# Phase transitions of extremal cuts for the configuration model

Souvik Dhara[*1], Debankur Mukherjee[†1], Subhabrata Sen[‡2]

[1]*Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands*
[2]*Department of Statistics, Stanford University, USA*

November 10, 2016

### Abstract

The $k$-section width and the Max-Cut for the configuration model are shown to exhibit phase transitions according to the values of certain parameters of the asymptotic degree distribution. These transitions mirror those observed on Erdős-Rényi random graphs, established by Luczak and McDiarmid (2001), and Coppersmith et al. (2004), respectively.

## 1 Introduction

Graph cut problems have a very rich history in Combinatorics and Theoretical Computer Science. Given a graph $G = (V, E)$, the $k$-section problem seeks to partition the vertices $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_k$ into $k$ equal sets (or differing by at most 1) such that the number of edges between the distinct sets is minimized. The minimum number of cross edges $w_k(G)$ thus obtained is referred to as the $k$-section width. The related Max-Cut problem seeks to divide the vertices into two sets (not necessarily equal) such that the number of edges between the two sets is maximized. These graph-partitioning problems are extremely important for numerous practical applications in network optimization, VLSI circuit design, computational geometry, and statistical physics [8, 9, 12, 17, 31, 34, 36]. On the other hand, from the perspective of Theoretical Computer Science, these problems are computationally hard, and even approximating the Max-Cut up to a constant factor is NP-hard [14, 16, 19, 32]. The study of these problems in the average case is also of considerable interest — mainly motivated by a desire to understand various graph partitioning heuristics. Problem instances are usually chosen to be the Erdős-Rényi random graph, or the random regular graph. An Erdős-Rényi random graph $\mathrm{ER}_n(d/n)$ is constructed on $n$ vertices, where any two vertices share an edge with probability $d/n$, independently of each other. A $d$-regular random graph is drawn uniformly at random from the space of all $d$-regular graphs on $n$ vertices. We note that these graph ensembles are *sparse*, in that typical

---

[*]s.dhara@tue.nl

[†]d.mukherjee@tue.nl

[‡]ssen90@stanford.edu

graphs on $n$ vertices have order $n$ edges and the degree of a typical vertex is of the constant order. See [4, 21, 22, 30] for a detailed review of the properties of these random graphs.

A crucial point to note in this context is that both sparse Erdős-Rényi and random regular graph ensembles lead to homogeneous instances, in the sense that any two vertices share an edge with equal probability. This is very different from the instances actually encountered in practical applications. Real networks are extremely inhomogeneous, and often display certain characteristic features, such as a power-law decay in the tail of the degree distribution [1, 2, 13, 22, 37]. Thus, it is of natural interest to study the behavior of the "extremal" cuts for graphs with more general degree distributions. The configuration model provides a canonical scheme for generating uniform random graphs with *any* prescribed degree sequence. This model is thus attractive for studying real-world networks, and analysis of its structural properties have attracted considerable attention in recent years [23, 24, 26, 28, 29, 35]. It is worthwhile to mention that despite the presence of very high degree vertices, a plethora of modern research remarkably conveys a qualitatively similar behavior of various statistics in this model to Erdős-Rényi random graphs, confirming empirical evidences.

Both $k$-section width and Max-Cut undergo phase transitions on the sparse Erdős-Rényi random graphs. These transitions reflect certain structural characteristics of the underlying graphs. Consider the $k$-section width problem for $\mathrm{ER}_n(d/n)$, with $k = 2$. For $d < 2\ln(2)$, the minimum bisection is exactly 0 with high probability, while for $d > 2\ln(2)$, the minimum bisection is of the order $n$, with high probability [33]. The Max-Cut also undergoes a phase transition; for $d < 1$, the difference between the total number of edges and the Max-Cut is of the constant order, while it is of the order $n$ for $d > 1$ [10]. The distribution of the Max-Cut within the critical window is analyzed by Daudé et al. [11], while the critical behavior of the bisection width is largely unknown.

