

The Diameter of Sparse Random Graphs*

Daniel Fernholz, Vijaya Ramachandran

Department of Computer Sciences, University of Texas at Austin; Texas;

e-mail: fernholz@cs.utexas.edu; vlr@cs.utexas.edu

Received 13 December 2004; revised 29 August 2006; accepted 9 May 2007

Published online 3 October 2007 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/rsa.20197

ABSTRACT: We derive an expression of the form $c \ln n + o(\ln n)$ for the diameter of a sparse random graph with a specified degree sequence. The result holds asymptotically almost surely, assuming that certain convergence and supercriticality conditions are met, and is applicable to the classical random graph $G_{n,p}$ with $np = \Theta(1) + 1$, as well as certain random power law graphs. © 2007 Wiley Periodicals, Inc. *Random Struct. Alg.*, 31, 482–516, 2007

Keywords: random graph; diameter; degree sequence; configuration model

1. INTRODUCTION

The *diameter* of a graph is the maximum distance between two vertices in the same component. We derive an expression of the form

$$c \ln n + o(\ln n)$$

for the diameter of a sparse random graph with a specified degree sequence. Our result holds asymptotically almost surely (AAS) for any degree distribution that asymptotically satisfies certain sparseness and convergence conditions similar to those imposed by [17, 18] in the analysis of the size and degree distribution of the largest connected component of such a random graph, as well as a supercriticality condition which guarantees that a giant component — that is, a connected component of size $\Theta(n)$ — is AAS present. This includes sparse “power

Correspondence to: Daniel Fernholz

*Supported by NSF (CCR-9988160, CCF-0514876).

© 2007 Wiley Periodicals, Inc.

law” distributions which have received considerable recent attention (e.g. [1]), as well as the Poisson distribution, which corresponds to the random graph $G_{n,p}$ with constant expected degree.

Diameter results of such precision for sparse random graphs were known earlier only for regular graphs of constant degree [6]. Weaker results (to within a constant factor) were known for the diameter of sparse $G_{n,p}$ [8] and random “expected-degree” power law graphs [16]. A recent paper on inhomogeneous random graphs [7] includes diameter results which generalize the $G_{n,p}$ result of the present paper in a different direction. A slightly different derivation of diameter results in this paper, with a stronger probabilistic guarantee ($1 - n^{\Omega(1)}$ rather than $1 - o(1)$) can be found in the first author’s PhD thesis [11].

For the random graph $G_{n,p}$ with expected degree $d = np > 1$, we derive an expression for the leading constant c as a function of d . This function can be characterized asymptotically by $c = \frac{3-o(1)}{\ln d}$ as $d \rightarrow 1$ and $c = \frac{1}{\ln d} + \frac{2}{d} + O\left(\frac{\ln d}{d^2}\right)$ as $d \rightarrow \infty$. This resolves the open problem, posed in [8], of whether the diameter of $G_{n,d/n}$ satisfies $\frac{\ln n}{\ln d}(1 + o(1))$. We also show that, for all $d > 1$, the diameter of $G_{n,d/n}$ is AAS equal to the diameter of its giant component, thus resolving another open problem from [8].

The proof of the diameter result involves tracing the rate of neighborhood expansion during a breadth-first search (BFS); as such, our argument bears some resemblance to the arguments used in previous diameter papers including [8] and [16]. However, rather than perform BFS on the entire graph, we restrict our search to edges in the 2-core (defined as the maximal induced subgraph of minimum degree at least 2), as well as any edges connecting a particular vertex to the 2-core. Our argument thus relies heavily on results about the size and degree distribution of the 2-core of a random graph, which have been derived by several authors including [9, 12, 14].

Performing BFS in the 2-core is advantageous in that each vertex has degree at least 2, and therefore there are no “dead ends.” It follows that neighborhood sizes during 2-core BFS are generally nondecreasing, and this property facilitates the analysis of the neighborhood growth, particularly for small neighborhoods. As a result, we are able to attain relatively simple and tight bounds on worst-case neighborhood growth, which in turn lead an asymptotic expression of the form $c \ln n + o(\ln n)$ for the diameter of a random graph.

Our main theorem (Theorem 5.1) yields a method for computing the leading constant c in the above expression, as well as an intuitive interpretation of the resulting formula. Briefly, the leading constant can be expressed as $c = 2c_1 + c_2$, where $c_2 \ln n$ is the “average distance” between vertices in the giant component (or in the giant connected 2-core), while $c_1 \ln n$ is the maximum distance separating any vertex in the graph from the 2-core. A typical longest shortest path connecting vertices u, v will consist of a path of length $c_1 \ln n$ from u to the 2-core, a path of “average” length $c_2 \ln n$ across the 2-core, and a second path of length $c_1 \ln n$ from the 2-core to v .

This paper is organized as follows. In section 2 we present basic definitions and notation, and we derive some preliminary lemmas. In section 3 we analyze a breadth-first search on a random graph. In section 4 we recall previous results regarding the 2-core of a random graph, and we analyze a breadth-first search in the 2-core. Then, in section 5, we use the results in two previous sections to prove our main result. In section 6, we compute the diameter for the Poisson distribution, which corresponds to the random graph $G_{n,p}$ and for a certain power law distribution.

2. PRELIMINARIES

In this section, we introduce basic definitions and notation, and we present some preliminary results. The lemmas presented in this section are fairly simple, and as such the proofs are discussed in the appendix.

2.1. Probability

In this section, we introduce notation and basic definitions to describe probabilistic constructions. We adopt the convention that random variables and structures are denoted in boldface. We denote the distribution of a random element \mathbf{s} in a discrete set S by

$$\mathfrak{D}[\mathbf{s}](s) = \mathbb{P}[\mathbf{s} = s]$$

for $s \in S$. If two random elements \mathbf{s}, \mathbf{s}' are equal in distribution, we write $\mathbf{s} \stackrel{d}{=} \mathbf{s}'$.

Given a probability distribution μ on \mathbb{R} , we denote a generic random variable with distribution μ by \mathbf{X}_μ . Generic random variables are all considered independent, and as such we denote the sum of n independent μ -distributed random variables by $\sum_{i=1}^n \mathbf{X}_{\mu,i}$.

We denote the binomial distribution with parameters n, p by $\text{Bin}_{n,p}(j) = \binom{n}{j} p^j (1-p)^{n-j}$.

This paper deals exclusively with discrete probability spaces which are measure-theoretically trivial. Nevertheless, we shall use several constructs of probability theory, specifically σ -fields and filtrations, to deal with states of partial information; definitions of these concepts can be found in any standard probability text (e.g. [3, 15]).

We assume that every random element \mathbf{s} in a set S is a measurable mapping from an underlying probability space $\Omega \rightarrow S$. We let $\sigma[\mathbf{s}]$ denote the σ -field induced by \mathbf{s} , and we let $\sigma[\emptyset] = \{\Omega, \emptyset\}$ denote the trivial σ -field. If \mathcal{F} is a σ -field on Ω and \mathbf{s} is a random element, we denote by $\mathfrak{D}[\mathbf{s}|\mathcal{F}]$ the distribution of \mathbf{s} conditional on \mathcal{F} , noting that $\mathfrak{D}[\mathbf{s}|\mathcal{F}]$ is an \mathcal{F} -measurable random probability distribution on S .

2.2. The Configuration Model

We shall generate fixed-degree-sequence random graphs using the *configuration model* [4]. Our notation is as follows. We begin with an (even) set of *endpoints* A partitioned into a set of *vertices* V . A *configuration* on (A, V) is a triple $G = (A, V, E)$, where the edge set E is a perfect matching of the endpoint set A . We let $m = |A|$ denote the number of endpoints and $n = |V|$ the number of vertices; note that the number of edges in the resulting graph is $m/2$.

We now introduce notation to describe a configuration $G = (A, V, E)$.

- For a vertex $v \in V$, $A(v)$ denotes the set of endpoints which belong to v .
- For an endpoint $a \in A$, $V(a)$ denotes the vertex which contains a .
- For an endpoint $a \in A$, $\vec{E}(a)$ denotes the endpoint matched to a , so the pair $\{a, \vec{E}(a)\}$ occurs as an edge in E .
- For a vertex set $V' \subseteq V$, $A(V')$ denotes the set of all endpoints belonging to vertices in V' , so $A(V') = \cup_{v \in V'} \{A(v)\}$.
- For an endpoint set $A' \subseteq A$, we similarly write $V(A') = \cup_{a \in A'} \{V(a)\}$ and $\vec{E}(A') = \cup_{a \in A'} \{\vec{E}(a)\}$, noting that the union which defines $V(A')$ is not necessarily disjoint.

Now, given an endpoint partition (A, V) , we define the random configuration

$$\mathbf{G}(A, V) = (A, V, \mathbf{E}),$$

where \mathbf{E} is a uniformly random matching of the set A .

We note that the configuration model does not necessarily produce a simple graph. If we are specifically interested in random simple graphs, it is not difficult to argue that results achieved using the configuration model are still valid so long as some minor technical conditions are satisfied (see e.g., [4, 5]). In this paper, we ignore these concerns and focus on the random configuration $\mathbf{G}(A, V)$. Moreover, from this point forward, we shall use the terms “graph” and “configuration” interchangeably, and as such we may simply refer to $\mathbf{G}(A, V)$ as a “random graph.”

2.2.1. The Configuration Model Algorithm. We adopt an algorithmic approach to the configuration model similar to [13, 17]. The algorithmic approach is based on a simple observation.

Let E be a matching of A , and consider a subset $A_0 \subseteq A$. We define the *restricted matching* $E\langle A_0 \rangle$ to be the subset of E consisting of all edges $\{a, a'\} \in E$ such that either $a \in A_0$ or $a' \in A_0$. The following proposition is immediate.

Proposition 2.1. *Let \mathbf{E} be a uniformly random matching of an endpoint set A , and choose any $a \in A$. Then $\bar{\mathbf{E}}(a)$ is a uniformly random element in $A \setminus \{a\}$, and conditional on $\bar{\mathbf{E}}(a)$, the restricted matching $\mathbf{E}\langle A \setminus \{a, \bar{\mathbf{E}}(a)\} \rangle$ is a uniformly random matching of the set $A \setminus \{a, \bar{\mathbf{E}}(a)\}$.*

This proposition allows us to observe selected parts of a random configuration $\mathbf{G}(A, V) = (A, V, \mathbf{E})$ while maintaining the conditional uniform randomness of the parts we have not seen. In this section, we present notation for generating random configurations incrementally.

Let $\mathcal{F} \subseteq \sigma[\mathbf{E}]$ be a sub- σ -field of the σ -field induced by \mathbf{E} . Intuitively, \mathcal{F} represents a state of partial knowledge about \mathbf{E} . We now define the \mathcal{F} -measurable random partial matching

$$\mathbf{E}_{\mathcal{F}} = \{\{a, a'\} \subset A : \mathbb{P}[\{a, a'\} \in \mathbf{E} \mid \mathcal{F}] = 1\}.$$

Informally, $\mathbf{E}_{\mathcal{F}}$ consists of the set of edges which are known to occur in \mathbf{E} based on the information in \mathcal{F} . We say an edge in $\mathbf{E}_{\mathcal{F}}$ is \mathcal{F} -exposed; accordingly, we call any endpoint belonging to an \mathcal{F} -exposed edge an \mathcal{F} -exposed endpoint. We denote by $\mathbf{U}_{\mathcal{F}}$ the set of \mathcal{F} -unexposed endpoints, noting that $\mathbf{U}_{\mathcal{F}}$ is an \mathcal{F} -measurable random subset of A .

We say a sub- σ -field $\mathcal{F} \subseteq \sigma[\mathbf{E}]$ is *sub-uniform* if the conditional distribution $\mathfrak{D}[\mathbf{E}\langle \mathbf{U}_{\mathcal{F}} \rangle \mid \mathcal{F}]$ is always uniform on the set of matchings of $\mathbf{U}_{\mathcal{F}}$. Accordingly, a *sub-uniform filtration* is a filtration in which each σ -field is sub-uniform. It follows immediately from Proposition 2.1 that, for any subset $A_0 \subseteq A$, the induced σ -field $\sigma[\mathbf{E}\langle A_0 \rangle]$ is sub-uniform. By using this same technique repeatedly, we have the following proposition.

Proposition 2.2. *Consider a sequence of σ -fields $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \dots, \mathcal{F}_t$ such that for each $0 \leq i \leq t - 1$, there exists an \mathcal{F}_i -measurable random subset $\mathbf{B}_i \subseteq A$ such that $\mathcal{F}_{i+1} = \sigma[\mathcal{F}_i, \mathbf{E}\langle \mathbf{B}_i \rangle]$.*

Then $\mathcal{F}_0, \dots, \mathcal{F}_t$ is a sub-uniform filtration.

It is useful to think of sub-uniform filtrations constructed according to Proposition 2.2 algorithmically in the following sense. At time 0, we have no knowledge of the matching \mathbf{E}

whatsoever, as represented by the σ -field $\mathcal{F}_0 = \sigma[\emptyset]$. We now choose an \mathcal{F}_0 -measurable (i.e., non-random) subset $\mathbf{B}_0 \subseteq A$, and we determine the (random) matches of all endpoints in \mathbf{B}_0 . The σ -field $\mathcal{F}_1 = \sigma[\mathbf{E}(\mathbf{B}_0)]$ thus represents our knowledge about \mathbf{E} after these endpoints have been exposed. Then, we choose a second set of endpoints \mathbf{B}_1 , possibly using some information provided by \mathcal{F}_1 ; hence, the set \mathbf{B}_1 may be random, since it may depend on the restricted matching $\mathbf{E}(\mathbf{B}_0)$. We then let $\mathcal{F}_2 = \sigma[\mathcal{F}_1, \mathbf{E}(\mathbf{B}_1)]$ represent our newly updated state of knowledge, and so on.

2.3. Asymptotic Parametrization and the Degree Distribution

In this section, we present notation and definitions and preliminary results related to the asymptotic specification of the configuration model.

This paper studies large random graphs asymptotically as the number of vertices $n \rightarrow \infty$. Hence, we are in effect considering a sequence of random graphs $\mathbf{G}(A_1, V_1), \mathbf{G}(A_2, V_2), \dots$ where (A_n, V_n) is an endpoint partition with $|V_n| = n$. The asymptotic subscript “ n ” will generally remain implicit; in this section, we make the subscript explicit to give precise asymptotic definitions.

We shall use the usual “big O ” asymptotic notation, including the standard notation $\tilde{O}(f(n)) = O(f(n) \ln^{O(1)} n)$. All asymptotic notation is assumed to refer to the limit as $|V_n| = n \rightarrow \infty$ unless explicitly stated otherwise.

We shall reserve the symbols n and m for the number of vertices $n = |V|$ and endpoints $m = |A|$ in an endpoint partition (A, V) . We shall deal exclusively with sparse graphs, meaning that $m = \Theta(n)$. In particular, this means that the symbols “ m ” and “ n ” are interchangeable in many asymptotic expressions, for example $O(n) \equiv O(m)$ and $(\ln n) \cdot (1 + o(1)) \equiv (\ln m) \cdot (1 + o(1))$.

2.3.1. Asymptotics and Probability. Given a sequence of events H_1, H_2, \dots , we say H_n occurs AAS if $\mathbb{P}[H_n] = 1 - o(1)$, and H occurs with high probability (WHP), if $\mathbb{P}[H_n] = 1 - n^{-\omega(1)}$. Note that the conjunction of a constant number of AAS events occurs AAS, and the conjunction of a polynomial number of WHP events occurs WHP.

The AAS and WHP conditions can be used in concert with asymptotic notation in the natural way; for example, the statement $f(n) = o(n)$ AAS (WHP) indicates that for all $\epsilon > 0$, the event $f(n) < \epsilon n$ occurs AAS (WHP).

2.3.2. The Degree Distribution. In this section we describe our method for asymptotic specification of the configuration model. Our asymptotic assumptions are essentially equivalent to those of Molloy and Reed [17, 18].

