more shell V_{t^*+3}, and since its size is $o(n)$ it contributes $o(1)$ to H_w. Thus we have shown that w.h.p. in Case 1, $H_w \leqslant \widehat{H}$ for all w where

$$\widehat{H} = \max_{x \in [\alpha,\beta]} H(\mathrm{e}^{-x\gamma}) + o(1)\,.$$

This proves (20), and recalling from Lemma 12 that $\alpha = \beta = 1$ when $d = \omega(\log n)$ proves (21).

In Case 2, the contribution (24) of V_{t^*} to H_w matters, as does the contribution from V_{t^*+2}. As in the proof of Lemma 15, either the expectation of $|V_{t^*+2}|/n$ is $\Omega(1/d) = \Omega(1/\log^5 n)$, in which case Chernoff bounds ensure that $|V_{t^*+2}|/n \leqslant (1+o(1))\,\mathrm{e}^{-\alpha\gamma}$ with probability $1 - o(1/n)$, or $|V_{t^*+2}|/n = o(1/d)$ with the same probability. In either case, w.h.p. for all w the contribution of V_{t^*} and V_{t^*+2} to H_w can be bounded as

$$-\frac{V_{t^*}}{n}\log\frac{V_{t^*}}{n} - \frac{V_{t^*+2}}{n}\log\frac{V_{t^*+2}}{n} \leqslant (1+o(1))\left(-\frac{\beta\gamma}{d}\log\frac{\beta\gamma}{d} + \alpha\gamma\mathrm{e}^{-\alpha\gamma}\right), \tag{25}$$

where we pessimistically maximized the two terms separately.

Finally, since $-(1-p)\log(1-p) \leqslant p$, the contribution of V_{t^*+1} to H_w is at most $(|V_{\leqslant t^*}| + |V_{t^*+2}|)/n$, which due to the logarithmic terms is small compared to (25). This completes the proof of (22), and again recalling that $\alpha = \beta = 1$ when $d = \omega(\log n)$ gives (23). ☐

5 The upper bound

Our upper bound on $\mathrm{MD}(G(n,p))$ is essentially identical to that in [3]. The only difference is that, as in our lower bound, fluctuations in vertex degrees have to be taken into account when $d = O(\log n)$.

We follow a standard strategy. We will show that, with high probability, all pairs of vertices u, v have a large number of vertices w that separate them. Therefore, if we construct W by choosing enough landmarks w uniformly at random, there is a good chance that W separates every pair. The following lemma is standard, but we include a proof for the reader.

Lemma 18. *Let $G = (V,E)$ be a graph with $|V| = n$. For each distinct pair of vertices $u, v \in V$, let*

$$S(u,v) = \{w \in V : d(u,w) \neq d(v,w)\}$$

denote the set of vertices w that separate u, v. Suppose that for all u, v we have $S(u,v) \geqslant \sigma n$ for some $\sigma = \sigma(n)$. Then

$$\mathrm{MD}(G) \leqslant \left\lceil \frac{2\log n}{-\log(1-\sigma)} \right\rceil .$$

Proof. Let $Z = \lceil (2\log n)/(-\log(1-\sigma)) \rceil$ and let $W \subseteq V$ be a set of Z vertices chosen uniformly with replacement. For a given pair u, v, each $w \in W$ independently distinguishes them with probability at least σ. Therefore, the probability that u and v are not separated by any $w \in W$ is at most $(1-\sigma)^Z \leqslant \mathrm{e}^{-2\log n} = 1/n^2$. Taking the union bound over the $\binom{n}{2}$ distinct pairs, W is a separator with probability at least $1 - \binom{n}{2}/n^2 \geqslant 1/2$. Thus some separator W of size at most Z exists, and $\mathrm{MD}(G) \leqslant Z$. $\square$

We will apply Lemma 18 as follows. For a given t if we write

$$\Delta_t(u,v) = V_t(u) \triangle V_t(v)$$

where $\triangle$ denotes the symmetric difference, then

$$S(u,v) = \bigcup_{t=0}^{\mathrm{Diam}} \Delta_t(u,v)\,. \tag{26}$$

To lower bound $|S(u,v)|$, we first show that after t^* steps of expansion, the shells for any two distinct vertices are nearly disjoint, so that their symmetric difference is almost as large as their union:

Lemma 19. *Let $G = G(n,p)$ where $p = d/n$ and $c\log n \leqslant d \leqslant \log^5 n$ for some constant $c > 1$, let t^* be defined as in Definition 1, and let α be defined as in Lemma 11. Then w.h.p. the following holds for all distinct pairs of vertices $u, v \in V$:*

$$\begin{aligned} |\Delta_{t^*}(u,v)| &= (1 \pm o(1))\, d^{t^*-1}(\deg u + \deg v) \\ &\geqslant (1-o(1))\,\frac{2\alpha\gamma n}{d}\,. \end{aligned} \tag{27}$$

We prove Lemma 19 in Section 8 by combining the proof of Lemma 12 with the fact that w.h.p. no two vertices have more than two common neighbors.