In this paper, we initiate a study of these *extremal* cuts for the configuration model, and explore similar phase transition phenomena. The main takeaway of our results is that the phase transitions for the extremal cuts are robust, and are present in a large class of random graphs, viz. configuration models with finite second moment degrees. This emphasizes that in the class of non-spatial random graphs these phase transition phenomena are not intimately dependent on the precise model details, but are determined by the component sizes and the structures of the typical local neighborhoods. Technically, the proofs in the Erdős-Rényi case crucially utilize the independence and homogeneity in the model — while we rely on the recent insights about the structure of the configuration model [23, 26, 28, 29] to establish our results. We also prove several novel structural properties of the connected components (see Sections 4.2 and 4.3). Among many other intermediate results, we show that the largest connected component consists of a well-connected 2-core (Lemma 4.5) and several *thin* hanging trees (Lemma 4.6), and most of the connected components except the largest are finite (Lemma 4.10). Furthermore, we obtain that when $\nu > 1$, the largest connected component is *stable*, in the sense that $\Theta(n)$ edges *must* be deleted to separate out any $\Theta(n)$ vertices (Proposition 4.3). The latter notion is useful to study the stability of the largest connected component subject to intelligent attacks (edge deletion) on networks.

The rest of the paper is organized as follows: Section 2 formally introduces the configuration model along with the assumptions on the underlying degree sequence and summarizes certain preliminary properties of this model. Section 3 states the main results of this paper and offers several heuristics and key insights. The proofs of the main results are included in Sections 4 and 5.

2

# 2 Preliminaries

**The configuration model.** Consider a degree sequence $d = (d_1, d_2, \ldots, d_n)$ on the vertex set $[n] = \{1, 2, \ldots, n\}$. Equip vertex $j$ with $d_j$ stubs or *half-edges*. Two half-edges create an edge once they are paired. Therefore, initially there are $\ell_n = \sum_{i \in [n]} d_i$ half-edges. Pick any one half-edge and pair it with a uniformly chosen half-edge from the remaining unpaired half-edges. Keep repeating the above procedure until all the unpaired half-edges are exhausted. The random graph constructed in this way is called the configuration model, and will henceforth be denoted by $\mathrm{CM}_n(d)$.

Note that the graph constructed by the above procedure may contain self-loops and multiple edges. It can be shown that conditionally on $\mathrm{CM}_n(d)$ being simple, the law of such graphs is uniform over all possible simple graphs with degree sequence $d$ (cf. [22, Proposition 7.7], [27]).

A vertex chosen uniformly at random from the vertex set $[n]$, independently of the graph $\mathrm{CM}_n(d)$ is called a *typical* vertex. Let $D_n$ be the degree of a typical vertex. Throughout this paper we assume the following:

**Assumption 1.** We assume that the degree sequence is such that

a. $D_n \xrightarrow{d} D$ (*weak convergence of the degree of a typical vertex*);

b. $\mathbb{E}[D_n] \to \mathbb{E}[D]$, and $\mathbb{E}[D_n^2] \to \mathbb{E}[D^2]$ (*moment assumptions*);

c. $\mathbb{P}(D = 1) > 0$ (*positive proportion of degree one vertices*).

Like most other sparse random graph models, $\mathrm{CM}_n(d)$ exhibits a phase transition in terms of the size of its largest connected component, and this has been studied extensively in [29, 35]. The phase transition occurs when the value of the parameter

$$\nu := \frac{\mathbb{E}\left[D(D-1)\right]}{\mathbb{E}\left[D\right]} \tag{2.1}$$

exceeds one (cf. [29]). More precisely, let $g_D(x) := \mathbb{E}[x^D]$ be the probability generating function of $D$, and let $\xi$ be the unique nonzero solution to the equation $g'_D(x) = \mathbb{E}[D]x$. Define

$$\eta = 1 - g_D(\xi). \tag{2.2}$$

Denote the $i^{\mathrm{th}}$ largest component of $\mathrm{CM}_n(d)$ by $\mathscr{C}_{(i)}$, $i \geq 1$. Then the following theorem characterizes the asymptotic proportion of vertices in each component:

**Theorem 2.1** ([29, Theorem 2.3]). *Consider* $\mathrm{CM}_n(d)$ *satisfying* Assumption 1*. Then,*

(i) $|\mathscr{C}_{(1)}|/n \xrightarrow{\mathbb{P}} \eta$, *as* $n \to \infty$, *where* $\eta$ *is as in* (2.2)*. Further,* $\eta > 0$ *if and only if* $\nu > 1$.