Given a vertex v in an endpoint partition (A, V) , the *degree* $\deg_{(A,V)}(v) = |A(v)|$ is the number of endpoints contained in v . Accordingly, the *degree distribution*

$$\lambda_{(A,V)}(i) = \frac{|\{v \in V : \deg_{(A,V)}(v) = i\}|}{n}$$

is the probability distribution corresponding to the degree of a randomly chosen vertex.

We shall specify the configuration model asymptotically using convergence of the degree distribution; we require convergence in L_1 as well as convergence of either one or two moments. Given distributions λ, ν on a discrete set S , the L_1 distance is defined as usual

by $\|\lambda - \nu\|_1 = \sum_{s \in S} |\lambda(s) - \nu(s)|$. For any distribution λ on \mathbb{Z} , we denote the k 'th moment of λ by $M_k(\lambda) = E[\mathbf{X}_\lambda^k]$. Our asymptotic assumptions will take the following form.

Assumptions 2.1. *We are given a sequence $(A_1, V_1), (A_2, V_2), \dots$ of (even) endpoint partitions with $|V_n| = n$ and a limiting degree distribution λ , which satisfies $0 < M_1(\lambda) < \infty$, and such that*

1. $\lim_{n \rightarrow \infty} \|\lambda_{(A_n, V_n)} - \lambda\|_1 = 0$;
2. $\lim_{n \rightarrow \infty} M_1(\lambda_{(A_n, V_n)}) = M_1(\lambda)$;
3. $\lim_{n \rightarrow \infty} M_2(\lambda_{(A_n, V_n)}) = M_2(\lambda)$.

Note that we only require that the limiting distribution have one finite moment; however, if in fact $M_2(\lambda) < \infty$ then the second moment $M_2(\lambda_{(A_n, V_n)})$ must converge to $M_2(\lambda)$.

Our main theorem, proved in Section 5, relies on these asymptotic assumptions. However, the technical lemmas proved in Sections 3 and 4 are stated in the more explicit form “for all $\epsilon > 0$ there exists $\delta > 0 \dots$ ” to precisely specify dependencies among asymptotic constants.

2.4. The Residual Distribution

Our method of asymptotic parametrization, described above, is based on the degree distribution. However, our technical work relies mainly on a second, related distribution, namely the *residual distribution*. In this section, we define the residual distribution, and we derive some preliminary results regarding its sensitivity to small changes in an endpoint partition.

Given an endpoint partition (A, V) and an endpoint $a \in A$, the *residual degree* $\text{res}_{(A, V)}(a)$ is the number of other endpoints which belong to the same vertex as a . Formally, we have

$$\text{res}_{(A, V)}(a) = |A(V(a)) \setminus \{a\}| = \text{deg}_{(A, V)} V(a) - 1.$$

Accordingly, we define the *residual distribution* by $\mu_{(A, V)}(i) = \frac{| \{a \in A : \text{res}_{(A, V)}(a) = i \} |}{m}$. The residual distribution can be computed easily from the degree distribution, and in general, for any distribution λ on $\mathbb{Z}^* = \{0, 1, \dots\}$ with $0 < M_1(\lambda) < \infty$, we define the corresponding residual distribution by

$$\mu_\lambda(i) = \frac{(i + 1)\lambda(i + 1)}{M_1(\lambda)}. \tag{1}$$

The effects of our asymptotic assumptions on the residual distribution are given in the following proposition, which follows immediately from definitions.

Proposition 2.3. *Assumptions 2.1 have the following implications:*

1. $\lim_{n \rightarrow \infty} \|\mu_{(A_n, V_n)} - \mu_\lambda\|_1 = 0$, where μ_λ is computed from the limiting degree distribution λ according to Eq. (1).
2. $\lim_{n \rightarrow \infty} M_1(\mu_{(A_n, V_n)}) = M_1(\mu_\lambda)$.

We call the distribution μ_λ the *limiting residual distribution*. Note that, similarly to assumptions 2.1, it is not necessarily the case the $M_1(\mu_\lambda) < \infty$, but if the first moment is finite, then the convergence $M_1(\mu_{(A_n, V_n)}) \rightarrow M_1(\mu_\lambda)$ holds.

2.4.1. Sensitivity of the Residual Distribution We now present a technical lemma which states that, under our asymptotic assumptions, a vertex set of size $o(m)$ can only contain $o(m)$ endpoints.

Lemma 2.4. *Let μ be any distribution on \mathbb{Z}^* . Then for all $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

For any endpoint partition (A, V) such that the residual distribution $\mu_{(A,V)}$ satisfies $\|\mu_{(A,V)} - \mu\|_1 < \delta$, and any subset $V' \subseteq V$ satisfying $|V'| < \delta|A|$, we have $|A(V')| < \epsilon|A|$.

The proof of this lemma appears in the Appendix. We have the following immediate corollary, which essentially states that removing $o(m)$ endpoints changes the residual distribution by an L_1 distance of at most $o(1)$.

Corollary 2.5. *Let μ be any distribution on \mathbb{Z}^* . Then for all $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

For any endpoint partition (A, V) satisfying $\|\mu_{(A,V)} - \mu\|_1 < \delta$, and for any subset $A' \subseteq A$ satisfying $|A| - |A'| < \delta|A|$, we have

$$\|\mu_{(A',V)} - \mu\|_1 < \epsilon.$$

2.5. Probabilistic Tools

In our analysis of the random graph $\mathbf{G}(A, V)$, we will encounter certain technical difficulties which stem from two principal sources. First, our method of asymptotic parametrization does not provide precise knowledge of any particular endpoint partition (A, V) ; we can only assume that, for n sufficiently large, the degree distribution $\lambda_{(A,V)}$ lies in an arbitrarily small neighborhood of the limiting distribution λ . Second, the occurrences of edges in a random matching are not independent events, and as a result, we will frequently be confronted with sequences of random variables which are “not quite” independent and identically distributed.

Our general strategy for dealing with these difficulties will be to relate certain “messy” situations to “clean” probabilistic constructions, and then to derive bounds using well-known classical results. The material presented below will be used toward this end.

2.5.1. Domination and the Truncated Distribution. Given distributions μ, ν on \mathbb{Z} , we say μ dominates ν and write $\mu \triangleright \nu$ if $\mathbb{P}[\mathbf{X}_\mu \geq i] \geq \mathbb{P}[\mathbf{X}_\nu \geq i]$ for all i , where, as defined in Section 2.1, \mathbf{X}_μ and \mathbf{X}_ν are μ - and ν -distributed random variables, respectively. Accordingly, for random variables \mathbf{X}, \mathbf{Y} , we write $\mathbf{X} \stackrel{d}{\triangleright} \mathbf{Y}$ if $\mathfrak{D}[\mathbf{X}] \triangleright \mathfrak{D}[\mathbf{Y}]$ and say \mathbf{X} dominates \mathbf{Y} in distribution.

We shall frequently encounter distributions with support in the set $\mathbb{Z}^* = \{0, 1, \dots\}$ (e.g., the degree and residual distributions defined above). For technical reasons, we shall often consider these as distributions on the set $\mathbb{Z}^* \cup \{-1\}$. Our next lemma establishes that the closed ϵ -neighborhood of any such distribution μ (with respect to L_1 distance) contains a lower bound $\mu_{[\epsilon]}$ with respect to the partial order \triangleright . We call this lower bound the ϵ -truncated distribution.

Lemma 2.6. *For any distribution μ on $\mathbb{Z}^* \cup \{-1\}$ and any $\epsilon > 0$, there exists a unique distribution $\mu_{[\epsilon]}$ on $\mathbb{Z}^* \cup \{-1\}$ satisfying both $\|\mu - \mu_{[\epsilon]}\|_1 \leq \epsilon$ and $\nu \triangleright \mu_{[\epsilon]}$ for all distributions ν on $\mathbb{Z}^* \cup \{-1\}$ with $\|\mu - \nu\|_1 \leq \epsilon$.*

The proof of this lemma (which appears in the Appendix) is quite simple; briefly, $\mu_{[\epsilon]}$ is constructed by taking an amount $\epsilon/2$ of probability weight from the top of the distribution and placing this weight on the value -1 .

2.5.2. Applications. We now present three lemmas which apply the concepts defined above to achieve probabilistic bounds in specific situations; the proofs can be found in the Appendix.

The first lemma deals with a random process in which the conditional distributions of the increments always dominate a fixed distribution μ .

Lemma 2.7. *Let y_1, \dots, y_t be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \dots, \mathcal{F}_t$, and let $\mathbf{Y}_t = \sum_{i=1}^t y_i$. Consider a distribution μ such that $\mathcal{D}[y_{s+1} \mid \mathcal{F}_s] \triangleright \mu$ always for all $0 \leq s \leq t - 1$. Then \mathbf{Y}_t dominates in distribution the sum of t independent μ -distributed random variables.*

The next lemma considers a similar situation for a sequence of Bernoulli trials, but where the lower bound on the conditional distribution only holds before a given stopping time τ .

Lemma 2.8. *Consider a sequence of Bernoulli random variables z_1, z_2, \dots adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \mathcal{F}_2, \dots$, and let $\mathbf{Z}_t = \sum_{i=1}^t z_i$. Consider a stopping time τ and a constant $0 < p < 1$ such that*

$$P[\mathbf{z}_{t+1} = 1 \mid t < \tau, \mathcal{F}_t] \geq p$$

always for all $t \geq 0$. Then for any $t, r \geq 0$, we have

$$P[\tau > t \wedge \mathbf{Z}_t \leq r] \leq P[\mathbf{X}_{\text{Bin}_t, p} \leq r].$$

The final lemma in this section deals with the sum of \mathbf{Y} independent, identically distributed random variables, where \mathbf{Y} is itself a random variable, and such that the distribution $\mathcal{D}[\mathbf{Y}]$ dominates a binomial distribution.

Lemma 2.9. *Let μ be a distribution on $\mathbb{Z}^* \cup \{-1\}$, let $r \in \mathbb{Z}^*$ and $0 \leq \epsilon \leq 1$, and let \mathbf{Y} be a \mathbb{Z} -valued random variable such $0 \leq \mathbf{Y} \leq r$ always and such that $\mathbf{Y} \stackrel{d}{\triangleright} \mathbf{X}_{\text{Bin}_r, 1-\epsilon/2}$. Then*

$$\sum_{i=1}^{\mathbf{Y}} \mathbf{X}_{\mu, i} \stackrel{d}{\triangleright} \sum_{j=1}^r \mathbf{X}_{\mu_{[\epsilon]}, j}.$$

3. BREADTH-FIRST SEARCH ON A RANDOM GRAPH

In this section we analyze a breadth-first search on a random graph. Our notation is as follows. Let (A, V, E) be a configuration, and choose any $v \in V$. For any integer $i \geq 0$, we let $N_{v,i}$ denote the set of vertices at distance exactly i from v ; the set $N_{v,i}$ is the i 'th vertex neighborhood of v .

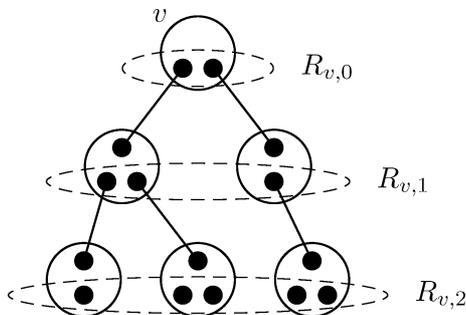
It is often more useful to consider neighborhoods as consisting of endpoints rather than vertices. Hence, we define the i 'th endpoint neighborhood $R_{v,i}$ to be the set of endpoints

belonging to vertices in $N_{v,i}$, and which are not matched to endpoints in $R_{v,i-1}$. Formally, then, we have

$$R_{v,0} = A(v)$$

$$R_{v,i} = \{a \in A(N_{v,i}) : \vec{E}(a) \notin R_{v,i-1}\}.$$

The endpoint neighborhoods of a vertex v are illustrated below.



In a random configuration $\mathbf{G}(A, V) = (A, V, \mathbf{E})$, the endpoint neighborhoods are random subsets of A , which according to our convention we denote in boldface by $\mathbf{R}_{v,i}$. For any v , and any i , we now define the σ -field $\mathcal{B}_{v,i}$ inductively by $\mathcal{B}_{v,0} = \sigma[\emptyset]$ and

$$\mathcal{B}_{v,i+1} = \sigma[\mathcal{B}_{v,i}, \mathbf{E}(\mathbf{R}_i)].$$

We call the filtration $\mathcal{B}_{v,0}, \mathcal{B}_{v,1}, \dots$, the *breadth-first search (BFS) filtration*. We note that, since $\mathbf{R}_{v,i+1}$ is clearly $\mathcal{B}_{v,i}$ -measurable, then by proposition 2.2 the BFS filtration is sub-uniform.

This section is devoted to analyzing the BFS filtration. To simplify our notation, we abbreviate the set of unexposed endpoints and exposed edges in the BFS filtration by $\mathbf{U}_{v,i} = \mathbf{U}_{\mathcal{B}_{v,i}}$ and $\mathbf{E}_{v,i} = \mathbf{E}_{\mathcal{B}_{v,i}}$, respectively. Intuitively, the set $\mathbf{E}_{v,i}$ consists of all edges explored during i iterations of a standard breadth-first search beginning at v . We note that the endpoints in $\mathbf{R}_{v,i}$ remain unexposed in the σ -field $\mathcal{B}_{v,i}$, and therefore $\mathbf{R}_{v,i} \subseteq \mathbf{U}_{v,i}$. Also, the set $\mathbf{U}_{v,i} \setminus \mathbf{R}_{v,i}$ consists of all endpoints belonging to vertices at distances greater than i from v .

We are particularly interested in the random sequence of neighborhood sizes $|\mathbf{R}_{v,0}|, |\mathbf{R}_{v,1}|, \dots$, which we call the *BFS process*. We thus simplify our notation by defining $\mathbf{r}_{v,i} = |\mathbf{R}_{v,i}|$.

Before we begin our formal analysis, we engage in a brief intuitive discussion about this process. Let us imagine that we are performing BFS on a random configuration. After i levels, we will have explored the edges in $\mathbf{E}_{v,i}$, and in the next iteration, we will explore all edges incident on the endpoint set $\mathbf{R}_{v,i}$. And, conditional on $\mathcal{B}_{v,i}$, the unexposed endpoints $\mathbf{U}_{v,i}$ will be matched uniformly at random.

Now, if $\mathbf{R}_{v,i}$ is “small” relative to $\mathbf{U}_{v,i}$, then in the typical situation, each endpoint in $\mathbf{R}_{v,i}$ will match to an endpoint outside of $\mathbf{R}_{v,i}$, and no two endpoints in $\mathbf{R}_{v,i}$ will match to the same vertex. In this case, if an endpoint $a \in \mathbf{R}_{v,i}$ matches to an endpoint $\vec{E}(a)$ on a vertex

of degree d , then the remaining $d - 1$ endpoints are contributed to \mathbf{R}_{i+1} . Hence, informally, the size of $\mathbf{R}_{v,i+1}$ is approximated by

$$\mathbf{r}_{v,i+1} \simeq \sum_{a \in \mathbf{R}_{v,i}} \deg(V(\vec{\mathbf{E}}(a))) - 1 = \sum_{a \in \mathbf{R}_{v,i}} \text{res}(\vec{\mathbf{E}}(a)),$$

where $\text{res}(\vec{\mathbf{E}}(a)) = \deg(V(\vec{\mathbf{E}}(a))) - 1$ is the residual degree of the endpoint $\vec{\mathbf{E}}(a)$, as defined in Section 2.4.

Moreover, if $\mathbf{R}_{v,i}$ is “small,” then the residual degrees of the endpoints in $\vec{\mathbf{E}}(a)$ will be “almost” independent random variables, and the distribution of each of these random variables will be approximately given by the residual degree distribution $\mu_{(A,V)}$. So, again quite informally, we have

$$\mathbf{r}_{v,i+1} \simeq \sum_{j=1}^{\mathbf{r}_{v,i}} \mathbf{X}_{\mu,j},$$

where the $\mathbf{X}_{\mu,j}$ are independent μ -distributed.

We note that, if strict equality were to hold, then the equation above would define a *branching process* (see e.g., [2]), and the expected rate of neighborhood growth would be given by the average residual degree $M_1(\mu)$. To make this approximation precise, we must account for the various informal assumptions that have been made above. Specifically, the two principal differences between the the BFS process and a pure branching process are the following.