Lemma 19 establishes that u and v have a fairly large number of separators in their t^*th shells. They might have even more in their (t^*+1)st shell: in fact, in Case 1 we will show that $|\Delta_{t^*+1}| = \Theta(n)$. Thus, analogous to Lemma 15, the next lemma lower bounds $|S(u,v)|$ either by including both these shells or just the (t^*+1)st, depending on the case.

Lemma 20. *Let $G = G(n,p)$ where $p = d/n$, and let $d, t^*, \gamma, \alpha, \beta$ be as above, with $c\log n \leqslant d \leqslant \log^5 n$ for some constant $c > 1$. In Case 1 where $\gamma = \Theta(1)$, w.h.p. for all distinct $u, v \in V$:*

$$\begin{aligned} \frac{|S(u,v)|}{n} &\geqslant \frac{|\Delta_{t^*+1}(u,v)|}{n} \\ &\geqslant (2-o(1))\,(1-\mathrm{e}^{-\alpha\gamma})\,\mathrm{e}^{-\beta\gamma}\,. \end{aligned} \tag{28}$$

In Case 2 where $\gamma = \omega(1)$, w.h.p. for all distinct $u, v \in V$:

$$\begin{aligned} \frac{|S(u,v)|}{n} &\geqslant \frac{|\Delta_{t^*}(u,v)| + |\Delta_{t^*+1}(u,v)|}{n} \\ &\geqslant (2-o(1))\left(\frac{\alpha\gamma}{d} + \mathrm{e}^{-\beta\gamma}\right)\,. \end{aligned} \tag{29}$$

We prove Lemma 20 in Section 8.

Finally, combining Lemmas 20 and 18 gives the following upper bounds on $\mathrm{MD}(G)$. In the following the function q mirrors that in Theorem 2 from [3].

Theorem 21. *Let $G = G(n,p)$ where $p = d/n$ and $c \log n \leqslant d \leqslant \log^5 n$ for some constant $c > 1$, and let α, β, and γ be as above.*

In Case 1 where $\gamma = \Theta(1)$, w.h.p.

$$\text{if } d = O(\log n), \quad \mathrm{MD}(G) \leqslant (1+o(1)) \frac{2 \log n}{-\log q(\alpha, \beta)}, \tag{30}$$

$$\text{if } d = \omega(\log n), \quad \mathrm{MD}(G) \leqslant (1+o(1)) \frac{2 \log n}{-\log q}. \tag{31}$$

where

$$q(\alpha,\beta) = 1 - 2\,(1 - \mathrm{e}^{-\alpha\gamma})\,\mathrm{e}^{-\beta\gamma}$$
$$q = q(1,1) = 1 - 2(1 - \mathrm{e}^{-\gamma})\,\mathrm{e}^{-\gamma} = (\mathrm{e}^{-\gamma})^2 + (1 - \mathrm{e}^{-\gamma})^2.$$

In Case 2, w.h.p.

$$\text{if } d = O(\log n), \quad \mathrm{MD}(G) \leqslant (1+o(1)) \frac{\log n}{\alpha\gamma/d + \mathrm{e}^{-\beta\gamma}}, \tag{32}$$

$$\text{if } d = \omega(\log n), \quad \mathrm{MD}(G) \leqslant (1+o(1)) \frac{\log n}{\gamma/d + \mathrm{e}^{-\gamma}}. \tag{33}$$

Proof. In Case 1, we apply Lemma 18 with the lower bound given by Lemma 20,

$$\sigma = (2 - o(1))\,(1 - \mathrm{e}^{-\alpha\gamma})\,\mathrm{e}^{-\beta\gamma}$$

and write $q(\alpha, \beta) = 1 - \sigma$ (absorbing the error term).

In Case 2, we likewise apply Lemma 18 but with the lower bound

$$\sigma = (2 - o(1)) \left(\frac{\alpha\gamma}{d} + \mathrm{e}^{-\gamma\beta} \right),$$

where we used $\mathrm{e}^{-\gamma\alpha} = o(1)$. Moreover, in this case we have $\sigma = o(1)$ since $\gamma/d = o(1)$ and $\gamma = \omega(1)$. Then we use $-\log(1-\sigma) = (1+o(1))\,\sigma$ for the denominator in Lemma 18.

In both cases, as per Lemma 11, when $\gamma = \omega(1)$ and $d = \omega(\log n)$ we can take $\alpha = \beta = 1$, yielding (31) and (33). In particular, $q = q(1,1)$ coincides with the quantity q used in Theorem 2 from [3]. $\square$

Along with Theorem 17, this completes the proof of our main Theorem 3.

6 Concentration: Proof of Lemma 11

We state two versions of the Chernoff bound. For the induction argument in Lemma 12, we will use the following:

Lemma 22. *Let X be a binomial random variable with mean μ. Then for any $0 < \epsilon < 1$,*

$$\Pr\left[\left|\frac{X}{\mu} - 1\right| > \epsilon\right] \leqslant 2\mathrm{e}^{-\epsilon^2\mu/3}\,. \tag{34}$$

In particular, if $\mu = \Omega(\log^t n)$ for $t \geqslant 2$, then there is a constant A such that

$$(1-\epsilon)\mu \leqslant X \leqslant (1+\epsilon)\mu \quad \textit{where} \quad \epsilon = \frac{A}{\log^{(t-1)/2} n} \tag{35}$$

with probability $1 - o(1/n^3)$.