(ii) *Furthermore,* $|\mathscr{C}_{(i)}|/n \xrightarrow{\mathbb{P}} 0$, *as* $n \to \infty$, *for all* $i \geq 2$.

**Notation.** We often write $\mathbb{E}[D]$, the (asymptotic) expected value of the degree of a typical vertex as $\mu$. For any graph $G$, the $k$-section width and Max-Cut are denoted by $w_k(G)$ and $\mathsf{MaxCut}(G)$, respectively. The degree of a vertex $v$ is denoted by $d_v$, and the number of vertices of degree $k$ by $n_k$, $k \geq 0$. If two vertices $u$ and $v$ share an edge, then we write

3

$u \longleftrightarrow v$. For a nonempty subset $U \subseteq [n]$ of vertices, the neighborhood (or 1-neighborhood) is defined as

$$\mathcal{N}[U, 1] := U \cup \{v \in [n] : u \longleftrightarrow v \text{ for some } u \in U\},$$

and the $r$-neighborhood is defined as $\mathcal{N}[U, r] := \mathcal{N}[\mathcal{N}[U, r - 1], 1]$, $r > 1$. For any subset of vertices $A$, we denote the half-edges incident to the vertices in $A$ by $S(A)$, and the number of edges between $A$ and $A^c$ by $E(A, A^c)$. For any integer $m \geq 1$, we denote $(2m)!! := (2m - 1)(2m - 3) \cdots 1$. All the limiting statements should be understood as $n \to \infty$, unless specified otherwise. For a sequence of probability measures $(\mathbb{P}_n)_{n \geq 1}$, the sequence of events $(\mathcal{E}_n)_{n \geq 1}$ is said to hold with high probability if $\mathbb{P}_n(\mathcal{E}_n) \to 1$. We use the usual Bachmann-Landau notations $o(\cdot)$, $O(\cdot)$, and $\Theta(\cdot)$ to write asymptotic comparisons. For two sequences of random variables $(X_n)_{n \geq 1}$, and $(Y_n)_{n \geq 1}$, we write $X_n = o_{\mathbb{P}}(Y_n)$ to denote that $X_n/Y_n \xrightarrow{\mathbb{P}} 0$.

# 3 Main results

In this section we state the main results of this paper, and discuss several heuristics.

**Theorem 3.1** (Phase transition of the $k$-section width). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying* Assumption 1, *and let* $k \geq 2$ *be an integer. Then* $w_k(\mathrm{CM}_n(\boldsymbol{d}))$ *exhibits a phase-transition around* $\eta = 1/k$. *More precisely,*

  (i) *if* $\eta < 1/k$, *then with high probability* $w_k(\mathrm{CM}_n(\boldsymbol{d})) \leq k/2$;

  (ii) *if* $\eta > 1/k$, *then there exists* $\zeta > 0$, *such that with high probability* $w_k(\mathrm{CM}_n(\boldsymbol{d})) > \zeta n$.

**Remark 1.** [33, Theorem 1] establishes that the min $k$-section width is exactly zero below a critical threshold given by $\eta = 1/k$. This holds for the Erdős-Rényi case due to the natural presence of many isolated vertices. For a general configuration model, this is not necessarily true, and therefore, Theorem 3.1 (i) is indeed the best possible result that we can hope for in this case. In particular, if we assume the presence of a positive fraction of isolated vertices in the degree sequence, then using Lemma 4.1 below, we recover the same result as in [33].