1. *Cross-edges* can occur during BFS in two ways: two endpoints in $\mathbf{R}_{v,i}$ might match to each other, or two (or more) endpoints in $\mathbf{R}_{v,i}$ might match to the same vertex outside of $\mathbf{N}_{v,i}$.
2. The residual distribution changes over time; and in fact, even at the beginning of the BFS process, we only have asymptotic knowledge of the residual distribution.

We now proceed as follows. In Section 3.1, we derive a lower bound on the growth of neighborhood sizes by relating the BFS process to independent sampling from a truncated residual distribution. Then, in Section 3.2, we derive a weaker upper bound on neighborhood growth using only the expectation. We note that faster rates of neighborhood growth generally correspond to a shorter graph distances, and as such, the *lower* bound on neighborhood sizes will be used to derive an *upper* bound on the diameter in Section 5, and vice versa.

3.1. Lower Bound on Neighborhood Sizes

In this section we derive a lower bound on the growth rate of neighborhoods during BFS. Our main result in this section is the following theorem, which establishes a bound on large deviation probabilities of neighborhood sizes.

Theorem 3.1. *For any fixed distribution μ and any $M < M_1(\mu)$, there exist $\delta > 0$ and $C > 0$ such that the following statement holds.*

Let (A, V) be an endpoint partition satisfying $\|\mu_{(A,V)} - \mu\|_1 < \delta$. For any $v \in V$ and any $i \geq 0$, if $\sum_{j=0}^i \mathbf{r}_{v,j} < \delta|A|$ then

$$\mathbf{P}[\mathbf{r}_{v,i+1} < M \cdot \mathbf{r}_{v,i} \mid \mathcal{B}_{v,i}] \leq e^{-C \cdot \mathbf{r}_{v,i}}.$$

This theorem establishes a uniform bound on large deviation probabilities, in the sense that same constant C can be used for an entire δ -neighborhood of residual distributions. Also, for a given v , the same constant can be used for every BFS iteration until the sum of the neighborhood sizes exceeds δm .

Now, if the BFS process were a true branching process generated by the residual distribution μ , then the result in Theorem 3.1 would follow immediately from well known large deviation results regarding sums of independent random variables. We thus proceed by dealing individually with each of the informal assumptions discussed above which cause the BFS process to differ from a pure branching process. Hence, in Section 3.1.1, we deal with the problem that two endpoints in $\mathbf{R}_{v,i}$ may match to each other (horizontal edges). In Section 3.1.2 we deal with the problem that two or more endpoints may match to the same vertex outside of $\mathbf{N}_{v,i}$ (diagonal edges). In Section 3.1.3 we combine these results to relate an iteration of the BFS process to independent sampling from the ϵ -truncated residual distribution, and in Section 3.1.4 we derive the large deviation result in Theorem 3.1.

3.1.1. Horizontal Edges. A horizontal edge occurs if two endpoints in $\mathbf{R}_{v,i}$ match to each other. In this section we derive a probabilistic upper bound to the number of horizontal edges which occur in an iteration of BFS.

Specifically, the following lemma shows that, if $|\mathbf{R}_{v,i}| = r \leq m/2$, then the number of endpoints in horizontal edges is dominated by the binomial distribution $\text{Bin}_{r, \sqrt{r/m}}$. Hence, if $\mathbf{R}_{v,i}$ is “small” then “most” of the endpoints will match outside of $\mathbf{R}_{v,i}$.

Lemma 3.2. *Consider a subset R of an endpoint set A , let $r = |R|$ and $m = |A|$, and assume $r \leq m/2$. For a uniformly random matching \mathbf{E} , we have*

$$\mathfrak{D}[|\vec{\mathbf{E}}(R) \cap R|] \triangleleft \text{Bin}_{r, \sqrt{r/m}}. \tag{2}$$

Proof. We first define

$$f(s, r, m) = \text{P}[|\vec{\mathbf{E}}(R) \cap R| \geq s]$$

for $r = |R|$ and let $m = |A|$, and similarly $g(s, p, r) = \text{P}[\mathbf{X}_{\text{Bin}_{r,p}} \geq s]$. Equation (2) is therefore equivalent to the statement that

$$f(s, r, m) \leq g(s, \sqrt{r/m}, r) \tag{3}$$

holds for all s .

We proceed by induction; clearly (3) holds for $r = 0, 1$ and any s and $m \geq 2r$, since if $|R| \leq 1$ then there cannot be any internally matched endpoints. Now, choose any r and assume inductively that (3) holds for all triples (s', r', m') where $r' < r$ and $m' \geq 2r'$.

We note that any given endpoint $a \in R$ matches to another endpoint within R with probability $(r - 1)/(m - 1)$. If this occurs, then $r - 2$ unexposed endpoints remain in R , and we have 2 internally matched endpoints. Otherwise, we have $r - 1$ remaining unexposed endpoints in R , and no internally matched endpoint. In both cases, the total number of unexposed endpoints remaining is $m - 2$. Hence, we inductively compute

$$\begin{aligned} f(s, r, m) &= \frac{r - 1}{m - 1} f(s - 2, r - 2, m - 2) + \frac{m - r}{m - 1} f(s, r - 1, m - 2) \\ &\leq \frac{r - 1}{m - 1} g\left(s - 2, \sqrt{\frac{r - 2}{m - 2}}, r - 2\right) + \frac{m - r}{m - 1} g\left(s, \sqrt{\frac{r - 1}{m - 2}}, r - 1\right) \\ &\leq (r/m)g(s - 2, \sqrt{r/m}, r - 2) + (1 - r/m)g(s, \sqrt{r/m}, r - 1). \end{aligned}$$

The inequality in (3) now follows by letting $p = \sqrt{r/m}$ and computing

$$\begin{aligned} g(s, p, r) &= p \cdot g(s - 1, p, r - 1) + (1 - p) \cdot g(s, p, r - 1) \\ &= (p^2 \cdot g(s - 2, p, r - 2) + p(1 - p) \cdot g(s - 1, p, r - 2)) + (1 - p) \cdot g(s, p, r - 1) \\ &\geq p^2 \cdot g(s - 2, p, r - 2) + (1 - p^2) \cdot g(s, p, r - 1). \end{aligned}$$

■

3.1.2. Diagonal Edges and Uniform Endpoint Sampling. In the previous subsection, we derived a probabilistic bound on the number of endpoints in $\mathbf{R}_{v,i}$, which match internally due to horizontal edges. In this subsection, we prove a lemma which deals with *diagonal edges*, which occur if two endpoints in $\mathbf{R}_{v,i}$ match to the same vertex outside of $N_{v,i}$.

Our approach is based on the following observation. For any integer $r \geq 0$, conditional on the event that exactly r endpoints in $\mathbf{R}_{v,i}$ match externally, the set $\mathbf{E}(\mathbf{R}_{v,i}) \setminus \mathbf{R}_{v,i}$ will be a uniformly random subset of size r from $U_{v,i} \setminus \mathbf{R}_{v,i}$. And if we let $\mathbf{S} = \mathbf{E}(\mathbf{R}_{v,i}) \setminus \mathbf{R}_{v,i}$, we have

$$\mathbf{R}_{v,i+1} = A(V(\mathbf{S})) \setminus \mathbf{S}.$$

Hence, in the following lemma, we derive a probabilistic lower bound on the size of the set $A(V(\mathbf{S})) \setminus \mathbf{S}$, where \mathbf{S} is a uniform sample of size r from an endpoint partition (A, V) .

Lemma 3.3. *For any fixed distribution μ and any $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

Let (A, V) be an endpoint partition satisfying $\|\mu_{(A,V)} - \mu\|_1 < \delta$, and consider a uniformly random subset $\mathbf{S} \subseteq A$ of size $|\mathbf{S}| = r < \delta|A|$. Then the random variable $|A(V(\mathbf{S})) \setminus \mathbf{S}|$ dominates in distribution the sum of r independent $\mu_{[\epsilon]}$ -distributed random variables.

Proof. We consider the random process $\mathbf{s}_1, \dots, \mathbf{s}_r$ corresponding to sampling from A without replacement, so at each step $1 \leq t \leq r$, we choose an endpoint \mathbf{s}_t uniformly at random from $A \setminus \mathbf{S}_{t-1}$, where $\mathbf{S}_i = \{\mathbf{s}_1, \dots, \mathbf{s}_i\}$ for $1 \leq i \leq r$ and $\mathbf{S}_0 = \emptyset$. We also define $\mathbf{Q}_t = A(V(\mathbf{S}_t)) \setminus \mathbf{S}_t$, and we write $\mathbf{q}_t = |\mathbf{Q}_t|$, so the lemma states that \mathbf{q}_r dominates in distribution the sum of r independent $\mu_{[\epsilon]}$ -distributed random variables.

At time t , we define the *effective residual degree* of any endpoint in $A \setminus \mathbf{S}_t$ by

$$\mathbf{res}_t^*(a) = \begin{cases} -1 & \text{if } a \in \mathbf{Q}_t; \\ \mathbf{res}(a) & \text{otherwise.} \end{cases}$$

We denote the distribution of effective residual degrees at time t by

$$\mathbf{v}_t(j) = \frac{|\{a \in A \setminus \mathbf{S}_t : \mathbf{res}_t^*(a) = j\}|}{|A \setminus \mathbf{S}_t|}.$$

We note that $\mathbf{q}_{t+1} - \mathbf{q}_t = \mathbf{res}_t^*(\mathbf{s}_{t+1})$, and therefore $\mathcal{D}[\mathbf{q}_{t+1} - \mathbf{q}_t \mid \mathcal{F}_t] = \mathbf{v}_t$, where $\mathcal{F}_1, \dots, \mathcal{F}_r$ is the induced filtration $\mathcal{F}_t = \sigma[\mathbf{s}_1, \dots, \mathbf{s}_t]$.

We claim that, for δ sufficiently small, we have $\|\mathbf{v}_t - \mu\| < \epsilon$ always for all t . Indeed, since $|\mathbf{S}_t| = t \leq r < \delta|A|$, then by Corollary 2.5, δ can be chosen sufficiently small that, for arbitrary $\epsilon_0 > 0$, the residual distribution at time t satisfies $\|\mu_{(A_t, V)} - \mu\|_1 < \epsilon_0$ always. And, the actual residual degree and the effective residual degree at time t differ only on the set \mathbf{Q}_t ; since \mathbf{Q}_t is contained in at most $t \leq r < \delta|A|$ vertices, then by Lemma 2.4, we can

again choose δ sufficiently small such that $|\mathbf{Q}_t| < \epsilon_0|A|$ always. By setting ϵ_0 appropriately, it follows that δ can be chosen sufficiently small that $\|\mathbf{v}_t - \mu\| < \epsilon$ always for all t .

It now follows by Lemma 2.6 that $\mathbf{v}_t \triangleright \mu_{[\epsilon]}$ always, and since $\mathfrak{D}[\mathbf{q}_{t+1} - \mathbf{q}_t \mid \mathcal{F}_t] = \mathbf{v}_t$, the proof is complete due to Lemma 2.7. \blacksquare

3.1.3. General Lower Bound on Neighborhood Sizes. In this section, we combine Lemmas 3.2 and 3.3, along with Lemma 2.9 from Section 2.5.2, to derive a general probabilistic lower bound relating the distribution of $\mathbf{r}_{v,i+1}$ given $\mathbf{r}_{v,i}$ to independent sampling from the truncated residual distribution.

Lemma 3.4. *For any fixed distribution μ and any $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

Let (A, V) an endpoint partition satisfying $\|\mu_{(A,V)} - \mu\|_1 < \delta$. Choose any $v \in V$ and any $i \geq 0$. If $\sum_{j=0}^i \mathbf{r}_{v,i} < \delta|A|$ then the distribution of $\mathbf{r}_{v,i}$ conditional on $\mathcal{B}_{v,i}$ satisfies

$$\mathfrak{D}[\mathbf{r}_{v,i+1} \mid \mathcal{B}_{v,i}] \triangleright \mathfrak{D} \left[\sum_{j=1}^{\mathbf{r}_{v,i}} \mathbf{X}_{\mu_{[\epsilon]}j} \right].$$

Proof. First, let $\mathbf{Q} = \mathbf{R}_{v,i} \setminus \vec{\mathbf{E}}(\mathbf{R}_{v,i})$ denote the set of endpoints which match outside of $\mathbf{R}_{v,i}$, and let $\mathbf{q} = |\mathbf{Q}|$. For arbitrary $\epsilon_0 > 0$, we may choose δ sufficiently small that $\sqrt{\mathbf{r}_{v,i}/|\mathbf{U}_{v,i}|} < \epsilon_0/2$, and thus by Lemma 3.2, we have

$$\mathfrak{D}[\mathbf{q} \mid \mathcal{B}_{v,i}] \triangleright \text{Bin}_{\mathbf{r}_{v,i} - \sqrt{\mathbf{r}_{v,i}/|\mathbf{U}_{v,i}|}} \triangleright \text{Bin}_{\mathbf{r}_{v,i} - \epsilon_0/2}. \tag{4}$$

Now, conditional on both $\mathcal{B}_{v,i}$ and \mathbf{q} , the set of endpoints $\vec{\mathbf{E}}(\mathbf{R}_{v,i}) \setminus \mathbf{R}_{v,i}$ is a uniformly random subset of size \mathbf{q} chosen from $\mathbf{U}_{v,i} \setminus \mathbf{R}_{v,i}$. Moreover, since $|A| - |\mathbf{U}_{v,i} \setminus \mathbf{R}_{v,i}| \leq 2 \sum_{j=0}^i \mathbf{r}_{v,i}$, then by making δ sufficiently small, by Corollary 2.5 we can ensure that the residual distribution of $(\mathbf{U}_{v,i} \setminus \mathbf{R}_{v,i}, V)$ satisfies $\|\mu_{(\mathbf{U}_{v,i} \setminus \mathbf{R}_{v,i}, V)} - \mu\|_1 < \delta_0$ for arbitrary $\delta_0 > 0$. Hence, by Lemma 3.3, we have

$$\mathfrak{D}[\mathbf{r}_{v,i+1} \mid \mathbf{q}, \mathcal{B}_{v,i}] \triangleright \mathfrak{D} \left[\sum_{j=1}^{\mathbf{q}} \mathbf{X}_{\mu_{[\epsilon]}j} \right] \tag{5}$$

for arbitrary $\epsilon_1 > 0$.

We now invoke Lemma 2.9 regarding the sum of a random number of random variables, and since $(\mu_{[\epsilon_1]})_{[\epsilon_0]} = \mu_{[\epsilon_0 + \epsilon_1]}$, it follows from (4) and (5) that $\mathfrak{D}[\mathbf{r}_{v,i+1} \mid \mathcal{B}_{v,i}] \triangleright \mathfrak{D} \left[\sum_{j=1}^{\mathbf{r}_{v,i}} \mathbf{X}_{\mu_{[\epsilon_0 + \epsilon_1]}j} \right]$. The proof is now complete since ϵ_0, ϵ_1 can be made arbitrarily small. \blacksquare

3.1.4. Large Deviations. We now use Lemma 3.4 to complete the proof of the large deviation bound in Theorem 3.1. We begin by recalling a classical large deviation result for independent, identically distributed random variables. Although the result is standard (e.g., see [10, 15]), we include a brief proof in the Appendix.

Lemma 3.5 ([10, 15]). *Let μ be any distribution on $\mathbb{Z}^* \cup \{-1\}$. Then for all $M < M_1(\mu)$ there exists a $C > 0$ such that for all $r > 0$,*

$$\mathbb{P} \left[\sum_{i=1}^r \mathbf{X}_{\mu,i} \leq Mr \right] < e^{-Cr}.$$

We now complete the proof of Theorem 3.1.