Proof. We define ϵ in (35) so that $\epsilon^2\mu = \Omega(\log n)$. If $\mu \geqslant C \log^t n$, then setting $A = 10/C$ suffices to make the error probability in (34) $o(1/n^3)$. □

We also state the following bound, which is more precise when $\mu = O(\log n)$. We need this bound to prove Lemma 11 for c arbitrarily close to 1; the standard Chernoff bound (35) doesn't suffice until $c > 3$.

Lemma 23. *Let X be a binomial random variable with mean μ. Let $0 < \delta < 1$ be a constant. Then there exist constants $0 < \alpha < 1 < \beta$ such that*

$$\Pr[X < \alpha\mu] \leqslant \mathrm{e}^{-\delta\mu} \tag{36}$$

$$\Pr[X > \beta\mu] \leqslant \mathrm{e}^{-\delta\mu}\,. \tag{37}$$

Proof. The moment generating function of $X \sim \mathrm{Bin}(n, p = \mu/n)$ is upper bounded by that of a Poisson variable with mean μ,

$$\mathbb{E}[\mathrm{e}^{\lambda X}] = \left(1 + p(\mathrm{e}^{\lambda} - 1)\right)^n \leqslant \mathrm{e}^{\mu(\mathrm{e}^{\lambda}-1)}\,.$$

Let $0 < \alpha \leqslant 1$. For any $\lambda < 0$, Markov's inequality gives

$$\Pr[X < \alpha\mu] = \Pr[\mathrm{e}^{\lambda X} > \mathrm{e}^{\lambda\alpha\mu}] \leqslant \mathbb{E}[\mathrm{e}^{\lambda X}]/\mathrm{e}^{\lambda\alpha\mu} = \mathrm{e}^{-\mu(1-\mathrm{e}^{\lambda}+\lambda\alpha)}\,,$$

and respectively for $\Pr[X > \beta\mu]$ with λ replaced by $-\lambda$. The right-hand side is minimized when $\lambda = \log\alpha$ (resp. $-\log\alpha$), giving

$$\Pr[X < \alpha\mu] \leqslant \mathrm{e}^{-\mu f(\alpha)} \quad \text{and} \quad \Pr[X > \beta\mu] \leqslant \mathrm{e}^{-\mu f(\beta)} \tag{38}$$

where

$$f(\alpha) = 1 - \alpha + \alpha\log\alpha\,.$$

The function f is continuous for $\alpha \geqslant 0$. Moreover $f(1) = 0$ and $f(0) = f(\mathrm{e}) = 1$, so for any $0 < \delta < 1$ there are constants α, β with $0 < \alpha < 1 < \beta < \mathrm{e}$ such that $f(\alpha) = f(\beta) = \delta$. □

Applying this to the degree of a vertex in $G(n,p)$ yields the following.

Proof of Lemma 11. Fix a vertex w. If $X = \deg w$, then $X \sim \mathrm{Bin}(n-1,p)$ and has mean $\mu = (n-1)\,p = (1-1/n)\,d \geqslant (1-1/n)\,c\log n$. Since $c > 1$, in Lemma 23 we can set $\delta = c^{-\kappa}$ for some $0 < \kappa < 1$. Then the error probabilities in (36) and (37) are

$$\mathrm{e}^{-c^{-\kappa}\mu} = (1-o(1))\,\mathrm{e}^{-c^{1-\kappa}\log n} = (1-o(1))\,n^{-c^{1-\kappa}} = o(1/n)\,,$$

letting us take the union bound over all w.

To show that α and β tend to 1 as c increases, we can use the bound

$$f(\alpha) \geqslant \frac{(1-\alpha)^2}{3} \quad \text{for } 0 \leqslant \alpha \leqslant \mathrm{e}\,,$$

which the reader may recognize as a step in the derivation of the usual Chernoff bound (35). This implies that $f(\alpha), f(\beta) \geqslant \delta$ if

$$\alpha = 1 - \sqrt{3\delta} \quad \text{and} \quad \beta = 1 + \sqrt{3\delta}\,.$$

If c tends to infinity then $\delta = c^{-\kappa}$ tends to zero, and α and β tend to 1.

Finally, if $d = \omega(\log n)$ then the standard Chernoff bound suffices to prove (13) with $\alpha = \beta = 1$. □

7 Expansion: Proof of Lemmas 12 and 15

In this section we use expansion arguments to prove our claims about the shells $V_t(w)$ behind our lower bound.