   Theorem 3.1 is proved in Section 4. This result is comparable to [33, Theorem 1], established in the context of Erdős-Rényi random graphs. The proof for the Erdős-Rényi case makes crucial use of the fact that the edge occupancies are independent and identically distributed — a feature that is absent in this case. The proof in this paper, on the other hand, is more robust, and depends on a clear understanding of the local neighborhood structure in these random graphs. Roughly speaking, when $\eta < 1/k$, the strategy is to distribute all the components of size at least 3 among $k$ partitions as evenly as possible, and then to add the components of size at most 2 to balance the partitions. Since the size of the largest component is smaller than $n/k$ and the other components are very small ($o(n)$) in size, a $k$-partition can be made using the components of size at least 3, with at most $n/k$ vertices in each part. Because there are sufficiently many components of size at most 2, these can be used to balance the partitions. The latter step results in at most $k/2$ cross edges between the partitions. The above proof outline for the subcritical case is formalized in Subsection 4.1. Alternatively, when $\eta > 1/k$, the size of the largest connected component is more than $n/k$. Therefore, in order to split the graph into $k$ equal partitions,

the largest component must be split into at least two (possibly unequal) halves, each containing a positive proportion of vertices, and from the structural properties of the largest component, we show that with high probability this creates $\Theta(n)$ cross edges. The proof for the supercritical case is provided in Subsection 4.2.

We continue to describe our results for $\mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d}))$:

**Theorem 3.2** (Phase transition of the Max-Cut). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying* Assumption 1. *Then* $\mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d}))$ *admits a phase transition around* $\nu = 1$. *More precisely,*

(i) *if* $\nu < 1$, *then as* $n \to \infty$,

$$\frac{\ell_n}{2} - \mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d})) \xrightarrow{\mathrm{d}} Z \sim \mathrm{Poisson}\left(\frac{1}{4}\ln\left(\frac{1+\nu}{1-\nu}\right)\right),$$

(ii) *if* $\nu > 1$, *then there exists* $c^\star(\mu) > 0$, *such that for any* $c > c^\star(\mu)$, *with high probability,*

$$\mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d})) \leq n\left(\frac{\mu}{4} + c\sqrt{\mu}\right),$$

The proof of Theorem 3.2 is included in Section 5. We first note that $\ell_n/2$ is precisely the number of edges present in the graph. The difference between the total number of edges and the MaxCut is often referred to as the *distance to bipartiteness* of a graph, which measures the minimum number of edges to be deleted to transform the graph into a bipartite one. Thus, Theorem 3.2 establishes the phase transition for the distance to bipartiteness in a wide class of random graphs with prescribed degree sequence. The heuristic behind this phase-transition is that when $\nu < 1$, $\mathrm{CM}_n(\boldsymbol{d})$ is *roughly* a collection of trees and a *finite* number of unicyclic components. The trees do not contribute any edge to $\mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d}))$ at all and the unicyclic components with an odd cycle contributes at most one to the $\mathrm{MaxCut}(\mathrm{CM}_n(\boldsymbol{d}))$. On the other hand, when $\nu > 1$, this is no longer true, and any partition must leave $\Theta(n)$ edges uncut. Results analogous to Theorem 3.2 (i) and (ii) were established for Erdős-Rényi random graphs by Daudé at al. [11] and Coppersmith et al. [10], respectively.

**Remark 2.** It was shown in [25] that under Assumption 1, the probability of the graph being simple is bounded away from zero. Thus the results in Theorems 3.1 and 3.2 hold for the uniformly chosen simple graph with a prescribed degree sequence. Thus, all the results proved in the paper are true also for $\mathrm{ER}_n(d/n)$, as well as the generalized random graphs under appropriate conditions [22, Theorem 6.15] on weight sequence $\boldsymbol{w}$. In fact the results can be extended to more general inhomogeneous random graph models (cf. [22, Theorem 6.18]).