Proof of Theorem 3.1. By lemma 3.4, we can choose δ sufficiently small so that

$$P[\mathbf{r}_{v,i+1} < M \cdot \mathbf{r}_{v,i} \mid \mathcal{B}_{v,i}] \leq P\left[\sum_{j=1}^{\mathbf{r}_{v,i}} \mathbf{X}_{\mu_{[\epsilon]j}} < M \cdot \mathbf{r}_{v,i}\right]$$

for arbitrary $\epsilon > 0$. Moreover, since $\lim_{\epsilon \rightarrow 0} M_1(\mu_{[\epsilon]}) = M_1(\mu)$, then for a given $M < M_1(\mu)$, we may choose ϵ sufficiently small that $M_1(\mu_{[\epsilon]}) > M$. The result now follows by applying Lemma 3.5 to the distribution $\mu_{[\epsilon]}$. ■

3.2. Upper Bound on Neighborhood Sizes

In this section we derive an upper bound on the rate of neighborhood growth in a BFS process. The upper bound is different in nature than the lower bound derived in Section 3.1 in two ways. First, we only use the expected neighborhood size to compute the bound. Second, we bound the probability that the sum of all neighborhoods up to time t exceeds a given amount, rather than considering individual iterations.

Lemma 3.6. *For any $1 < M < \infty$, and any $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

For any endpoint partition (A, V) with $|A| = m$ and such that $M_1(\mu_{(A,V)}) < M + \delta$, and for any $q < \delta m$ and $t \geq 0$, we have

$$P\left[\sum_{i=0}^t \mathbf{r}_{v,i} > q\right] < \frac{\deg(v)(M + \epsilon)^t}{q(1 - 1/M)}.$$

Proof. For any $i \geq 0$, note that $\mathbf{r}_{v,i+1}$ is at most equal to the sum of the residual degrees of the endpoints $\vec{\mathbf{E}}(a)$ for $a \in \mathbf{R}_{v,i}$ (with equality if no cross-edges occur). Hence, by linearity of expectation, we have

$$E[\mathbf{r}_{v,i+1} \mid \mathcal{B}_{v,i}] \leq \sum_{a \in \mathbf{R}_{v,i}} E[\text{res}(\vec{\mathbf{E}}(a)) \mid \mathcal{B}_{v,i}].$$

Now for any given a , we compute

$$E[\text{res}(\vec{\mathbf{E}}(a)) \mid \mathcal{B}_{v,i}] = \frac{\sum_{a' \in \mathbf{U}_{v,i} \setminus \{a\}} \text{res}(a')}{|\mathbf{U}_{v,i}| - 1} \leq \frac{\sum_{a' \in A} \text{res}(a')}{|\mathbf{U}_{v,i}| - 1} = M_1(\mu_{(A,V)}) \frac{m}{|\mathbf{U}_{v,i}| - 1},$$

and therefore, for δ sufficiently small, we can guarantee $E[\mathbf{r}_{v,i+1} \mid \mathcal{B}_{v,i}] \leq \mathbf{r}_{v,i}(M + \epsilon)$ whenever both $M_1(\mu_{(A,V)}) < M + \delta$ and $\sum_{j=0}^i \mathbf{r}_{v,j} < \delta m$.

We now define the random variables

$$\mathbf{r}'_{v,i} = \begin{cases} \mathbf{r}_{v,i} & \text{if } \sum_{j=0}^{i-1} \mathbf{r}_{v,j} < \delta m, \\ 0 & \text{otherwise,} \end{cases}$$

and we deduce that $E[\mathbf{r}'_{v,j+1} \mid \mathcal{B}_{v,j}] \leq \mathbf{r}'_{v,j}(M + \epsilon)$ always, and therefore $E[\mathbf{r}'_{v,i}] \leq \deg(v)(M + \epsilon)^i$ for all i .

For $q < \delta m$, using Markov's inequality, we now compute

$$\begin{aligned} \mathbb{P}\left[\sum_{i=0}^t \mathbf{r}_{v,i} > q\right] &= \mathbb{P}\left[\sum_{i=0}^t \mathbf{r}'_{v,i} > q\right] \leq \frac{1}{q} \mathbb{E}\left[\sum_{i=0}^t \mathbf{r}'_{v,i}\right] \leq \frac{\deg(v)}{q} \sum_{i=0}^t (M + \epsilon)^i \\ &\leq \frac{\deg(v)(M + \epsilon)^t}{q(1 - 1/M)}. \end{aligned}$$

■

4. THE 2-CORE

The k -core of a graph is the maximal induced subgraph of minimum degree at least k . The sizes and degree distributions of the k -cores of a random graph have been determined by various authors [9, 12, 14], and our diameter result relies significantly on these results for the specific case of $k = 2$.

The relevance of the 2-core to the diameter of a random graph can be appreciated intuitively by considering the structure of a typical longest shortest path. In general, any longest shortest path in a random graph (with minimum degree 1) will occur between a pair vertices u, v of degree 1. Moreover, this path will consist of three segments: a path from u to the 2-core, a path through the 2-core, a path from the 2-core to v . It follows that, if we wish to uncover such a path using BFS, we only need to search through the set of edges in the 2-core, as well as the edges connecting u and v to the 2-core.

In this section, motivated by the informal discussion in the previous paragraph, we introduce notation and some simple results related to BFS in the 2-core of the random graph $\mathbf{G}(A, V)$. As we shall see, restricting BFS to the 2-core greatly simplifies the analysis of neighborhood growth, particularly when neighborhood sizes are small. This is because every vertex in the 2-core contains at least two endpoints, and thus if we “enter” a particular vertex through one of its endpoints, at least one other endpoint typically remains unexposed as an “exit.” It follows that, if no cross-edges occur, neighborhood sizes are strictly nondecreasing. Moreover, if no cross-edges occur, then a (strict) increase in neighborhood size will occur unless every endpoint matched during a given BFS iteration has residual degree 1. These two properties allow for a simpler and more accurate analysis of the growth of small neighborhoods than can be achieved using the tools of the previous section.

We now proceed as follows. In Section 4.1, we introduce our notation for describing the 2-core of a random graph, and we review some known results about the structure of the 2-core. We also define a related structure which we call the *augmented 2-core*, and we translate relevant results about the 2-core to the augmented 2-core. In Section 4.2, we discuss BFS in the augmented 2-core, and we derive some preliminary results related to the nondecreasing neighborhood size property discussed above. These results will enable us to pin down our main diameter result in Section 5.

4.1. The 2-Core and the Augmented 2-Core

As noted above, the 2-core of a graph is the maximal induced subgraph with minimum degree at least 2. We shall denote the 2-core of a graph $G = (A, V, E)$ by G^{κ} ; and, in general, we shall use the superscript “ κ ” to refer to the structures related to the 2-core of

G , so A^k , V^k and E^k denote the endpoint, vertex, and edge sets of the 2-core, respectively. Also, $\text{deg}^k(v)$ denotes the degree of a vertex in the 2-core, λ^k denotes the degree distribution of the 2-core, and similarly for the residual distribution.

The size and degree distribution of the 2-core of a random graph have been determined previously by several authors [9, 12, 14]. In particular, we have the following theorem, which implies that if (A, V) satisfies the asymptotic assumptions 2.1 with limiting residual distribution μ , then the 2-core of $\mathbf{G}(A, V)$ also satisfies assumptions 2.1 WHP, with a limiting residual distribution η_μ which can be computed from μ .

Theorem 4.1 ([9, 12, 14]). *Let (A, V) be an endpoint partition which asymptotically satisfies assumptions 2.1 with limiting residual distribution $\mu = \mu_\lambda$ satisfying $M_1(\mu) > 1$. Then,*

1. *there exists a constant $0 < C_\mu < 1$ determined by μ such that $|A^k|/|A| = C_\mu + o(1)$, WHP;*
2. *there exists a distribution η_μ determined by μ , such that $\|\mu^k - \eta_\mu\|_1 = o(1)$ WHP;*
3. *the distribution η_μ satisfies $M_1(\eta_\mu) = M_1(\mu)$, and in fact $M_1(\mu^k) = M_1(\mu) + o(1)$ WHP.*

Let $G = (A, V, E)$ be a graph and consider a subset $W \subseteq V$. We define the W -augmented 2-core to be the maximal induced subgraph of G such that every vertex in $V \setminus W$ has degree at least 2. Hence, the vertices in W are “exempt” from the minimum degree requirement. We denote the W -augmented 2-core of a graph by $G^{k,W}$, and related structures are denoted accordingly.

For this paper, we shall often examine W -augmented 2-core where W is a “small” subset of V . Our next lemma follows almost immediately from the proof of Theorem 4.1 (in fact the bound on $|A(W)|$ can be weakened to $o(m)$). However, for completeness, we include an independent proof in the Appendix.

Lemma 4.2. *Consider an endpoint partition (A, V) as in Theorem 4.1, and consider a subset $W \subseteq V$ such that $|A(W)| = o(m/\ln m)$. Then WHP $|A^{k,W} \setminus A^k| = o(m)$.*

If, in addition, $|A(W)| = m^{\Omega(1)}$ then AAS a constant fraction of the endpoints in $A(W)$ belong to $A^{k,W}$.

It follows immediately from Lemma 4.2, that for any subset $W \subseteq V$ of size $|A(W)| = o(m/\ln m)$, the statement of Theorem 4.1 holds if we replace A^k and μ^k with $A^{k,W}$ and $\mu^{k,W}$, respectively.

Given a random graph $\mathbf{G}(A, V)$, we define the 2-core σ -field to be the σ -field induced by the set of edges which are *not* in the 2-core. We denote this σ -field by $\mathcal{F}^k = \sigma[\mathbf{E} \setminus \mathbf{E}^k]$. Intuitively, the 2-core σ -field represents a state of information in which all the edges which are not in the 2-core are known, but the structure of the 2-core remains unknown. It follows from the proof of Theorem 4.1 in [12] (and it is not difficult to see) that the 2-core σ -field \mathcal{F}^k is subuniform. Similarly, for any $W \subseteq V$, we define the W -augmented 2-core σ -field $\mathcal{F}^{k,W} = \sigma[\mathbf{E} \setminus \mathbf{E}^{k,W}]$, noting that $\mathcal{F}^{k,W}$ is also subuniform.

4.2. BFS and the 2-Core

In this section, we present some definitions and a simple lemma related to BFS in the augmented 2-core of a random graph.

For any vertex v in a graph G , we define the *2-core endpoint neighborhoods* of v , denoted by $R_{v,i}^k$, to be the endpoint neighborhoods of v in the v -augmented 2-core $G^{k,v}$. It is easy to see that any path in the original graph G , which connects v to a vertex $u \in G^{k,v}$ is itself contained in $G^{k,v}$, and it follows that the 2-core endpoint neighborhoods can be expressed simply as

$$R_{v,i}^k = R_{v,i} \cap A^{k,v}. \tag{6}$$

As usual, the 2-core endpoint neighborhoods of a vertex in a random graph are denoted in boldface by $\mathbf{R}_{v,i}^k$. We also define the *2-core BFS filtration* $\mathcal{B}_{v,i}^k$ analogously to the ordinary BFS filtration, by $\mathcal{B}_{v,0}^k = \mathcal{F}^{k,v}$ and $\mathcal{B}_{v,i+1}^k = \sigma[\mathcal{B}_{v,i}^k, \mathbf{E}(\mathbf{R}_{v,i}^k)]$. The 2-core BFS filtration thus corresponds to first exposing all edges outside the v -augmented 2-core and then performing BFS beginning at v . Moreover, using the expression in (6) for 2-core endpoint neighborhoods, we may express the 2-core BFS filtration in terms of the ordinary BFS filtration by

$$\mathcal{B}_{v,i}^k = \sigma[\mathcal{F}^{k,v}, \mathcal{B}_{v,i}].$$

As discussed at the beginning of Section 4, BFS in the 2-core exhibits the useful property that neighborhood sizes are generally nondecreasing. In the following lemma, we make this property precise, and we also bound the probability of repeating a given neighborhood size.

Lemma 4.3. *For any fixed distribution μ with $M_1(\mu) > 1$ and any $\epsilon > 0$, there exist constants C and $\delta > 0$ such that the following holds WHP.*

Let (A, V) an endpoint partition with $|A| = m$, and assume that $\|\mu_{(A,V)} - \mu\|_1 < \delta$. Choose any $v \in V$ and any $i \geq 0$, and assume that $\sum_{j=0}^i \mathbf{r}_{v,i}^k < \delta m$. Then

1. $\mathbf{P}[\mathbf{r}_{v,i+1}^k < \mathbf{r}_{v,i}^k \mid \mathcal{B}_{v,i}^k] \leq C \cdot (\mathbf{r}_{v,i}^k)^3 / m$.
2. $\mathbf{P}[\mathbf{r}_{v,i+1}^k < \mathbf{r}_{v,i}^k - 2 \mid \mathcal{B}_{v,i}^k] \leq C \cdot (\mathbf{r}_{v,i}^k)^6 / m^2$.
3. $\mathbf{P}[\mathbf{r}_{v,i+1}^k = \mathbf{r}_{v,i}^k \mid \mathcal{B}_{v,i}^k] \leq (\eta_\mu(1) + \epsilon) \mathbf{r}_{v,i}^k + C \cdot (\mathbf{r}_{v,i}^k)^3 / m$.

Proof. First, note that by choosing $\delta > 0$ sufficiently small and invoking Theorem 4.1 and Lemma 4.2, we can guarantee that the fraction of $\mathcal{B}_{v,i}^k$ -unexposed endpoints which have unexposed residual degree 1 is at most $\eta_\mu(1) + \epsilon$ WHP. Now, for each $a \in \mathbf{R}_{v,i}^k$, define the $\mathcal{B}_{v,i+1}^k$ -measurable random variable $\mathbf{y}(a)$ by letting $\mathbf{y}(a) = 1$ if either

1. $\bar{\mathbf{E}}(a) \in \mathbf{R}_{v,i}^k$, or
2. $V(\bar{\mathbf{E}}(a)) = V(\bar{\mathbf{E}}(a')) = w$ for some vertex w with $\deg^{k,v}(w) \leq 2\mathbf{r}_{v,i}^k$,

and $\mathbf{y}(a) = 0$ otherwise. We also define $\mathbf{Y} = \sum_{a \in \mathbf{R}_{v,i}^k} \mathbf{y}(a)$, and since every vertex other than v has unexposed degree at least 2, it follows that $\mathbf{r}_{v,i+1}^k \geq \mathbf{r}_{v,i}^k - \mathbf{Y}$.

Now, for any given pair $a, a' \in \mathbf{R}_{v,i}^k$, we have $\mathbf{P}[\bar{\mathbf{E}}(a) = a' \mid \mathcal{B}_{v,i}^k] = O(1/m)$. Also, the probability, conditional on $\mathcal{B}_{v,i}^k$, that a and a' both match to a vertex w with $\deg^{k,v}(w) \leq 2\mathbf{r}_{v,i}^k$ is $O(\mathbf{r}_{v,i}^k/m)$. By considering all pairs $a, a' \in \mathbf{R}_{v,i}^k$, it follows that

$$\mathbf{P}[\mathbf{Y} > 0 \mid \mathcal{B}_{v,i}^k] = O((\mathbf{r}_{v,i}^k)^3 / m).$$

Similarly, by considering all possible ways that at least three endpoints in $\mathbf{R}_{v,i}^k$ can be involved in cross-edges of the kinds described above, we easily deduce

$$P[\mathbf{Y} > 2 \mid \mathcal{B}_{v,i}] = O((r_{v,i}^k)^6/m^2).$$

For the final statement of the lemma, we note that the probability, conditional on $\mathcal{B}_{v,i}^k$, that every endpoint in $\mathbf{R}_{v,i}^k$ matches to an endpoint of residual degree 1 outside of $\mathbf{R}_{v,i}^k$ is at most

$$\mu_{v,i}^k(1)^{r_{v,i}^k} \leq (\eta_\mu(1) + \epsilon)^{r_{v,i}^k}.$$

And, the event $\mathbf{r}_{v,i+1}^k = \mathbf{r}_{i,v}^k$ can only occur if either $\mathbf{Y} > 0$ or if every endpoint in $\mathbf{R}_{v,i}^k$ matches to an endpoint of residual degree 1, hence the proof is complete. ■

5. THE DIAMETER OF A RANDOM GRAPH

Given vertices u, v in a graph G , let $\delta(u, v)$ denote the distance from u to v , that is, the length of a shortest path from u to v . We set $\delta(u, v) = \infty$ if u and v do not belong to the same connected component. The diameter $\Delta(G)$ of a graph G is the maximum finite distance between any pair of vertices in G . In this section, we compute the diameter of a random graph $\mathbf{G}(A, V)$ with asymptotic precision.