We prove Lemma 12 by induction on t, using Lemma 11 as the base case $t = 1$. We rely on the fact that, for a fixed vertex w, if we have revealed the shells $V_1(w), \ldots, V_{t-1}(w)$ and the edges between them, then the edges between $V_{t-1}(w)$ and $V_{\geqslant t}(w)$ remain independent. Let us write V_t instead of $V_t(w)$ for simplicity. Then conditioned on $|V_1|, \ldots, |V_{t-1}|$ we have

$$|V_t| \sim \mathrm{Bin}(n', p')\,,$$

where

$$n' = |V_{\geqslant t}| = n - |V_{<t}| \quad \text{and} \quad p' = 1 - (1-p)^{|V_{t-1}|}\,, \tag{39}$$

since p' is the probability that at least one edge exists connecting V_{t-1} to a a given vertex in $V_{\geqslant t}$.

We will then use Chernoff bounds to show inductively that $|V_t|$ is close to its expectation, which is turn is close to $d|V_{t-1}|$. Specifically, we define $E_t(w)$ as the event that

$$\Pi_t^- d^{t-1} \deg w \leqslant |V_t(w)| \leqslant \Pi_t^+ d^{t-1} \deg w\,, \tag{40}$$

where $\Pi_t^{\pm} = 1 \pm o(1)$ are multiplicative errors determined below. For each w, the total probability that any step in the induction fails will be $o(1/n)$, so we can take the union bound over all $w \in V$.

Very similar reasoning appears in Lemma 2.1 of [3], which they prove for $d = \omega(\log n)$. In the case $d = O(\log n)$ which we include, the vertex degrees have significant fluctuations as in Lemma 11, but this only affects the base case of the induction where $|V_1(w)| = \deg w$. This base case will change slightly in Lemma 19, where we consider $V_t(u) \cup V_t(v)$ for a pair of vertices u, v.

Proof of Lemma 12. We start by fixing a vertex w and revealing $V_1(w)$. Importantly, Lemma 11 shows that $|V_1(w)| = \deg w = \Omega(\log n)$ with probability $1 - o(1/n)$, and we will condition on this event.

For each $1 \leqslant t \leqslant t^*$, let E_t denote the event (40) where, for some $\{\delta_t\}$ and $\{\epsilon_t\}$ to be determined below,

$$\Pi_t^- = \prod_{t'=2}^{t} (1-\delta_{t'})(1-\epsilon_{t'}) \quad \text{and} \quad \Pi_t^+ = \prod_{t'=2}^{t} (1+\epsilon_{t'})\,.$$

Of course, $\Pi_1^- = \Pi_1^+ = 1$. For simplicity, we write E_t and V_t instead of $E_t(w)$ and $V_t(w)$. We write $E_{\leqslant t} = \bigcap_{t'=1}^{t} E_{t'}$ and $E_{<t}$ similarly.

Our induction step will show that, for $2 \leqslant t \leqslant t^*$, E_t holds w.h.p. conditioned on $E_{<t}$. Since $|V_1(w)| = \deg w$, the base case E_1 holds trivially. Thus

$$\Pr[E_{\leqslant t^*}] = \Pr\left[\bigcap_{t=2}^{t^*} E_t\right] = \prod_{t=2}^{t^*} \Pr[E_t \mid E_{<t}]\,. \tag{41}$$

We will show that
$$\Pr[E_t \mid E_{<t}] = 1 - o(1/n^3)\,,$$
so we can take the union bound over all $t^* < \log n$ steps of the induction, as well as over all n vertices w.

As stated above, if we condition on $|V_{t'}|$ for all $t' < t$ then $|V_t|$ is a binomial random variable. We denote its expectation μ_t. There are two sources of multiplicative error that we need to track over the t^* steps of expansion. First, μ_t differs slightly from $d|V_{t-1}|$, and second, $|V_t|$ differs slightly from μ_t. The first source of error increases as t approaches t^*, and the second decreases as t and μ_t increase and the Chernoff bound becomes tighter. Happily, the product of all these error terms will turn out to be $1 \pm o(1)$.

Specifically, for each $2 \leqslant t \leqslant t^*$ we will show that, conditioned on $E_{<t}$, the expectation μ_t of $|V_t|$ is deterministically almost $d|V_{t-1}|$,

$$(1-\delta_t)\, d|V_{t-1}| \;\leqslant\; \mu_t \;\leqslant\; d|V_{t-1}|\,. \tag{42}$$

We will also show that, with probability $1 - o(1/n^3)$,

$$(1-\epsilon_t)\, \mu_t \;\leqslant\; |V_t| \;\leqslant\; (1+\epsilon_t)\, \mu_t\,. \tag{43}$$

Together these imply the induction step

$$(1-\delta_t)(1-\epsilon_t)\, d|V_{t-1}| \leqslant |V_t| \leqslant (1+\epsilon_t)\, d|V_{t-1}|\,, \tag{44}$$

and thus (taking products to get $\Pi_t^{\pm}$) establish E_t. Finally, the values of δ_t and ϵ_t that we will give below in Eqs. (47) and (48) decrease geometrically as we count backwards from t^* or forwards from $t = 2$ respectively, so that the products $\Pi_t^{\pm}$ are $1 \pm o(1)$.