**Remark 3.** As we will see in the proof of Theorem 3.2, the value of $c^*(\mu)$ is given by

$$c^\star(\mu) = \inf_{c>0}\left\{c : \left(\frac{1}{4} - \frac{c}{\sqrt{\mu}}\right)^{-\left(\frac{1}{4} - \frac{c}{\sqrt{\mu}}\right)}\left(\frac{1}{4} + \frac{c}{\sqrt{\mu}}\right)^{-\left(\frac{1}{4} + \frac{c}{\sqrt{\mu}}\right)} < 2^{-\frac{1}{\mu}+1}\right\}.$$

Numerically it can be seen that $c^\star(\mu)$ increases with $\mu$, and is approximately equal to $\sqrt{\ln(2)}/2$ for large enough $\mu$. Figure 1 shows the values of $c^\star(\mu)$ for $3 \leq \mu \leq 50$. Notice that even for $\mu$-values as low as 30, $c^\star(\mu)$ is sufficiently close to $\sqrt{\ln(2)}/2$. This value agrees with the upper bound of Max-Cut for the Erdős-Rényi random graphs and random regular graphs observed in [10, Theorem 20] and [3, Theorem 2], respectively. Thus, we again obtain the results for a large class of inhomogeneous random graphs (see Remark 2) as special cases.
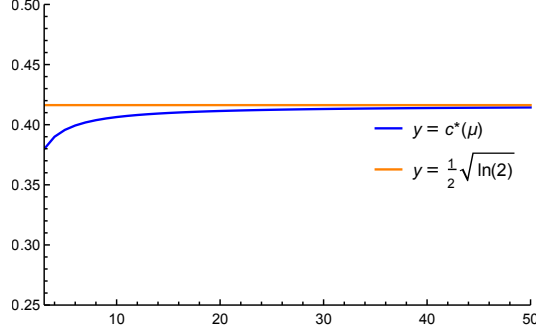
Figure 1: Numerical values of $c^\star(\mu)$ for $3 \leq \mu \leq 50$.

To further illustrate the usefulness of the above phase transition results, we consider graphs obtained by random deletion of edges from a given graph. Such results are crucial for studying the stability of networks to link failures. Percolation refers to keeping the edges of a graph with a given probability $p_n$, independently among each other and independent of the underlying (random) graph. Using Theorems 3.1 and 3.2 we are able to characterize the threshold of the percolation probability for the configuration model, with respect to the $k$-section width and the Max-Cut. Let $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ be the graph obtained by retaining the edges of $\mathrm{CM}_n(\boldsymbol{d})$ with probability $p_n$. An important property of $\mathrm{CM}_n(\boldsymbol{d})$ is that $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ is again distributed as a configuration model conditionally on its degree sequence [15, 24]. Therefore, one can deduce the phase transition results for the extremal cuts of $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ from Theorems 3.1 and 3.2. In fact, since the percolated graphs always have a positive proportion of isolated vertices in the sparse regime (Assumption 1), the minimum bisection below the threshold $\eta = 1/k$ becomes exactly zero with high probability, see Remark 1.

Let $k \geq 2$ be an integer. Then the phase transition for $w_k(\mathrm{CM}_n(\boldsymbol{d}, p_n))$ with $p_n \to p$, occurs at $p = p_{\min}(k, \boldsymbol{d})$, such that the asymptotic proportion of vertices in the largest connected component of $\mathrm{CM}_n(\boldsymbol{d}, p_n)$ is precisely equal to $1/k$. For an arbitrary degree sequence, the explicit solution for $p(k)$ is not immediate from [24, Theorem 3.9]. However, in the particular case of percolation on the $d$-regular graph (i.e. $\boldsymbol{d} = d\boldsymbol{1} = (d, d, \ldots, d)$) with $d \geq 3$, notice that by [24, (3.13),(3.14)], $p_{\min}(k, d)$ can be obtained as a solution for $p$ in the following system of equations:

$$\sqrt{p}d(1 - \sqrt{p} + \sqrt{p}\xi)^{d-1} + (1 - \sqrt{p})d = d\xi, \quad 1 - (1 - \sqrt{p} + \sqrt{p}\xi)^d = \frac{1}{k}, \qquad (3.1)$$

and thus,

$$p_{\min}(k, d) = \frac{1 - \left(1 - \frac{1}{k}\right)^{\frac{1}{d}}}{1 - \left(1 - \frac{1}{k}\right)^{\frac{d-1}{d}}}. \qquad (3.2)$$