We begin by stating our main theorem regarding the diameter of a random graph. Recall that, for a distribution μ , η_μ denotes the limiting residual distribution of the 2-core of a random graph with limiting residual distribution μ , as determined by Theorem 4.1.

Theorem 5.1. *Let (A, V) be an endpoint partition which asymptotically satisfies assumptions 2.1, with limiting degree distribution λ and corresponding residual distribution $\mu = \mu_\lambda$. Assume further that both $M_1(\mu) > 1$ and $\lambda(1) > 0$.*

Then, the diameter of $\mathbf{G}(A, V)$ AAS satisfies $\Delta(\mathbf{G}(A, V)) = \Delta_\mu \ln n + o(\ln n)$, where

$$\Delta_\mu = \frac{2}{-\ln \eta_\mu(1)} + \frac{1}{\ln M_1(\mu)}. \tag{7}$$

Due to a result of Molloy and Reed [17, 18], the assumption that $M_1(\mu) > 1$ implies that the largest connected component of $\mathbf{G}(A, V)$ AAS has size $\Theta(n)$ —this is the so-called “giant” component—while the second largest component AAS has size $o(n)$ (and therefore all of the “non-giant” components are “small”). An immediate consequence of the proof of Theorem 5.1 is that the diameter of $\mathbf{G}(A, V)$ is equal to the diameter of the giant component.

The assumption that $\lambda(1) > 0$ ensures that the minimum (positive) degree is 1. The behavior of the diameter in the case $\lambda(1) = 0$ is qualitatively different, and we deal with generalizations to higher minimum degree in Section 5.3.

Informally, the two terms on the right hand side of Eq. (7) correspond to two different characteristics of a random graph that determine its diameter. The term $\frac{1}{\ln M_1(\mu)}$ measures the “average” distance across the graph. The term $\frac{2}{-\ln \eta_\mu(1)}$ accounts for long isolated paths or cycles which can cause the distance between a particular pair of vertices to be significantly longer than the average. A typical longest shortest path will consist of an “average” shortest path of length $\frac{\ln n}{\ln M_1(\mu)}$ extended on both ends by two long “strands,” each having length $\frac{\ln n}{-\ln \eta_\mu(1)}$.

We prove the upper bound of Theorem 5.1 in Section 5.1 and the lower bound in Section 5.2. Then, in Section 5.3 we generalize to higher minimum degrees.

5.1. Upper Bound Proof for Theorem 5.1

In this section, we prove the upper bound of Theorem 5.1. Our general approach is to use the results from Sections 3 and 4 to trace the rate of neighborhood growth during BFS in the augmented 2-core. We maintain the unstated assumption that we are considering a random graph $\mathbf{G}(A, V)$ such that the endpoint partition (A, V) asymptotically satisfies assumptions 2.1 with asymptotic degree distribution λ and corresponding residual distribution $\mu = \mu_\lambda$. And, as usual the number of endpoints and vertices in (A, V) are denoted by $m = |A|$ and $n = |V|$, respectively.

We begin by observing that “large” endpoint neighborhoods will AAS be connected by an edge. Specifically, we have the following lemma; the proof is straightforward and appears in the Appendix.

Lemma 5.2. *Consider disjoint subsets $S_1, S_2 \subseteq A$ such that both $|S_1|, |S_2| > m^{1/2} \ln^2 m$. Then*

$$P[\vec{\mathbf{E}}(S_1) \cap S_2 = \emptyset] = m^{-\omega(1)}.$$

For our purposes, then, a “large” neighborhood consists of at least $m^{1/2} \ln^2 m$ endpoints. Since this value plays a central role in the upper bound proof of Theorem 5.1, we abbreviate by defining

$$\xi = m^{1/2} \ln^2 m.$$

Now, for any pair of vertices u, v , and any integers t_1, t_2 , if both $\mathbf{r}_{u,t_1} > \xi$ and $\mathbf{r}_{v,t_2} > \xi$, it follows immediately from lemma 5.2 that $\delta(u, v) \leq t_1 + t_2 + 1$ WHP. Moreover, since 2-core neighborhood sizes satisfy $\mathbf{r}_{v,i}^k \leq \mathbf{r}_{v,i}$ always, we have the following WHP implication:

$$(\mathbf{r}_{u,t_1}^k > \xi) \wedge (\mathbf{r}_{v,t_2}^k > \xi) \implies \delta(u, v) \leq t_1 + t_2 + 1. \tag{8}$$

For any integer $r \geq 0$, we now define the random variables

$$\begin{aligned} \tau_v(r) &= \min \{i : \mathbf{r}_{v,i}^k \geq r\}, \\ \rho_v(r) &= \min \{i : \mathbf{r}_{v,i}^k \geq r \text{ or } \mathbf{r}_{v,i}^k = 0\}. \end{aligned} \tag{9}$$

Intuitively, $\tau_v(r)$ is the closest 2-core neighborhood which contains at least r endpoints, and $\rho_v(r)$ is the closest 2-core neighborhood which is either empty or has size at least r . Both $\tau_v(r)$ and $\rho_v(r)$ are clearly stopping times with respect to the 2-core BFS filtration $\mathcal{B}_{v,i}^k$. We note that $\tau_v(r) = \infty$ if and only if $\mathbf{r}_{v,i}^k < r$ for all i , while $\rho_v(r) < \infty$ always. Moreover, if $\tau_v(r) < \infty$ then in fact $\tau_v(r) = \rho_v(r)$.

It follows from the WHP implication in (8) that the inequality

$$\delta(u, v) \leq \tau_v(\xi) + \tau_u(\xi) + 1 \tag{10}$$

holds WHP for all pairs $u, v \in V$. Hence, if we are able to bound the maximum finite value of $\tau_v(\xi)$ for all $v \in V$, then (10) will yield an upper bound on the distance between any pair of vertices which both satisfy $\tau_v(\xi) < \infty$.

Of course, the inequality (10) is only useful for vertices satisfying $\tau_v(\xi) < \infty$. A different method is necessary to deal with vertices for which $\tau_v(\xi) = \infty$. Intuitively, the condition $\tau_v(\xi) < \infty$ holds if and only if the vertex v belongs to the giant component (in fact, this statement holds AAS for all v). As we shall see, it is somewhat easier to bound the distance between vertices satisfying $\tau_v(\xi) = \infty$, and the substance of the upper bound proof is the bound on the maximum finite value of $\tau_v(\xi)$.

We now proceed as follows. In Section 5.1.2, we derive an upper bound on the distance between any pair of vertices satisfying $\tau_v(\xi) < \infty$ as described above, or equivalently, a bound on the diameter of the giant component. Then, in Section 5.1.3, we bound the maximum distance between any pair of vertices which are not in the giant component. Moreover, the upper bound we derive in Section 5.1.3 is strictly lesser than the expression for the diameter of $\mathbf{G}(A, V)$ in Theorem 5.1. Hence, this bound, in conjunction with the upper bound proof of Theorem 5.1, will establish that the diameter of $\mathbf{G}(A, V)$ is AAS equal to the diameter of the giant component.

5.1.1. Neighborhood Growth in the 2-Core. In this section, we present a concise summary of key results from Sections 3 and 4, stated in the form that will be used in the upper bound proof of Theorem 5.1. We note that for any $t = O(\ln n)$, if $t \leq \rho_v(\xi)$, then all neighborhoods $\mathbf{r}_{v,i}^k$ for $i \leq t$ have size at most $\xi = m^{1/2} \ln^2 m$, and hence the sum of their sizes is $o(m)$. This observation allows us to invoke both Theorem 3.1 (with respect to the 2-core residual distribution η_μ) and Lemma 4.3. The following proposition is therefore immediate and serves essentially as a summary of these earlier results.

Proposition 5.3. *For any $v \in V$ and $t = O(\ln n)$, the following hold WHP whenever $t < \rho_v(\xi)$.*

1. $\mathbb{P}[\mathbf{r}_{v,t+1}^k < \mathbf{r}_{v,t}^k \mid \mathcal{B}_{v,t}^k] = \tilde{O}(n^{-1})$.
2. $\mathbb{P}[\mathbf{r}_{v,t+1}^k < \mathbf{r}_{v,t}^k - 2 \mid \mathcal{B}_{v,t}^k] = \tilde{O}(n^{-2})$.
3. $\mathbb{P}[\mathbf{r}_{v,t+1}^k = \mathbf{r}_{v,t}^k \mid \mathcal{B}_{v,t}^k] \leq \eta_\mu(1) + o(1)$.
4. For any constant $M < M_1(\eta_\mu) = M_1(\mu)$,

$$\mathbb{P}[\mathbf{r}_{v,t+1}^k < M \cdot \mathbf{r}_{v,t}^k \mid \mathcal{B}_{v,t}^k] = e^{-\Omega(\mathbf{r}_{v,t}^k)}.$$

From the first two statements in this proposition, we easily deduce the following corollary regarding decreases in neighborhood sizes.

Corollary 5.4. *For any $v \in V$, and any $t = O(\ln n)$, with probability $1 - o(n^{-1})$:*

1. The event $\mathbf{r}_{v,s+1}^k < \mathbf{r}_{v,s}^k$ occurs at most once for all s satisfying both $s < t$ and $s < \rho_v(\xi)$;
2. The event $\mathbf{r}_{v,s+1}^k < \mathbf{r}_{v,s}^k - 2$ never occurs for any s satisfying both $s < t$ and $s < \rho_v(\xi)$.

In particular, note that if the value of $\mathbf{r}_{v,s}^k$ exceeds 2 for any $s \geq 0$, then with probability $1 - o(n^{-1})$, we have $\mathbf{r}_{v,t}^k > 0$ for all $t = O(\ln n)$ and $t \leq \rho_v(\xi)$. Moreover, Theorem 5.5, which is proved in the next section, establishes that $\rho_v(\xi) = O(\ln n)$ with probability $1 - o(n^{-1})$. It follows that, with probability $1 - o(n^{-1})$, for any vertex v , if $\tau_v(3) < \infty$ then $\tau_v(\xi) < \infty$, and therefore the giant component AAS consists of exactly the set of vertices satisfying $\tau_v(3) < \infty$.

5.1.2. *The Giant Component.* In this section, we prove the following theorem, which establishes an upper bound of $\frac{\Delta_\mu}{2} \ln n + o(\ln n)$ for the stopping time $\rho_v(\xi)$ (where Δ_μ is defined in Eq. (7)).

Theorem 5.5. *For any fixed $\epsilon > 0$, and for all $v \in V$,*

$$P \left[\rho_v(\xi) > \left(\frac{1}{-\ln \eta_\mu(1)} + \frac{1}{2M_1(\mu)} + \epsilon \right) \ln n \right] = o(n^{-1}). \tag{11}$$

Since $\tau_v(r) = \rho_v(r)$ if and only if $\tau_v(r) < \infty$ (see (9)), then we also have an upper bound on the maximum finite value of $\tau_v(\xi)$ for all $v \in V$; and, by Lemma 5.2, this yields an upper bound on the diameter of giant component. Moreover, as noted above, Theorem 5.5 and Corollary 5.4 jointly imply that $\tau_v(3) < \infty \implies \tau_v(\xi) < \infty$ with probability $1 - o(n^{-1})$. Hence (pending the proof of Theorem 5.5), we have the following corollary.

Corollary 5.6. *For all $\epsilon > 0$, the maximum distance between any pair of vertices u, v satisfying $\tau_u(3) < \infty$ and $\tau_v(3) < \infty$ is AAS at most $(\Delta_\mu + \epsilon) \ln n$.*

To derive the upper bound in (11), we split the interval $0, \dots, \rho_v(\xi)$ into two subintervals using the intermediate stopping time $\rho_v(\ln \ln n)$. We bound the length of each of these subintervals in Lemmas 5.7 and 5.8. The proofs of these lemmas are quite similar; in each case, we exploit the fact that neighborhood sizes are generally nondecreasing, and we simply count the number of “good” iterations for which the increase in neighborhood size exceeds a given amount. We then use Proposition 5.3 to derive a lower bound on the probability that any given iteration is “good,” and we determine the maximum number of good iterations which can occur before the neighborhood size must exceed the given limit. Finally, we use Lemma 2.8 from Section 2.5.2 to attain bounds on the corresponding stopping times from the binomial distribution.

We first consider neighborhoods of size $0 < \mathbf{r}_{v,t}^\kappa < \ln \ln n$. Here, a “good” iteration occurs whenever $\mathbf{r}_{v,t}^\kappa > \mathbf{r}_{v,t-1}^\kappa$, and the probability of a good iteration is bounded below by $1 - \eta_\mu(1) - o(1)$.

Lemma 5.7. *For any vertex v and any $\epsilon > 0$,*

$$P \left[\rho_v(\ln \ln n) > \left(\frac{1}{-\ln \eta_\mu(1)} + \epsilon \right) \ln n \right] = o(n^{-1}).$$

Proof. For $t \geq 1$, define the $\mathcal{B}_{v,t}^\kappa$ -measurable random variable

$$\mathbf{z}_t = \begin{cases} 0 & \text{if } \mathbf{r}_{v,t}^\kappa \leq \mathbf{r}_{v,t-1}^\kappa, \\ 1 & \text{if } \mathbf{r}_{v,t}^\kappa > \mathbf{r}_{v,t-1}^\kappa, \end{cases}$$

and let $\mathbf{Z}_t = \sum_{i=1}^t \mathbf{z}_i$. It follows from the nondecreasing neighborhood size property expressed in Corollary 5.4 that for any $t = O(\ln n)$ and $t < \rho_v(\ln \ln n)$, if $\mathbf{r}_{v,0}^\kappa > 0$ then $\mathbf{r}_{v,t}^\kappa \geq \mathbf{Z}_t - 1$ with probability $1 - o(n^{-1})$. Hence, for $t = O(\ln n)$,

$$P[\rho_v(\ln \ln n) > t \wedge \mathbf{Z}_t > \ln \ln n] = o(n^{-1}). \tag{12}$$

Now, since $\rho_v(\ln \ln n) \leq \rho_v(\xi)$, then by Statement 3 of Proposition 5.3 we have

$$P[\mathbf{z}_{t+1} = \mathbf{0} \mid \rho_v(\ln \ln n) > t, \mathcal{B}_{v,t}^\kappa] \leq \eta_\mu(1) + o(1)$$

WHP. It thus follows from (12) and Lemma 2.8 that

$$\begin{aligned} P[\rho_v(\ln \ln n) > t] &\leq P[\rho_v(\ln \ln n) > t \wedge \mathbf{Z}_t \leq \ln \ln n] + o(n^{-1}) \\ &\leq P[\mathbf{X}_{\text{Bin}_{t,1-\eta_\mu(1)-o(1)}} \leq \ln \ln n] + o(n^{-1}). \end{aligned}$$

We now invoke a simple binomial inequality

$$P[\mathbf{X}_{\text{Bin}_{t,1-p}} \leq r] \leq \binom{t}{r} p^{t-r} \leq \frac{t^r p^t}{p^r},$$

and for $t = c \ln n$, $r = \ln \ln n$, and $p = \eta_\mu(1) + o(1)$, we deduce

$$\begin{aligned} P[\rho_v(\ln \ln n) > c \ln n] &\leq \left(\frac{c \ln n}{\eta_\mu(1) + o(1)} \right)^{\ln \ln n} (\eta_\mu(1) + o(1))^{c \ln n} + o(n^{-1}) \\ &\leq n^{c \ln \eta_\mu(1) + o(1)} + o(n^{-1}). \end{aligned}$$

Hence, for $c = \frac{1}{-\ln \eta_\mu(1)} + \epsilon$, we have $P[\rho_v(\ln \ln n) > c \ln n] = o(n^{-1})$, and the proof is complete. ■

Next, we consider neighborhood sizes $\ln \ln n$ up to $\xi = m^{1/2} \ln^2 m$. In this range, a “good” iteration occurs if $\mathbf{r}_{v,t+1}^\kappa > M \cdot \mathbf{r}_{v,t}^\kappa$ for arbitrary $M < M_1(\mu)$. We note that if $\mathbf{r}_{v,\rho_v(\ln \ln n)}^\kappa = \mathbf{0}$ then $\rho_v(\xi) = \rho_v(\ln \ln n)$, so the duration of this interval is 0. Thus, in the following lemma we consider the nontrivial case where $\mathbf{r}_{v,\rho_v(\ln \ln n)}^\kappa \geq \ln \ln n$, and therefore $\rho_v(\ln \ln n) = \tau_v(\ln \ln n)$.