To prove (42) and (43), let $t \leqslant t^*$ and suppose inductively that $E_{<t}$ holds. We have

$$|V_t| \sim \mathrm{Bin}(n', p') \quad \text{with expectation} \quad \mu_t = n'p'\,,$$

where

$$n' = n - |V_{<t}| \quad \text{and} \quad p' = 1 - (1-p)^{|V_{t-1}|}\,.$$

Clearly $\mu_t \leqslant d|V_{t-1}|$ since each vertex in V_{t-1} has d neighbors in expectation. To see that μ_t is not much smaller than this, note that since $E_{<t}$ holds $|V_{<t}|$ is a geometric sum dominated by its largest term, i.e.,

$$|V_{<t}| = O(|V_{t-1}|) = O(d^{t-1}) = O(d^{t^*} d^{-(t^*-t+1)}) = o(d^{-(t^*-t+1)} n)\,,$$

where we used $d^{t^*} = o(n)$. Thus

$$n' = \left(1 - o(d^{-(t^*-t+1)})\right) n\,. \tag{45}$$

For p', the error term comes from having more than one neighbor in V_{t-1}. Taking the Taylor series gives

$$\begin{aligned} p' &= \left(1 - O(p|V_{t-1}|)\right) p|V_{t-1}| \\ &= \left(1 - O(d^t/n)\right) p|V_{t-1}| \\ &= \left(1 - O\big((\gamma/d)\, d^{-(t^*-t)}\big)\right) p|V_{t-1}|\,, \end{aligned} \tag{46}$$

where we used $d^{t^*}/n = \gamma/d$. Combining (45) and (46) gives

$$\mu_t = n'p' = \left(1 - O\big((\gamma/d)\, d^{-(t^*-t)}\big)\right)\, d|V_{t-1}|\,.$$

This proves (42) with

$$\delta_t = B(\gamma/d)\, d^{-(t^*-t)} \tag{47}$$

for some constant $B > 0$.

To prove that (43) holds with probability $1 - o(1/n^3)$, we invoke the Chernoff bound in Lemma 22. Conditioning on $E_{<t}$ implies $\mu_t = \Theta(d^t) = \Omega(\log^t n)$, so (35) implies that (43) holds with probability $1 - o(1/n^3)$, where

$$\epsilon_t = A(\log n)^{-(t-1)/2} \tag{48}$$

for some constant $A > 0$.

Finally, we have

$$\begin{aligned} -\log \Pi_t^- &= -\sum_{t'=2}^{t} \log(1-\delta_{t'}) - \sum_{t'=2}^{t} \log(1-\epsilon_{t'}) \\ &= O\left(\sum_{t'=2}^{t} \delta_{t'} + \sum_{t'=2}^{t} \epsilon_{t'}\right)\,. \end{aligned}$$

Both these sums are geometric and are dominated by their largest term. Thus for any $2 \leqslant t \leqslant t^*$, recalling $\gamma/d = o(1)$ gives

$$-\log \Pi_t^- = O(\delta_t + \epsilon_2) = O\big((\gamma/d)\, d^{-(t^*-t)} + (\log n)^{-1/2}\big) = o(1)\,.$$

This implies $\Pi_t^- = 1 - o(1)$, and similarly for Π_t^+. This completes the proof of (14), which along with Lemma 11 yields (15). □

Proof of Lemma 15. We write V_t for $V_t(w)$ for simplicity. Recall that $V_{\leqslant t^*}$ is the union of all the shells whose sizes are upper and lower bounded by Lemma 12. Since $|V_t|$ grows by a factor of $\Theta(d)$ at each step, $|V_{\leqslant t^*}|$ is a geometric sum dominated by its largest term $|V_{t^*}|$. Thus

$$|V_{\leqslant t^*}| = \sum_{t=0}^{t^*} |V_t| = (1 + O(1/d))\, |V_{t^*}| = O(d^{t^*}) = o(n)\,. \tag{49}$$

Now let n' be the number of vertices outside these shells,

$$n' = |V_{>t^*}| = n - |V_{\leqslant t^*}| = (1 - o(1))n\,. \tag{50}$$

As in the proof of Lemma 12, conditioned on $|V_1|, \ldots, |V_{t^*}|$ we have

$$|V_{t^*+1}| \sim \mathrm{Bin}(n', p') \tag{51}$$

where p' is the probability a given vertex in $V_{>t^*}$ is a neighbor of at least one vertex in V_{t^*}. That is,

$$p' = 1 - (1-p)^{|V_{t^*}|} = 1 - \mathrm{e}^{-p|V_{t^*}|} + O(p^2 |V_{t^*}|)\,, \tag{52}$$

where we used the Taylor series. We can rewrite (15) for $t = t^*$ as

$$(1 - o(1))\, \alpha\gamma \leqslant p|V_{t^*}| \leqslant (1 + o(1))\, \beta\gamma\,. \tag{53}$$

Case 1. If $\gamma = \Theta(1)$, then Eqs. (50), (52), (53), and $p = o(1)$ tell us that

$$1 - \mathrm{e}^{-\alpha\gamma} - o(1) \leqslant p' \leqslant 1 - \mathrm{e}^{-\beta\gamma} + o(1)\,.$$