It was shown in [24, Theorem 3.9] that, when $p_n \to p$, the phase transition for the largest connected component occurs at $p < 1/\nu$. Thus, the phase transition for $\mathsf{MaxCut}(\mathrm{CM}_n(\boldsymbol{d}, p_n))$ also occurs at $1/\nu$, which for $d$-regular random graphs equals

$$p_{\max}(d) := \frac{1}{d-1}. \qquad (3.3)$$

Therefore, given the phase transition results in Theorems 3.1 and 3.2, we have proved the following theorem:

6

**Theorem 3.3** (Extremal cuts for percolation on random $d$-regular graphs). *For any $d \geq 3$,*

(i) *If $p_n \to p < p_{\min}(k,d)$, then with high probability, $w_k(\mathrm{CM}_n(d\mathbf{1}, p_n)) = 0$. Furthermore, if $p_n \to p > p_{\min}(k,d)$, then with high probability there exists $\zeta > 0$, such that $w_k(\mathrm{CM}_n(d\mathbf{1}, p_n)) > \zeta n$.*

(ii) *If $p_n \to p < p_{\max}(d)$, then*

$$\mathrm{E}_n - \mathsf{MaxCut}(\mathrm{CM}_n(d\mathbf{1}, p_n)) \xrightarrow{\mathrm{d}} Z \sim \mathsf{Poi}\left(\frac{1}{4} \ln\left(\frac{1 + (d-1)p}{1 - (d-1)p}\right)\right),$$

*where $\mathrm{E}_n$ is the number of edges in $\mathrm{CM}_n(d\mathbf{1}, p_n)$. Further, if $p_n \to p > p_{\max}(d)$, then for any $c > c^\star(dp)$,*

$$\mathsf{MaxCut}(\mathrm{CM}_n(d\mathbf{1}, p_n)) \leq n\left(\frac{dp}{4} + c\sqrt{dp}\right),$$

*with high probability, where $c^\star(\cdot)$ is as given by Theorem 3.2.*

## 4 Proof for the $k$-section width

In this section we prove the phase transition of the $k$-section width stated in Theorem 3.1.

### 4.1 Subcritical case

In this subsection we present the proof of Theorem 3.1 (i). In Lemma 4.1 we first state a useful graph theoretic result, which ensures that if (i) the size of the largest component is smaller than $n/k$, (ii) there are $\Theta(n)$ *small* components (i.e., of size at most 2), and (iii) the sizes of every other component is smaller than the $k^{\mathrm{th}}$ fraction of the number of small components, then the $k$-section width is at most $k/2$. This lemma is an extension of [33, Lemma 9] to fit in the scenario when there are possibly no isolated vertices. Then in Lemma 4.2 we show that under Assumption 1, $\Theta(n)$ such *small* components are created. This will complete the proof of the subcritical regime.

**Lemma 4.1.** *Consider a graph $G$ on $n$ vertices, with $m$ components of sizes $c_1 \geq \cdots \geq c_m$, such that (i) $c_1 \leq n/k$, (ii) $\#\{i : c_i \leq 2\} \geq rn$ for some $r > 0$, and (iii) $c_i \leq rn/k$ for all $i > k$. Then, $w_k(G) \leq k/2$. In addition, if $\#\{i : c_i = 1\} \geq k - 1$, then $w_k(G) = 0$.*

*Proof.* Let $G$ contain $m_2$ components of size more than 2, and enumerate them as $C_1, C_2, \ldots, C_{m_2}$ with sizes $c_1 \geq \cdots \geq c_{m_2}$, respectively (ties can be broken arbitrarily). We construct $k$ partitions $V_1, V_2, \ldots, V_k$ sequentially as follows. Define $V_1(1) = C_1$, and $V_i(1) = \emptyset$ for $i = 1, \ldots, k$. For $2 \leq t \leq m_2$,

$$V_i(t) = \begin{cases} V_i(t-1) \cup C_t & \text{if} \quad i = \min\{j : |V_j(t-1) \cup C_t| \leq n/k\} \\ V_i(t-1) & \text{otherwise,} \end{cases}$$

i.e., sequentially at each step add all the vertices in components of size more than 2, to the partitions in a way such that the size of each partition does not exceed $n/k$. The claim below establishes that the above steps are feasible.