Lemma 5.8. *For all $\epsilon > 0$, there exists $\delta > 0$ such that, for all v ,*

$$P\left[\infty > \tau_v(\xi) > \tau_v(\ln \ln n) + \left(\frac{1}{2M_1(\mu)} + \epsilon\right) \ln n\right] = o(n^{-1}).$$

Proof. We begin by considering an arbitrary constant M satisfying $1 < M < M_1(\mu)$; the value of M will be specified further on. For each $t > \tau(v, \ln \ln n)$, we now define

$$\mathbf{y}_t = \begin{cases} 1 & \text{if } \mathbf{r}_{v,t+1}^\kappa / \mathbf{r}_{v,t}^\kappa \geq M, \\ 0 & \text{if } \mathbf{r}_{v,t+1}^\kappa / \mathbf{r}_{v,t}^\kappa < M, \end{cases}$$

and $\mathbf{Y}_t = \sum_{i=\tau_v(\ln \ln n)}^t \mathbf{y}_i$. Corollary 5.4, now implies that for $t = O(\ln n)$ and $\tau_v(\ln \ln n) < t < \rho_v(\xi)$, with probability $1 - o(n^{-1})$, we have (for n sufficiently large)

$$\mathbf{r}_{v,t}^\kappa \geq M^{\mathbf{Y}_t} (\ln \ln n - 2) > e^{\mathbf{Y}_t \ln M}.$$

It follows that for $s = O(\ln n)$,

$$P\left[(\rho_v(\xi) - \tau_v(\ln \ln n)) > s \wedge \left(\mathbf{Y}_{\tau_v(\ln \ln n)+s} > \frac{\ln \xi}{\ln M}\right)\right] = o(n^{-1}). \tag{13}$$

Using Statement 4 of Proposition 5.3, it follows that with probability $1 - o(n^{-1})$, for $t = O(\ln n)$,

$$P[\mathbf{y}_{t+1} = 0 \mid \tau_v(\ln \ln n) < t < \tau_v(\xi), \mathcal{B}_{v,t}^k] \leq e^{-\Omega(\ln \ln n)} = o(1),$$

and hence by (13) and Lemma 2.8, for $s = O(\ln n)$,

$$P[\tau_v(\xi) - \tau_v(\ln \ln n) > s] \leq P\left[\mathbf{X}_{\text{Bin}_{s,1-o(1)}} \leq \frac{\ln \xi}{\ln M}\right] + o(n^{-1}).$$

Since

$$\ln \xi = \ln m^{1/2} \ln^2 m = (1/2 + o(1)) \ln n,$$

we set $s = c \ln n$ for a constant $c > \frac{1}{2 \ln M}$ and deduce

$$\begin{aligned} P[\tau_v(\xi) - \tau_v(\ln \ln n) > c \ln n] &\leq P\left[\mathbf{X}_{\text{Bin}_{c \ln n, 1-o(1)}} \leq \left(\frac{1}{2 \ln M} + o(1)\right) \ln n\right] + o(n^{-1}) \\ &= P\left[\mathbf{X}_{\text{Bin}_{c \ln n, o(1)}} \geq \left(c - \frac{1}{2 \ln M} - o(1)\right) \ln n\right] + o(n^{-1}) \\ &= P[\mathbf{X}_{\text{Bin}_{c \ln n, o(1)}} \geq \Omega(\ln n)] + o(n^{-1}) \\ &= o(n^{-1}). \end{aligned}$$

Moreover, by making $M_1(\mu) - M$ arbitrarily small, we can achieve this bound for any $c = \frac{1}{2 \ln M_1(\mu)} + \epsilon$, and the proof is complete. \blacksquare

The proof of Theorem 5.5 is now complete, and, as discussed above, this establishes an upper bound on the diameter of the giant component, as stated in Corollary 5.6.

5.1.3. Small Components. We now deal with distances between vertices which are not in the giant component, or, more precisely, between vertices for which $\tau(\xi) = \infty$. We consider the two possible types of small components in $\mathbf{G}(A, V)$ separately: tree components and components that contain cycles.

We first compute an upper bound on the diameter of the tree components.

Lemma 5.9. *For any $\epsilon > 0$, the maximum diameter of any tree component is AAS less than*

$$\left(\frac{1}{-\ln \eta_\mu(1)} + \epsilon\right) \ln n. \tag{14}$$

Proof. Note that if u, v belong to the same tree component, then the augmented 2-core $\mathbf{G}^{\kappa, \{u,v\}}$ consists of the ordinary 2-core \mathbf{G}^κ along with a path connecting u, v which is disjoint from \mathbf{G}^κ . Also, the unexposed degrees of u, v must satisfy

$$\text{deg}^{\kappa, \{u,v\}}(u) = \text{deg}^{\kappa, \{u,v\}}(v) = 1.$$

Now, if we perform BFS beginning from u in $\mathbf{G}^{\kappa, \{u,v\}}$, to reveal a path of length r connecting u to v which does not connect to the 2-core, we must encounter $r - 1$ consecutive endpoints of unexposed residual degree 1, and then in the r 'th iteration, we must encounter the remaining unexposed endpoint on the vertex v . Since the fraction of endpoints of unexposed residual degree 1 is at WHP at most $\eta_\mu(1) + o(1)$, then the probability of choosing

$\left(\frac{1}{-\ln \eta_\mu(1)} + \epsilon\right) \ln n$ consecutive endpoints of residual degree 1 is $o(n^{-1})$. And, the probability of choosing the remaining endpoint on v is $O(n^{-1})$, so the overall probability is $o(n^{-2})$. The lemma now follows by considering all $O(n^2)$ possible pairs of endpoints and invoking the first moment method. ■

We now consider small components which contain cycles. More precisely, we recall that corollary 5.6 establishes an upper bound for the distance between any pair of vertices for which $\tau_v(3) < \infty$. Hence, we consider vertices for which $\tau_v(3) = \infty$.

Lemma 5.10. *For any $v \in V$ and any $\epsilon > 0$,*

$$P[\tau_v(3) = \infty \wedge \rho_v(3) > \epsilon \ln n] = O(n^{-(1+\Omega(1))}).$$

Proof. We first note that if $\tau_v(3) = \infty$, then we must have $\deg^{k,v}(v) = \mathbf{r}_{v,0}^k \leq 2$. Moreover, in order for $\rho_v(3) > \epsilon \ln n$, the neighborhood size $\mathbf{r}_{v,i}^k$ must be either 1 or 2 for $\epsilon \ln n$ consecutive iterations, and since the probability of repeating a neighborhood size is WHP at most $\eta_\mu(1) + o(1)$, we have

$$P[\tau_v(3) = \infty \wedge \rho_v(3) > \epsilon \ln n] \leq (\eta_\mu + o(1))^{\epsilon \ln n - 1} + O(n^{-1}) = O(n^{-\Omega(1)}).$$

Then, in order for $\tau_v(3) = \infty$, the neighborhood size must drop to 0 due to a cross edge; this occurs with probability $O(n^{-1})$. Hence, the probability of at least $\epsilon \ln n$ neighborhoods of size either 1 or 2 followed by a neighborhood of size 0 is $O(n^{-(1+\Omega(1))})$. ■

This lemma immediately implies if $\tau_v(3) = \infty$ then with probability $1 - o(n^{-1})$ the maximum finite distance between v and any vertex in the 2-core is less than $\epsilon \ln n$. It follows that the maximum finite distance between any two such vertices is AAS less than $2\epsilon \ln n$ for arbitrary $\epsilon > 0$.

This lemma also implies that if $\tau_v(3) = \infty$, then with probability $1 - o(n^{-1})$, most $\epsilon \ln n$ vertices in the 2-core belong to the same connected component as v . Therefore, v cannot be connected by a path to any vertex u with $\tau_u(3) < \infty$, since any such u will AAS belong to the giant component as shown previously.

These observations, which are summarized in the following corollary, complete the upper bound proof of Theorem 5.1.

Corollary 5.11. *The following occur AAS simultaneously for all pairs of vertices u, v , neither of which belong to tree components:*

1. if $\tau_u(3) = \infty$ and $\tau_v(3) = \infty$ then either $\delta(u, v) = o(\ln n)$ or $\delta(u, v) = \infty$;
2. if $\tau_u(3) < \infty$ and $\tau_v(3) = \infty$ then $\delta(u, v) = \infty$.

5.2. Lower Bound Proof for Theorem 5.1

In this section we prove the lower bound of Theorem 5.1 by showing that for any $\epsilon > 0$, there AAS exist vertices $u, v \in V$ such that $\delta(u, v) > (\Delta_\mu - \epsilon) \ln n$. Although we are ultimately interested in the limiting residual distribution μ , we first consider the diameter of a random graph in which the limiting residual distribution is the 2-core distribution η_μ , and then we extend this to the desired result.

We first prove that almost all pairs of vertices in a graph with residual distribution η_μ are connected by a path of length at least $\frac{\ln n}{\ln M_1(\mu)} - o(\ln n)$.

Lemma 5.12. *Let (A, V) be an endpoint partition which asymptotically satisfies assumptions 2.1 with limiting residual distribution η_μ , and let u, v be vertices of degree $O(1)$. Then, for all $\epsilon > 0$*

$$P \left[\infty > \delta(u, v) > \left(\frac{1}{\ln M_1(\mu)} - \epsilon \right) \ln n \right] = 1 - o(1).$$

Proof. It is shown in [17, 18] that the giant component of such a graph AAS contains all but $o(n)$ vertices, and it follows easily that any given vertex belongs to the giant component with probability $1 - o(1)$. Hence, for any given pair $u, v \in V$, we have $P[\delta(u, v) < \infty] = 1 - o(1)$.

For the lower bound on $\delta(u, v)$, we use the upper bound on neighborhood growth from Lemma 3.6. Specifically, since $M_1(\eta_\mu) = M_1(\mu)$, then for arbitrary $\epsilon_0 > 0$ there exists $\delta > 0$ such that

$$P \left[\sum_{i=0}^t r_{v,i} > q \right] = O \left(\frac{(M_1(\mu) + \epsilon_0)^t}{q} \right)$$

for all $q < \delta m$. Hence, by making ϵ_0 sufficiently small, for all $t \leq \left(\frac{1}{\ln M_1(\mu)} - \epsilon \right) \ln n$ we deduce $P \left[\sum_{i=0}^t r'_{v,i} = o(m) \right] = 1 - o(1)$.

In other words, with probability $1 - o(1)$, the first t neighborhoods have combined size $o(m)$, and thus the probability that a path from v to an arbitrary vertex of degree $O(1)$ is exposed is also $o(1)$. We conclude that the distance from u to v is at least t with probability $1 - o(1)$. ■

We now consider the diameter a random graph with limiting distribution η_μ but which contains at least $m^{1-o(1)}$ vertices of degree 1.

Lemma 5.13. *Let (A, V) be an endpoint partition which asymptotically satisfies assumptions 2.1 with limiting residual distribution η_μ and assume that V contains at least $m^{1-o(1)}$ vertices of degree 1. Then, for any $\epsilon > 0$, the diameter of $\mathbf{G}(A, V)$ is AAS at least*

$$\left(\frac{2}{-\ln \eta_\mu(1)} + \frac{1}{\ln M_1(\mu)} - \epsilon \right) \ln n.$$

Proof. Let $V_0 \subseteq V$ denote the set of vertices of degree 1, and for any $v \in V_0$, let \mathbf{w}_v denote closest vertex to v such that $\deg(\mathbf{w}_v) \neq 2$ (other than v itself). Note that \mathbf{w}_v can be found easily by performing BFS from v until a vertex of some degree other than 2 is found.

We now choose an integer $C \geq 2$ such that $\eta_\mu(C) > 0$, and we let $t = \left\lfloor \left(\frac{1}{-\ln \eta_\mu(1)} - \epsilon_0 \right) \ln n \right\rfloor$ for arbitrary $\epsilon_0 > 0$. For each $v \in V_0$, we define the following events:

- $\mathbf{H}_1(v)$ occurs if both $\deg(\mathbf{w}_v) = C + 1$ and $\delta(v, \mathbf{w}_v) = t$;
- $\mathbf{H}_2(v)$ occurs if $\mathbf{w}_v \neq \mathbf{w}_u$ for all $u \in V_0 \setminus \{v\}$;
- $\mathbf{H}(v) = \mathbf{H}_1(v) \wedge \mathbf{H}_2(v)$.

We claim that there AAS exist at least two vertices u, v such that $\mathbf{H}(u)$ and $\mathbf{H}(v)$ both occur.

Let us denote the exact fraction of endpoints of residual degree 1 by $z = \mu_{(A',V')}(1)$, noting that $z = \eta_\mu(1) + o(1)$ by assumption, and let us denote $y = \eta_\mu(C) > 0$. Since the event $\mathbf{H}_1(v)$ occurs if and only if, during BFS, we encounter $t - 1$ endpoints of residual degree 1 followed by an endpoint of residual degree C , we have

$$P[\mathbf{H}_1(v)] = z^{t-1}y(1 + o(1)).$$

Also, it is clear that, for any $u \in V_0 \setminus \{v\}$, we have $P[\mathbf{w}_u = \mathbf{w}_v \mid \mathbf{H}_1(v)] = O(m^{-1})$, and since $\eta_\mu(0) = 0$, then we must have $|V_0| = o(m)$ by assumption, and therefore $P[\mathbf{H}_2(v) \mid \mathbf{H}_1(v)] = 1 - o(1)$, and hence

$$P[\mathbf{H}(v)] = P[\mathbf{H}_1(v) \wedge \mathbf{H}_2(v)] = z^{t-1}y(1 + o(1)).$$

And, since $|V_0| = m^{1-o(1)}$ and $z^{t-1} = m^{\Omega(1)-1}$, it follows that the expected number of vertices for which $\mathbf{H}(v)$ occurs is $m^{\Omega(1)}$.

For any pair of vertices $u, v \in V_0$, we similarly deduce $P[\mathbf{H}(v) \wedge \mathbf{H}(u)] = (z^{t-1}y)^2(1 + o(1))$, and by the second moment method, it follows that with probability $1 - o(1)$ the event $\mathbf{H}(v)$ occurs for $\omega(1)$ vertices $v \in V_0$.

In particular, there AAS exist $u, v \in V$ such that both $\mathbf{H}(u)$ and $\mathbf{H}(v)$ occur. And, any path connecting such vertices u and v must pass through \mathbf{w}_u and \mathbf{w}_v , and therefore

$$\begin{aligned} \delta(u, v) &= \delta(u, \mathbf{w}_u) + \delta(\mathbf{w}_u, \mathbf{w}_v) + \delta(v, \mathbf{w}_v) \\ &= \delta(\mathbf{w}_u, \mathbf{w}_v) + 2 \cdot \left[\left(\frac{1}{-\ln \eta_\mu(1)} - \epsilon_0 \right) \ln n \right]. \end{aligned}$$

Now, let \mathcal{F}^* denote the σ -field induced by all edges connecting v to \mathbf{w}_v for all $v \in V_0$. Note that \mathcal{F}^* is clearly subuniform, and since AAS we have $\delta(v, \mathbf{w}) = O(\ln n)$ for all $v \in V_0$, then $|\mathbf{U}_{\mathcal{F}^*}| = m - o(m)$ a.a.s. Hence, for vertices u, v such that $\mathbf{H}(u)$ and $\mathbf{H}(v)$ both occur, by lemma 5.12 we have

$$P \left[\delta(\mathbf{w}_u, \mathbf{w}_v) > \left(\frac{1}{\ln M_1(\mu)} - \epsilon_1 \right) \ln n \mid \mathcal{F}^* \right] = 1 - o(1)$$

AAS for arbitrary $\epsilon_1 > 0$. The lemma now follows by choosing ϵ_0, ϵ_1 sufficiently small. ■

To complete the lower bound proof of Theorem 5.1, we consider a set $V_0 \subseteq V$ such that $|V_0| = \Theta(n / \ln^2 n)$ and such that V_0 consists exclusively of vertices of degree 1. By Lemma 4.2, WHP a constant fraction of the vertices in V_0 have degree 1 in the V_0 -augmented 2-core \mathbf{G}^{κ, V_0} . Hence, we may apply Lemma 5.13 to \mathbf{G}^{κ, V_0} and conclude that there AAS exist vertices $u, v \in V_0$ such that $\delta(u, v) > (\Delta_\mu - \epsilon) \ln n$ for arbitrary $\epsilon > 0$.