Since $|V_{t^*+1}|$ is binomial with expectation $\Omega(n)$, the Chernoff bound (Lemma 22) implies that it is within $O(n^{2/3})$ of its expectation $n'p' = (1-o(1))np'$ with probability $1-\mathrm{e}^{-\Omega(n^{1/3})}$. Thus w.h.p.

$$1 - \mathrm{e}^{-\alpha\gamma} - o(1) \leqslant \frac{|V_{t^*+1}|}{n} \leqslant 1 - \mathrm{e}^{-\beta\gamma} + o(1)\,, \tag{54}$$

so that V_{t^*+1} contains a positive fraction of the vertices. The probability that a vertex in the rest of the graph does not have a neighbor in V_{t^*+1} is then

$$(1-p)^{|V_{t^*+1}|} \leqslant \mathrm{e}^{-p|V_{t^*+1}|} = \mathrm{e}^{-\Omega(d)} = o(1)\,,$$

so w.h.p. $|V_{>t^*+2}| = o(n)$. Thus we have

$$\frac{|V_{t^*+1}|}{n} + \frac{|V_{t^*+2}|}{n} = 1 - \frac{|V_{\leqslant t^*}|}{n} - \frac{|V_{>t^*+2}|}{n} = 1 - o(1)\,,$$

which with (54) gives

$$\mathrm{e}^{-\beta\gamma} - o(1) \leqslant \frac{|V_{t^*+2}|}{n} \leqslant \mathrm{e}^{-\alpha\gamma} + o(1)\,.$$

The error probability is $o(1/n)$, so applying the union bound completes the proof for Case 1.

Case 2. If $\gamma = \omega(1)$, then (52) and (53) give

$$p' \geqslant 1 - \mathrm{e}^{-\alpha\gamma} = 1 - o(1)\,,$$

so V_{t^*+1} contains almost all the remaining n' vertices outside $V_{\leqslant t^*}$. The number of vertices in the remainder $V_{>t^*+1}$, which by Lemma 13 is w.h.p. just V_{t^*+2}, is distributed as $\mathrm{Bin}(n', 1-p')$ and has expectation at most

$$n'(1-p') \leqslant n(1-p') \leqslant n\mathrm{e}^{-\alpha\gamma}\,.$$

However, to apply the union bound over all w we need a bound on $|V_{t^*+2}|$ that holds with probability $1 - o(1/n)$. Since $\gamma = \omega(1)$, we have $\gamma/d = \omega(1/d) = \omega(1/\log^5 n)$. If $\mathrm{e}^{-\alpha\gamma} = \Omega(1/\log^5 n)$, then standard Chernoff bounds imply that $|V_{t^*+2}|/n \leqslant (1+o(1))\,\mathrm{e}^{-\alpha\gamma}$ with probability $1 - o(1/n)$. Alternatively, if $\mathrm{e}^{-\alpha\gamma} = o(1/\log^5 n)$, Chernoff bounds (Lemma 22) let us absorb $|V_{t^*+2}|/n$ into the error term $o(1)\,(\beta\gamma/d)$ in (19) with the same probability. Either way we have

$$\frac{|V_{\leqslant t^*}| + |V_{t^*+2}|}{n} \leqslant (1+o(1))\left(\frac{\beta\gamma}{d} + \mathrm{e}^{-\alpha\gamma}\right)\,,$$

having used Eq. 49. This completes the proof of (19). □

8 Separators: Proof of Lemmas 19 and 20

Proof of Lemma 19. Following [3], we first show that with high probability, for all distinct $u, v \in V$,

$$|V_1(u) \cap V_1(v)| \leqslant 2\,,$$

i.e., no two vertices have more than two common neighbors. This follows from the classic fact that sparse random graphs w.h.p. have no small subgraphs with more than one loop, i.e., with more edges than vertices. A pair u, v with three common neighbors forms a subgraph with 5 vertices and 6 edges, namely a copy of a complete bipartite graph $K_{2,3}$. The expected number of such pairs is

$$\binom{n}{2}\binom{n-2}{3}p^6 = O(n^5p^6) = O(d^6/n) = o(1)\,,$$

so w.h.p. no such pair exists. This implies w.h.p. for all distinct u, v

$$\deg u + \deg v - 2 \ \leqslant\ |V_1(u) \cup V_1(v)| \ \leqslant\ \deg u + \deg v\,,$$

and in particular (with Lemma 11)

$$|V_1(u) \cup V_1(v)| = (1 - o(1))(\deg u + \deg v)\,.$$

For each t, we can view $V_t(u) \cup V_t(v)$ as the tth shell around the set $\{u, v\}$. As in the proof of Lemma 12, if we condition on the sizes of the previous shells $|V_{t'}(u) \cup V_{t'}(v)|$ for $t' = 1, \ldots, t-1$, then $|V_t(u) \cup V_t(v)|$ is binomially distributed. *Mutatis mutandis*, the main fluctuations in the size of these shells are due to the first step, and w.h.p. each subsequent step expands by a factor close to d. Thus we have w.h.p. for all u, v,

$$\begin{aligned}|V_{t^*}(u) \cup V_{t^*}(v)| &= (1 \pm o(1))d^{t^*-1}|V_1(u) \cup V_1(v)| \\ &= (1 \pm o(1))\, d^{t^*-1}(\deg u + \deg v)\,.\end{aligned}$$