**Claim 1.** *For all $2 \leq t \leq m_2$, $\min\{j : |V_j(t-1) \cup C_t| \leq n/k\} \leq k$.*

*Proof.* Note that, due to condition (i) in the lemma, $|V_i(1)| \leq n/k$. Now, if possible assume that at step $t_0 \leq m_2$, $|V_j(t_0 - 1) \cup C_{t_0}| = |V_j(t_0 - 1)| + |C_{t_0}| > n/k$ for all $j = 1, \ldots, k$. Summing over $j$, we have

$$\sum_{j=1}^{k} |V_j(t_0 - 1)| + k|C_{t_0}| > n \implies |C_{t_0}| > \frac{n}{k} - \frac{1}{k}\sum_{j=1}^{k} |V_j(t_0 - 1)| \geq \frac{n}{k} - \frac{n - rn}{k},$$

due to condtion (ii). This in turn implies $|C_{t_0}| > \frac{rn}{k}$, which contradicts condition (iii). ⌟

After step $t = m_2$, we first add the components of size 2 and finally components of size 1 (the isolated vertices), if any. Observe that components of size 1, 2 can be added to the partitions such that each partition is of size $\lfloor n/k \rfloor$ or $\lfloor n/k \rfloor + 1$, there are no cross edges between the partitions, and the number of vertices remaining to be included in any partition is at most $k - 1$. Now, if $\#\{i : c_i = 1\} \geq k - 1$, then at the last step the remaining vertices must be isolated ones, and these do not create any cross edge, and thus the $k$-section width is exactly zero. Otherwise, the remaining $k - 1$ vertices can form at most $k/2$ cross edges (the worst case being there are no isolated vertices). $\square$

We now show that $\mathrm{CM}_n(\boldsymbol{d})$ with $\eta < 1/k$ satisfies all the conditions of Lemma 4.1, with high probability. Condition (i) follows from Theorem 2.1 (i) and the fact that $\eta < 1/k$. In Lemma 4.2 below, we show that the number of components of size 2 converges in probability to a positive constant which verifies Condition (ii). Condition (iii) is a consequence of Theorem 2.1 (ii). The proof of Theorem 3.1 (i) is now complete by Lemma 4.1. $\square$

Recall that $n_1$ denotes the number of vertices in $\mathrm{CM}_n(\boldsymbol{d})$ with degree one, and $n_1/n \to \mathbb{P}(D = 1) = p_1 > 0$. Suppose that the degree one vertices are indexed as $1, 2, \ldots, n_1$. We say that a *pair* is created if a degree one vertex is joined with another degree one vertex. Thus, the pairs are the components of size 2 in $\mathrm{CM}_n(\boldsymbol{d})$.

**Lemma 4.2.** *Let* $P_n := \sum_{1 \leq i < j \leq n_1} \mathbb{1}\{i \leftrightsquigarrow j\}$ *be the number of pairs in* $\mathrm{CM}_n(\boldsymbol{d})$. *Then, as* $n \to \infty$,

$$\frac{P_n}{n} \xrightarrow{\mathbb{P}} \frac{p_1^2}{2\mathbb{E}[D]}.$$

*Proof.* Note that, by Assumption 1

$$\frac{1}{n}\mathbb{E}[P_n] = \frac{1}{n}\sum_{1 \leq i < j \leq n_1} \mathbb{P}(i \leftrightsquigarrow j) = \frac{1}{n}\binom{n_1}{2}\frac{1}{\ell_n - 1} \to \frac{p_1^2}{2\mathbb{E}[D]}. \tag{4.1}$$