5.3. Other Cases

We now consider generalizations of Theorem 5.1 to graphs of higher minimum degree. The proofs of the generalizations in this section are quite similar to the proof of Theorem 5.1 given above. Hence we only offer proof sketches, which describe how to adapt the original proof to handle particular generalizations.

Recall that the term $\frac{2}{-\ln \eta_\mu(1)}$ in the constant Δ_μ reflects the fact that the longest shortest path will include two long “strands” of length $\frac{1}{-\ln \eta_\mu(1)}$. Moreover, the length of these strands

is determined by the fact that the probability of repeating a neighborhood size during 2-core BFS is $\eta_\mu(1) + o(1)$ for a neighborhood of size 1.

The length of these “strands” changes when the minimum degree is at least 2. First, obviously, the 2-core of a graph with minimum degree at least 2 is simply the entire graph, so $\eta_\mu(1) = \mu(1)$. More significantly, for graphs with minimum degree exactly 2, the initial neighborhood size is also at least 2, so to repeat a neighborhood size, in general we must choose two consecutive endpoints of residual degree 1. As a result, the longest “strands” will have length $\frac{1}{-2 \ln \mu(1)}$. For graphs with minimum degree at least 3, there are no endpoints of residual degree 1, and hence a neighborhood size can only be repeated when cross edges occur. In this case, the “strands” have length $o(\ln n)$, and the diameter is determined by $\frac{1}{\ln M_1(\mu)}$. We now consider these two cases formally in the two theorems stated below.

First, we consider graphs with $\mu(0) = 0$ and $\mu(1) > 0$. In this case, our asymptotic assumptions are not sufficient to determine the diameter; we must also require that the minimum degree is in fact 2, otherwise the presence of vertices of degree 1 may increase the diameter.

Theorem 5.14. *Let (A, V) asymptotically satisfy assumptions 2.1 with limiting residual distribution μ and assume that $\mu(1) > 0$ and $\mu(0) = 0$, and that the minimum degree in (A, V) is 2. Then AAS*

$$\Delta(\mathbf{G}(A, V)) = \left(\frac{1}{-\ln \mu(1)} + \frac{1}{\ln M_1(\mu)} + o(1) \right) \ln n.$$

Proof Sketch. For the upper bound, we note that the first endpoint neighborhood during BFS contains at least two endpoints, and the probability of repeating a neighborhood size of two is $\mu(1)^2 + o(1)$. Using this observation, it is straightforward to adapt the proof of Lemma 5.7 to show that for any v and any $\epsilon_0 > 0$, we have

$$\rho_v(\ln \ln n) < \left(\frac{1}{-2 \ln \mu(1)} + \epsilon_0 \right) \ln n$$

with probability $1 - o(n^{-1})$. The only minor technicality is the fact that it is possible to encounter a situation where the neighborhood size drops to 1 due to cross-edges. However, the probability of such an occurrence is $O(n^{-1})$ for any given BFS iteration, and hence the probability that an endpoint neighborhood of size 1 is either preceded or followed by more than $\epsilon_0 \ln n$ nonempty neighborhoods of size less than $\ln \ln n$ is $o(n^{-1})$.

The lower bound proof can be modified similarly by finding a pair of vertices u, v , such that both u and v have $\left(\frac{1}{-2 \ln \mu(1)} - \epsilon_0 \right) \ln n$ consecutive endpoint neighborhoods of size 2. ■

We now consider graphs with $\mu(0) = \mu(1) = 0$ and with minimum degree 3.

Theorem 5.15. *Let (A, V) asymptotically satisfy assumptions 2.1 with limiting residual distribution μ and assume that $\mu(1) = \mu(0) = 0$, and that the minimum degree in (A, V) is at least 3. Then AAS*

$$\Delta(\mathbf{G}(A, V)) = \left(\frac{1}{\ln M_1(\mu)} + o(1) \right) \ln n.$$

Proof Sketch. In this case, the minimum degree is at least 3, and hence the minimum residual degree is at least 2. It follows that, for a neighborhood of size $r_{v,i} = O(\ln \ln n)$, we have $r_{v,i+1} > r_{v,i}$ with probability $1 - o(1)$ and from this we can deduce $\rho_v(\ln \ln n) = o(\ln n)$ with probability $1 - o(n^{-1})$, and the upper bound follows. And, in this situation the lower bound follows directly from lemma 5.12. ■

6. COMPUTING THE DIAMETER FOR SPECIFIC DEGREE DISTRIBUTIONS

In this section, we describe a technique for computing the value of $\eta_\mu(1)$, which is necessary for computing the constant

$$\Delta_\mu = \frac{2}{-\ln \eta_\mu(1)} + \frac{1}{\ln M_1(\mu)}$$

from Theorem 5.1. We then illustrate by computing the Δ_μ for some special degree distributions.

The *probability generating function (PGF)* of a distribution μ is given by

$$f_\mu(z) = E[z^{X^\mu}] = \sum_{i=0}^{\infty} z^i \mu(i).$$

The PGF is a central tool in the theory of branching processes, and the use of the PGF for analyzing algorithms on random graphs is described in [11].

In particular, for any distribution μ with $M_1(\mu) > 1$, the PGF will exhibit exactly one fixed point in the range $z \in [0, 1)$; let us denote this fixed point by z_μ . A straightforward computation, using the expression for the 2-core residual distribution η_μ given in [12], shows that the PGF of η_μ is given by

$$f_{\eta_\mu}(z) = \frac{f_\mu(z_\mu + (1 - z_\mu)z) - z_\mu}{1 - z_\mu}. \tag{15}$$

Details of this computation are given in [11]. We also note that the distribution η_μ occurs in the theory of branching processes, where η_μ describes the number of surviving children in a survival-conditioned branching process generated by μ [2].

From Eq. (15), we immediately deduce

$$\eta_\mu(1) = f'_{\eta_\mu}(0) = f'_\mu(z_\mu).$$

Note also that the equality of the first moments of μ and η_μ stated in Theorem 4.1 can also be deduced using the PGF by

$$M(\eta_\mu) = f'_{\eta_\mu}(1) = f'_\mu(1) = M(\mu).$$

Hence, in terms of the PGF, we have

$$\Delta_\mu = \frac{2}{-\ln f'_\mu(z_\mu)} + \frac{1}{\ln f'_\mu(1)}. \tag{16}$$

6.1. The Diameter of $G_{n,p}$

We now consider the diameter of the classical random graph $G_{n,p}$, for $p = \frac{d}{n}$, where $d > 1$. It is known (e.g., [17, 18]) that $G_{n,p}$ for $p = \frac{d}{n}$ can be studied using the configuration model by specifying a Poisson degree distribution $\pi_d(i) = \frac{e^{-d} d^i}{i!}$. Also, the Poisson distribution has the property that the residual distribution is the same as the original distribution, so $\mu_{\pi_d} = \pi_d$. Hence, the diameter of $G_{n,d/n}$ for $1 < d = O(1)$ is AAS given by

$$\Delta(G_{n,d/n}) = \Delta_{\pi_d} \ln n + o(\ln n).$$

The PGF for the Poisson distribution π_d has the simple expression $f_{\pi_d}(z) = e^{d(z-1)}$. The fixed point of this function is given by $z_{\pi_d} = \frac{-W(-de^{-d})}{d}$, where the Lambert W -function $W(z)$ is the principal inverse of $f(z) = ze^z$. This gives the following expression for Δ_{π_d} :

$$\Delta_{\pi_d} = \frac{2}{\ln - W(de^{-d})} + \frac{1}{\ln d}. \tag{17}$$

The ‘‘average distance’’ between vertices in the giant component of $G_{n,d/n}$ is given by $\ln n / \ln d = \log_d n$. The actual diameter hence exceeds the average distance by $\frac{2 \ln n}{\ln - W(de^{-d})}$. The qualitative behavior of the diameter of sparse $G_{n,p}$ can be understood by examining the plot in Fig. 1, which shows the ratio of the diameter to the average distance as a function of the average degree.

From Eq. (17), it can be shown that $\Delta_{\pi_d} \ln d \rightarrow 3$ as $d \rightarrow 1$ and $\Delta_{\pi_d} \ln d \rightarrow 1$ as $d \rightarrow \infty$ as the above plot suggests, and it is a simple exercise to derive increasingly accurate asymptotic characterizations of Δ_{π_d} , as in

$$\Delta_{\pi_d} = \frac{1}{\ln d} + \frac{2}{d} + O\left(\frac{\ln d}{d^2}\right) \text{ as } d \rightarrow \infty$$

and so on.

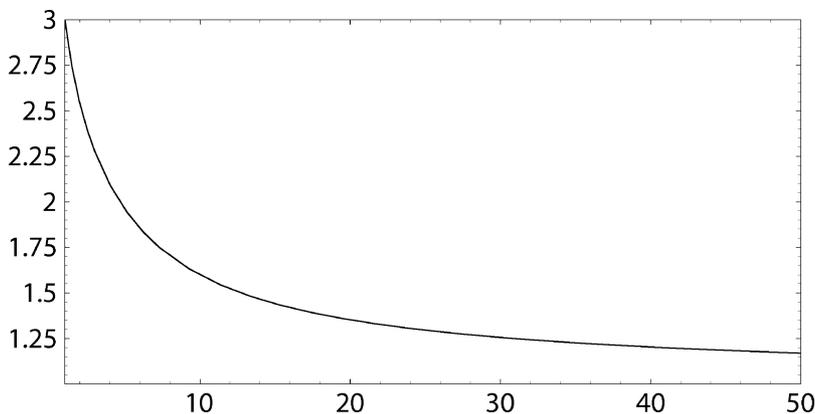


Fig. 1. The function $\frac{\Delta_{\pi_d}}{1/\ln d}$ as a function of the average degree d , which measures the ratio of the diameter to the average distance in the random graph $G_{n,d/n}$.

6.2. The Diameter of Power Law Graphs

We now compute the diameter of random graphs with the so-called “power-law” distributions, in which the number of vertices of degree d is proportional to $d^{-\beta}$ for a given constant β . The degree distribution for the random power law graph model proposed by [1] is

$$\lambda_\beta(i) = \frac{i^{-\beta}}{\zeta(\beta)},$$

with the zeta function $\zeta(\beta) = \sum_{n=1}^\infty n^{-\beta}$. The residual distribution is therefore given by

$$\mu_\beta(i) = \frac{(i + 1)^{1-\beta}}{\zeta(1 - \beta)},$$

and the corresponding PGF is

$$f_{\mu_\beta}(z) = \frac{\text{Li}(\beta - 1, z)}{z\zeta(\beta - 1)},$$

with the polylogarithm $\text{Li}(x, z) = \sum_{n=1}^\infty \frac{z^n}{n^x}$.

Note that both the average degree and the average residual degree decrease as β increases. For $\beta > 2$, the average degree is $O(1)$, so the graph is sparse, and for $\beta > 3$, the average residual degree is $O(1)$ so the “average distance” is $\Omega(\ln n)$. In addition, the giant component phase transition for this model occurs at the point $\beta = 3.479\dots$, where $M_1(\mu_\beta) = 1$, so for $\beta > 3.479\dots$, the graph no longer exhibits a giant component.

Hence, the results of this paper are relevant for the range $2 < \beta < 3.479$. Moreover, for $2 < \beta \leq 3$, the “average distance” is $o(\ln n)$, and therefore the leading constant for the diameter is simply given by $\frac{2}{-\ln f'_\mu(z_\mu)}$. The leading constant for the diameter of a random power law graphs is plotted in Fig. 2 as a function of β , along with the function $\frac{2}{-\ln f'_\mu(z_\mu)}$ for $\beta > 3$.

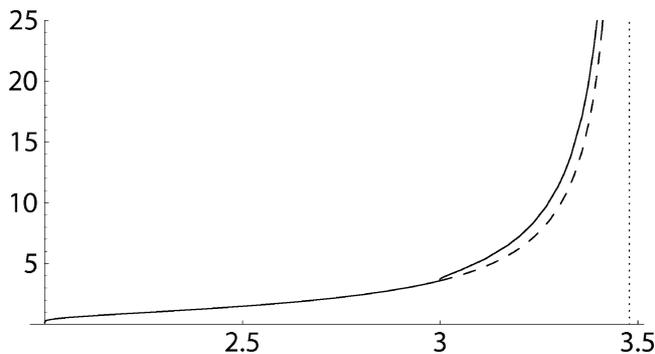


Fig. 2. The leading constant Δ_μ for the diameter of random graphs generated by the power law model introduced in [1], as a function of β . For $\beta < 3$, the average distance is $o(\ln n)$, and therefore the leading constant reflects only the contribution of long isolated paths. For $\beta > 3$, the dashed curve shows the contribution of long isolated paths, given by $\frac{2}{-\ln f'_\mu(z_\mu)}$, while the vertical distance between the solid and dashed curves represents the average distance. The dotted vertical line shows the phase transition, which occurs at $\beta = 3.479\dots$

APPENDIX

Proofs of Selected Lemmas

Lemma A.1 (Lemma 2.4). *Let μ be any distribution on \mathbb{Z}^* . Then for all $\epsilon > 0$, there exists a $\delta > 0$ such that the following statement holds.*

For any endpoint partition (A, V) satisfying $\|\mu_{(A,V)} - \mu\|_1 < \delta$, and for any subset $V' \subseteq V$ satisfying $|V'| < \delta|A|$, we have $|A(V')| < \epsilon|A|$.

Proof. Choose an integer J sufficiently large that $\sum_{j \geq J} \mu(j) < \epsilon/2$, and partition the set A into disjoint subsets $A_{\geq J}, A_{< J}$ according to residual degree:

$$A_{\geq J} = \{a \in A : \text{res}_{(A,V)}(a) \geq J\}$$

$$A_{< J} = \{a \in A : \text{res}_{(A,V)}(a) < J\}.$$

Now, let $B = A(V')$, and note that since $\|\mu - \mu_{(A,V)}\|_1 < \delta$, it follows that

$$\frac{|A_{\geq J}|}{|A|} = \sum_{j \geq J} \mu_{(A,V)}(j) < \epsilon/2 + \delta$$

and therefore $|B \cap A_{\geq J}| \leq (\epsilon/2 + \delta)|A|$. On the other hand, each endpoint in $A_{< J}$ has residual degree at most $J - 1$, and therefore belongs to a vertex of degree at most J . Therefore, each vertex in V' contributes at most J endpoints to $B \cap A_{< J}$, and thus

$$|B \cap A_{< J}| \leq J|V'| < J\delta|A|.$$

The lemma follows by choosing $\delta < \frac{\epsilon}{2(J+1)}$. ■

Lemma A.2 (Lemma 2.6). *For any distribution μ on $\mathbb{Z}^* \cup \{-1\}$ and any $\epsilon > 0$, there exists a unique distribution $\mu_{[\epsilon]}$ on $\mathbb{Z}^* \cup \{-1\}$ satisfying both $\|\mu - \mu_{[\epsilon]}\|_1 \leq \epsilon$ and $\nu \triangleright \mu_{[\epsilon]}$ for all distributions ν on $\mathbb{Z}^* \cup \{-1\}$ with $\|\mu - \nu\|_1 \leq \epsilon$.*

Proof. We denote the cumulative distribution function of μ by

$$F_\mu(j) = \mathbb{P}[\mathbf{X}_\mu \leq j] = \sum_{i=-1}^j \mu(i)$$

for $j \in \mathbb{Z}^* \cup \{-1\}$. We note the function F_μ completely specifies the distribution μ . In addition, for distributions μ, ν on $\mathbb{Z}^* \cup \{-1\}$, we have $\mu \triangleright \nu$ if and only if $F_\mu(j) \leq F_\nu(j)$ for all j .