Along with Lemma 12 this implies that $V_{t^*}(u)$ and $V_{t^*}(v)$ have a small intersection, i.e.,

$$\begin{aligned}|V_{t^*}(u) \cap V_{t^*}(v)| &= |V_{t^*}(u)| + |V_{t^*}(v)| - |V_{t^*}(u) \cup V_{t^*}(v)| \\ &= o(1)\, d^{t^*-1}(\deg u + \deg v)\,,\end{aligned}$$

so the symmetric difference is almost as large as the union,

$$\begin{aligned}|\Delta_{t^*}(u, v)| &= |V_{t^*}(u) \cup V_{t^*}(v)| - |V_{t^*}(u) \cap V_{t^*}(v)| \\ &= (1 \pm o(1))\, d^{t^*-1}(\deg u + \deg v)\,.\end{aligned}$$

Finally, Lemma 11 gives $\deg u + \deg v \geqslant (2-o(1))\alpha d$, whereupon $d^{t^*} = \gamma n/d$ gives (27). $\square$

Proof of Lemma 20. A vertex w is in $\Delta_{t^*+1}(u, v)$ if it is outside $V_{\leqslant t^*}(u) \cup V_{\leqslant t^*}(v)$ and has an edge to either $V_{t^*}(u)$ or $V_{t^*}(v)$ but not both. Similar to our previous proofs, if we reveal the edges between vertices in $V_{\leqslant t^*}(u) \cup V_{\leqslant t^*}(v)$, the edges connecting $V_{t^*}(u) \cup V_{t^*}(v)$ to the rest of the graph are independent. Therefore, the events that $w \in \Delta_{t^*+1}$ for different w are independent as well. Thus if we condition on $|V_t(u)|$, $|V_t(v)|$, and $|V_t(u) \cap V_t(v)|$ for $1 \leqslant t \leqslant t^*$, we have

$$|\Delta_{t^*+1}(u, v)| \sim \mathrm{Bin}(n', p') \tag{55}$$

where

$$n' = n - |V_{\leqslant t^*}(u) \cup V_{\leqslant t^*}(v)| = (1 - o(1))\, n$$

and

$$\begin{aligned}p' = &\left(1 - (1-p)^{|V_{t^*}(u) \setminus V_{t^*}(v)|}\right)(1-p)^{|V_{t^*}(v)|} \\ &+ \left(1 - (1-p)^{|V_{t^*}(v) \setminus V_{t^*}(u)|}\right)(1-p)^{|V_{t^*}(u)|}\,.\end{aligned}$$

Lemma 19 implies that w.h.p.,

$$|V_{t^*}(u) \cap V_{t^*}(v)| = o(|V_{t^*}(u)|)\,,$$

and so

$$(1-p)^{|V_{t^*}(u) \setminus V_{t^*}(v)|} = (1-p)^{|V_{t^*}(u)|(1+o(1))} = (1-p)^{|V_{t^*}(u)|} + o(1)\,,$$

and similarly for the term with u and v swapped. As in the proof of Lemma 15, the Taylor series and (53) gives

$$
\begin{aligned}
p' &= (1-o(1))\left[\left(1-(1-p)^{|V_{t^*}(u)|}\right)(1-p)^{|V_{t^*}(v)|}\right. \\
&\qquad \left. + \left(1-(1-p)^{|V_{t^*}(v)|}\right)(1-p)^{|V_{t^*}(u)|}\right] \\
&= (1-o(1))\left[\left(1-\mathrm{e}^{-p|V_{t^*}(u)|}\right)\mathrm{e}^{-p|V_{t^*}(v)|} + \left(1-\mathrm{e}^{-p|V_{t^*}(v)|}\right)\mathrm{e}^{-p|V_{t^*}(u)|}\right] \\
&\geqslant (2-o(1))\left(1-\mathrm{e}^{-\alpha\gamma}\right)\mathrm{e}^{-\beta\gamma}\,,
\end{aligned}
\tag{56}
$$

where we pessimistically minimized the two terms in the penultimate line separately. Thus the conditional expectation of $|\Delta_{t^*+1}|$ is bounded below by

$$\nu := n'p' \geqslant (2-o(1))\left(1-\mathrm{e}^{-\alpha\gamma}\right)\mathrm{e}^{-\beta\gamma}n\,.$$

Case 1. If $\gamma = \Theta(1)$, then $\nu = \Theta(n)$ and Chernoff bounds imply (28).

Case 2. Since $\gamma = \omega(1)$ we have $\mathrm{e}^{-\alpha\gamma} = o(1)$, so we simply write $\nu = (2-o(1))\,\mathrm{e}^{-\beta\gamma}n$.

Now the first term of (29) corresponds to Δ_{t^*} and comes from Lemma 19. For the second term involving Δ_{t^*+1}, we have a situation similar to Case 2 of the lower bound. If $\nu \geqslant n/\log^{10} n$, then Chernoff bounds imply that $|\Delta_{t^*+1}| \geqslant (1-o(1))\,\nu$ with probability $1-o(1/n^2)$. Alternatively, if $\nu < n/\log^{10} n$, Chernoff bounds give

$$|\Delta_{t^*+1}| \leqslant \nu \log^5 n < n/\log^5 n < n/d = o(n\gamma/d)\,,$$

with probability $1-o(1/n^2)$. In that case, we absorb $|\Delta_{t^*+1}|/n$ into the error term of $|\Delta_{t^*}|/n = (2-o(1))\,\alpha\gamma/d$.

In either situation we can take the union bound over all $u, v \in V$, completing the proof of (29). □

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation, project PID2020-112581GB-C21 MOTION (J. Díaz), the Santa Fe Institute (H. Hartle, C. Moore), and the Robert Wood Johnson Foundation (C. Moore).

References

[1] Z. Beerliova, F. Eberhard, T. Erlebach, A. Hall, M. Hoffmann, M. Mihalák, and L. Ram, Network discovery and verification. *IEEE J. Sel. Area Comm.* **24**, 2168–2181, 2006.

[2] R. Belmonte, F. Fomin, P. Golovach, and M. Ramanujan, Metric dimension of bounded width graphs. In *40th Math. Foundations Comp. Science*, Lecture Notes in Computer Science, vol. 9235, 115–126, Springer, 2015.

[3] B. Bollobás and D. Mitsche and P. Prałat, Metric dimension for random graphs. *Electronic Journal of Combinatorics* **20**, 4, #P1, 2013.

[4] B. Bollobás, The diameter of random graphs. *Transactions of the American Mathematical Society* **267**, 41–52, 1981.

[5] É. Bonnet and N. Purohit, Metric dimension parameterized by treewidth. In *14th Parameterized and Exact Computation* Leibniz International Proceedings in Informatics (LIPIcs), vol. 148, pp. 5:1–5:15, Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[6] G. Chartrand, L. Eroh, M.A. Johnson, and O.R. Oellemann, Resolvability in graphs and the metric dimension of a graph. *Discrete Applied Mathematics*, **105**, 99–113, 2000.

[7] F. Chung and L. Lu, The diameter of sparse random graphs. *Advances in Applied Mathematics* **26**, 257–279, 2001.

[8] T. Cover and J. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1999.

[9] J. Díaz, O. Pottonen, M. Serna, and E. van Leeuwen, On the complexity of metric dimension, *Journal of Computer and System Sciences* **83**, 1, 132–158, 2017.

[10] L. Epstein, A. Levin, and G. Woeginger, The (weighted) metric dimension of graphs: hard and easy cases. In *38th International Workshop on Graph-Theoretic Concepts in Computer Science, 2012*, Lecture Notes in Computer Science, vol. 7551, 114–125, Springer, 2012.

[11] H. Fernau, P. Heggernes, P. van't Hof, D. Meister, and R. Saei, Computing the metric dimension for chain graphs. *Information Processing Letters* **115**, 9, 671–676, 2015.

[12] F. Foucaud, G. Mertzios, R. Naserasr, A. Parreau, and P. Valicov, Identification, location-domination and metric dimension on interval and permutation graphs: algorithms and complexity, *Algorithmica* **78**, 3, 914–944, 2017.

[13] M.R. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, 1979.

[14] F. Harary, R.A. Melter, On the metric dimension of a graph, *Ars Combinatoria* **2**, 191–195, 1976.

[15] M. Hauptmann, R. Schmied, and C. Viehmann, Approximation complexity of metric dimension problem, *Journal of Discrete Algorithms.* **14**, 214–222, 2012.

[16] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs.* John Wiley & Sons, 2011.

[17] S. Khuller, B. Raghavachari, and A. Rosenfeld, Landmarks in graphs. *Discrete Applied Mathematics* **70**, 217–229, 1996.

[18] J. Komjáthy, and G. Ódor, Metric dimension of critical Galton-Watson trees and linear preferential attachment trees. *European Journal of Combinatorics* **95**, 103317, 2021.

[19] C. Moore and S. Mertens, *The Nature of Computation.* Oxford University Press, 2011.

[20] G. Ódor and P. Thiran, Sequential metric dimension for random graphs. *Journal of Applied Probability.* **58**, 4, 909–951, 2021.

[21] G. Ódor, *The role of adaptivity in source identification with time queries.* Thèse de doctorat en informatique et communications, n. 9083, EPFL, 2022.

[22] J. Radhakrishnan, *Entropy and counting*, Computational Mathematics, Modelling and Algorithms (Ed. J. C. Mishra), Narosa Publishers, 2001.

[23] O. Riordan and N. Wormald, The diameter of sparse random graphs. *Combinatorics, Probability and Computing* **19**, 835–926, 2010.

[24] C. Shannon, A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423, 1948.

[25] P. Slater, Leaves of trees, *Congressus Numerantium* **14**, 549–559, 1975.