We define the truncated distribution in by its distribution function

$$F_{\mu_{[\epsilon]}}(j) = \max\{F_\mu(j) + \epsilon/2, 1\}$$

for all $j \in \mathbb{Z}^* \cup \{-1\}$.

It is easy to verify that $\|\mu - \mu_{[\epsilon]}\|_1 \leq \epsilon$. And, for any pair of distributions μ, ν and any j ,

$$\|\mu - \nu\|_1 = \sum_{i=-1}^j |\mu(i) - \nu(i)| + \sum_{i=j+1}^{\infty} |\mu(i) - \nu(i)| \leq 2|F_\nu(j) - F_\mu(j)|.$$

It follows that if $\|\mu - \nu\|_1 \leq \epsilon$, then $F_\nu(j) \leq F_\mu(j) + \epsilon/2$ for all j , and hence $\nu \triangleright \mu_{[\epsilon]}$. ■

Lemma A.3 (Lemma 2.7). *Let $\mathbf{y}_1, \dots, \mathbf{y}_t$ be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \dots, \mathcal{F}_t$, and let $\mathbf{Y}_t = \sum_{i=1}^t \mathbf{y}_i$. Consider a distribution μ such that $\mathcal{D}[\mathbf{y}_{s+1} | \mathcal{F}_s] \triangleright \mu$ always for all $0 \leq s \leq t - 1$. Then \mathbf{Y}_t dominates in distribution the sum of t independent μ -distributed random variables.*

Proof. By induction, it suffices to prove that for a pair of random variables $\mathbf{y}_1, \mathbf{y}_2$, and for distributions μ_1, μ_2 , if that $\mathbf{y}_1 \triangleright \mu_1$, and $\mathcal{D}[\mathbf{y}_2 | \mathbf{y}_1] \triangleright \mu_2$ always then

$$\mathbf{y}_1 + \mathbf{y}_2 \triangleright \mathbf{X}_{\mu_1} + \mathbf{X}_{\mu_2}$$

where as usual \mathbf{X}_{μ_1} and \mathbf{X}_{μ_2} are independent.

In this case, for any y such that $P[\mathbf{y}_1 = y] > 0$, and for any s , we have

$$P[\mathbf{y}_1 + \mathbf{y}_2 > s | \mathbf{y}_1 = y] \geq P[\mathbf{X}_{\mu_2} > s - y] = 1 - F_{\mu_2}(s - y),$$

where as above F_{μ_2} denotes the cumulative distribution function. Hence,

$$P[\mathbf{y}_1 + \mathbf{y}_2 > s] = E[P[\mathbf{y}_1 + \mathbf{y}_2 > s | \mathbf{y}_1]] \geq E[1 - F_{\mu_2}(s - \mathbf{y}_1)],$$

and since F_{μ_2} is nondecreasing, then

$$P[\mathbf{y}_1 + \mathbf{y}_2 > s] \geq E[1 - F_{\mu_2}(s - \mathbf{X}_{\mu_1})] = P[\mathbf{X}_{\mu_1} + \mathbf{X}_{\mu_2} > s].$$

■

Lemma A.4 (Lemma 2.8). *Consider a sequence of Bernoulli random variables $\mathbf{z}_1, \mathbf{z}_2, \dots$, adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \mathcal{F}_2, \dots$, and let $\mathbf{Z}_t = \sum_{i=1}^t \mathbf{z}_i$. Consider a stopping time τ and a constant $0 < p < 1$ such that*

$$P[\mathbf{z}_{t+1} = 1 | t < \tau, \mathcal{F}_t] \geq p$$

always for all $t \geq 0$. Then, for any $t, r \geq 0$, we have

$$P[\tau > t \wedge \mathbf{Z}_t \leq r] \leq P[\mathbf{X}_{\text{Bin}_t, p} \leq r].$$

Proof. We define a second adapted sequence of Bernoulli trials

$$\mathbf{y}_t = \begin{cases} \mathbf{z}_t & \text{if } t \leq \tau \\ 1 & \text{if } t > \tau, \end{cases}$$

and we let $\mathbf{Y}_t = \sum_{i=1}^t \mathbf{y}_i$. Clearly, $P[\mathbf{y}_{t+1} = 1 | \mathcal{F}_t] \geq p$ always, and thus by lemma 2.7, $\mathbf{Y}_t \triangleright \text{Bin}_{t,p}$ for all t . Moreover, since $\mathbf{Y}_t = \mathbf{Z}_t$ for $t \leq \tau$, then

$$P[\tau > t \wedge \mathbf{Z}_t \leq r] \leq P[\mathbf{Y}_t \leq r] \leq P[\mathbf{X}_{\text{Bin}_t, p} \leq r].$$

■

Lemma A.5 (Lemma 2.9). *Let μ be a distribution on $\mathbb{Z}^* \cup \{-1\}$, let $r \in \mathbb{Z}^*$ and $0 \leq \epsilon \leq 1$, and let \mathbf{Y} be a \mathbb{Z} -valued random variable such $0 \leq \mathbf{Y} \leq r$ always and such that $\mathbf{Y} \stackrel{d}{\triangleright} \mathbf{X}_{\text{Bin}_{r, 1-\epsilon/2}}$. Then*

$$\sum_{i=1}^{\mathbf{Y}} \mathbf{X}_{\mu, i} \stackrel{d}{\triangleright} \sum_{j=1}^r \mathbf{X}_{\mu_{[\epsilon]j}}.$$

Proof. We shall in fact prove the stronger statement

$$\sum_{i=1}^Y (\mathbf{X}_{v,i} + 1) \triangleright^d \sum_{j=1}^r (\mathbf{X}_{v_{[\epsilon]j}} + 1).$$

Note that since $\mathbf{X}_v + 1 \geq 0$ always, and since $Y \triangleright^d \mathbf{X}_{\text{Bin}_{r,1-\epsilon/2}}$, we have

$$\sum_{i=1}^Y (\mathbf{X}_{v,i} + 1) \triangleright^d \sum_{j=1}^{\mathbf{X}_{\text{Bin}_{r,1-\epsilon/2}}} (\mathbf{X}_{v,j} + 1).$$

It therefore suffices to prove

$$\sum_{i=1}^{\mathbf{X}_{\text{Bin}_{r,1-\epsilon/2}}} (\mathbf{X}_{v,i} + 1) \triangleright^d \sum_{j=1}^r (\mathbf{X}_{v_{[\epsilon]j}} + 1). \tag{18}$$

Now, let $\mathbf{I}_{1-\epsilon/2}$ denote a Bernoulli random variable with $\mathbb{P}[\mathbf{I}_{1-\epsilon/2} = 1] = 1 - \epsilon/2$ and $\mathbb{P}[\mathbf{I}_{1-\epsilon/2} = 0] = \epsilon/2$. Since $\mathbf{X}_{\text{Bin}_{r,1-\epsilon/2}} \stackrel{d}{=} \sum_{j=1}^r \mathbf{I}_{1-\epsilon/2,j}$, where as usual the $\mathbf{I}_{1-\epsilon/2,j}$ are independent, we can express the left hand side of the above expression as

$$\sum_{i=1}^{\mathbf{X}_{\text{Bin}_{r,1-\epsilon/2}}} (\mathbf{X}_{v,i} + 1) \stackrel{d}{=} \sum_{j=1}^r \mathbf{I}_{1-\epsilon/2,j} (\mathbf{X}_{v,j} + 1). \tag{19}$$

It is clear that

$$\|\mathfrak{D}[\mathbf{I}_{1-\epsilon/2}(\mathbf{X}_v + 1)] - \mathfrak{D}[\mathbf{X}_v + 1]\|_1 \leq \epsilon,$$

and therefore

$$\mathbf{I}_{1-\epsilon/2}(\mathbf{X}_v + 1) \triangleright^d \mathbf{X}_{v_{[\epsilon]}} + 1 \tag{20}$$

by Lemma 2.6. Equation (18) now follows from (20) and (19), and the proof is complete. ■

Lemma A.6 (Lemma 3.5, [10, 15]). *Let μ be any distribution on $\mathbb{Z}^* \cup \{-1\}$. Then for all $M < M_1(\mu)$ there exists a $C > 0$ such that for all $n > 0$,*

$$\mathbb{P}\left[\sum_{i=1}^n \mathbf{X}_{\mu,i} \leq Mn\right] < e^{-Cn}.$$

Proof. Let us denote the cumulant generating function of the random variable $-\mathbf{X}_\mu$ by

$$\Lambda(z) = \ln \mathbb{E}[e^{-z\mathbf{X}_\mu}],$$

and we note that since μ has support in $[-1, \infty)$, then $\Lambda(z)$ is convergent for all $z \geq 0$. Also, let $\mathbf{Y} = \sum_{i=1}^n \mathbf{X}_{\mu,i}$ and note that $\ln \mathbb{E}[e^{-z\mathbf{Y}}] = n \ln \mathbb{E}[e^{-z\mathbf{X}_\mu}] = n\Lambda(z)$.

For any $z \geq 0$, by Markov's inequality we have

$$\mathbb{P}[\mathbf{Y} \leq Mn] = \mathbb{P}[e^{-z\mathbf{Y}} \geq e^{-zMn}] \leq \mathbb{E}[e^{-z\mathbf{Y}}]e^{zMn} = e^{n(\Lambda(z)+zM)}.$$

Since $\lim_{z \rightarrow 0^+} \frac{\Lambda(z)}{z} = \Lambda'(0) = -M_1(\mu)$, then for z_0 sufficiently small we have $\Lambda(z_0) < -z_0M$, and hence by setting $C = \Lambda(z_0) + z_0M < 0$, we conclude

$$P \left[\sum_{i=1}^n \mathbf{X}_{\mu,i} \leq Mn \right] = P[\mathbf{Y} \leq Mn] \leq e^{n(\Lambda(z_0) + z_0M)} = e^{-Cn}.$$

Lemma A.7 (Lemma 4.2). *Consider an endpoint partition (A, V) as in Theorem 4.1 and consider a subset $W \subseteq V$ such that $|A(W)| = o(m/\ln m)$. Then WHP $|\mathbf{A}^{\kappa,W} \setminus \mathbf{A}^\kappa| = o(m)$.*

If, in addition, $|A(W)| = m^{\Omega(1)}$ then AAS a constant fraction of the endpoints in $A(W)$ belong to $\mathbf{A}^{\kappa,W}$.

Proof. As noted in the main text, this lemma follows almost immediately from any proof of the 2-core result in either of [9, 12, 14]. We give a brief self-contained proof that relies only on the statement of Theorem 4.1 and the results of [17, 18].

For the first statement, we first show that for any set W and any $w \in W$, we have

$$P[|\mathbf{A}^{\kappa,W}| - |\mathbf{A}^{\kappa,W \setminus \{w\}}| = O(\ln n)] = 1 - o(n^{-1}).$$

The first statement of the lemma then follows by removing one vertex from W at a time.

We note that if $\deg^{\kappa,W}(w) \geq 2$ then $\mathbf{A}^{\kappa,W} = \mathbf{A}^{\kappa,W \setminus \{w\}}$, so we consider the case where $\deg^{\kappa,W}(w) \geq 1$. In this case $\mathbf{A}^{\kappa,W} \setminus (\mathbf{A}^{\kappa,W \setminus \{w\}})$ will contain all endpoints on the path connecting w to the nearest vertex of any degree other than 2 in $\mathbf{G}^{\kappa,w}$. Moreover, this path can be found by performing BFS, and the length of the path is equal to the number of consecutive endpoints of residual degree 1. It follows by Theorem 4.1 and by the fact that $\mathbf{A}^{\kappa,W} \supseteq \mathbf{A}^\kappa$ that WHP, the fraction of endpoints of residual degree 1 in $\mathbf{A}^{\kappa,W}$ is $1 - \Omega(1)$. Hence, there exists a constant C such that the probability choosing more than $C \ln n$ such endpoints consecutively is $o(n^{-1})$.

For the second statement of the lemma, it is only necessary to compare that the size of the 2-core computed in [9, 12, 14] and the size of the giant component, computed in [17, 18]. Specifically, the giant component contains $\Omega(m)$ more endpoints and $\Omega(n)$ more vertices than the 2-core. It follows that a constant fraction of the vertices (and endpoints) which are not in the 2-core still belong to the giant component and therefore are connected by a path to the 2-core, and hence if the set $A(W)$ is sufficiently large (i.e., $m^{\Omega(1)}$), then by standard means it follows easily that AAS a constant fraction of these endpoints will be connected by a path to the 2-core, and thus belong to $\mathbf{A}^{\kappa,W}$. ■

Lemma A.8 (Lemma 5.2). *Consider disjoint subsets $S_1, S_2 \subseteq A$ such that both $|S_1|, |S_2| > m^{1/2} \ln^2 m$. Then*

$$P[\vec{\mathbf{E}}(S_1) \cap S_2 = \emptyset] = m^{-\omega(1)}.$$

Proof. First, note that any given endpoint in S_1 matches into S_2 with probability at least $|S_2|/m \geq m^{-1/2}$. Now, if a particular endpoint in S_1 does not match into S_2 , it may match to another endpoint in S_1 . Nevertheless, if we sequentially expose the matches all of the endpoints in S_1 , there are at least $|S_1|/2$ chances to find a connection to S_2 . The probability that no connection is found is therefore at most

$$(1 - m^{-1/2})^{\frac{m^{1/2} \ln^2 m}{2}} = \left(\left(1 - \frac{1}{m^{1/2}} \right)^{m^{1/2}} \right)^{\frac{\ln^2 m}{2}} = \left(\frac{1}{e - o(1)} \right)^{\Omega(\ln^2 m)} = m^{-\Omega(\ln m)} = m^{-\omega(1)}.$$

REFERENCES

- [1] W. Aiello, F. Chung, and L. Lu, A random graph model for massive graphs. In Proceedings of the 32nd Annual ACM Symposium on Theory of Computation (STOC), 2000, pp. 171–180.
- [2] K. B. Athreya and P. E. Ney, Branching processes, Springer-Verlag, Berlin, 1972.
- [3] P. Billingsley, Probability and measure, John Wiley & Sons, New York, 1995.
- [4] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J Combin* 1 (1980), 311–316.
- [5] B. Bollobás, Random graphs, Academic Press, London, 1985.
- [6] B. Bollobás and W. F. de la Vega, The diameter of random regular graphs, *Combinatorica* 2 (1982), 125–134.
- [7] B. Bollobás, S. Janson, and O. Riordan, The phase transition in inhomogeneous random graphs, *Random Struct Algorithms* 31 (2007), 3–122.
- [8] F. Chung and L. Lu, The diameter of random sparse graphs, *Adv Appl Math*, 26 (2001), 257–279.
- [9] C. Cooper, The cores of random hypergraphs with a given degree sequence, *Random Struct Algorithms* 25 (2004), 353–375.
- [10] A. Dembo and O. Zeitouni, Large deviations techniques and applications, Springer-Verlag, New York, 1998.
- [11] D. Fernholz, Sparse random graphs: Methods, structure, and heuristics, PhD Thesis, Department of Computer Science, University of Texas at Austin, December 2007.
- [12] D. Fernholz and V. Ramachandran, Cores and connectivity in sparse random graphs. Technical Report UTCS TR04-13, Department of Computer Science, University of Texas at Austin, 2004.
- [13] A. Frieze and S. Suen, On the independence number of random cubic graphs, *Random Struct Algorithms* 5 (1994), 649–664.
- [14] S. Janson and M. J. Luczak, A simple solution to the k -core problem, *Random Struct Algorithms* 30 (2007), 50–62.
- [15] O. Kallenberg, Foundations of modern probability, Springer-Verlag, New York, 2002.
- [16] L. Lu, The diameter of random massive graphs, In Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA), 2001, pp. 912–921.
- [17] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct Algorithms* 6 (1995), 161–179.
- [18] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combin Probab Comput* 7 (1998), 295–305.