Further, if $I = \{(i_1, j_1, i_2, j_2) : 1 \leq i_1 < j_1 \leq n_1, 1 \leq i_2 < j_2 \leq n_1, i_1, i_2, j_1, j_2 \text{ are distinct}\}$, then

$$\begin{aligned}
\frac{1}{n^2}\mathbb{E}[P_n^2] &= \frac{1}{n^2}\left(\sum_{i_1, j_1, i_2, j_2 \in I} \mathbb{P}(i_1 \leftrightsquigarrow j_1, i_2 \leftrightsquigarrow j_2) + \sum_{1 \leq i < j \leq n_1} \mathbb{P}(i \leftrightsquigarrow j)\right) \\
&= \frac{1}{n^2}\left(\frac{1}{(\ell_n - 1)(\ell_n - 3)}\binom{n_1}{2}\binom{n_1 - 2}{2} + \binom{n_1}{2}\frac{1}{\ell_n - 1}\right) \longrightarrow \left(\frac{p_1^2}{2\mathbb{E}[D]}\right)^2.
\end{aligned} \tag{4.2}$$

Therefore,

$$\frac{1}{n^2}\mathrm{Var}(P_n) \to 0, \tag{4.3}$$

and an application of Chebyshev's inequality completes the proof. $\square$

8

## 4.2 Supercritical case

In this subsection we prove the supercritical phase of the $k$-section width stated in Theorem 3.1 (ii). As mentioned earlier, since $\eta > 1/k$, the fraction of vertices in the largest component is more than $1/k$, with high probability. Therefore, in any balanced $k$-partition of the graph $G$, there must exist two distinct partitions each containing an asymptotically positive proportion of vertices from the largest component. It is thus enough to show that if the largest component is partitioned into two sets $V_1$, $V_2$, each containing a positive proportion of vertices, then with high probability, there exist $\Theta(n)$ cross-edges between $V_1$ and $V_2$. The following key definition formalizes this cut-property:

**Definition 1** ($\varepsilon$-$\delta$ cut). Given $\varepsilon, \delta > 0$, an $(\varepsilon, \delta)$-cut of a graph $G = (V, E)$ is a partition of $V$ in two sets $V_1$, and $V_2$ such that $|V_1|, |V_2| > \varepsilon|V|$, and the number of edges between $V_1$ and $V_2$ is at most $\delta|V|$.

Now observe that the following proposition is enough to conclude Theorem 3.1 (ii):

**Proposition 4.3.** *Consider a supercritical* $\mathrm{CM}_n(\boldsymbol{d})$. *For any* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon) > 0$ *such that with high probability the giant component* $\mathscr{C}_{(1)}$ *does not have an* $(\varepsilon, \delta)$-*cut.*

We now briefly sketch the outline of the proof of Proposition 4.3. The idea was first introduced by Bollobás et al. [5] in the context of stability of the largest connected component of inhomogeneous random graphs. We leverage their technique for the configuration model, and in conjunction with suitable structural properties of the giant component, prove Proposition 4.3. The application of the technique to the configuration model poses a substantial challenge, and due to the dependence among edges, the methods for inhomogeneous random graphs [5] or the Erdős-Rényi random graphs [33] are challenging to apply. In this paper, we therefore present some novel arguments that establish the necessary structural properties for the proof techniques to work. In particular, we come with a sequential construction of the configuration model in Subsection 4.2.1, that facilitates the comparison between $\mathrm{CM}_n(\boldsymbol{d})$ and the graph with one deleted vertex.

For any graph $G$ with vertex set $V$, define the *$k$-core* to be the maximal set of vertices $V^k \subseteq V$, such that in the subgraph induced by $V^k$, each vertex has degree at least $k$. Note that the $k$-core of any graph is unique, although it can possibly consist of an empty graph only. It is worthwhile to note that the 2-core of any connected graph is also connected. Algorithmically, the $k$-core of a graph can be obtained by sequentially deleting the vertices of degree less than $k$ along with all their incident edges, until all the vertices in the remaining graph have degree more than $k$. Observe that, $V^k \supseteq V^{k+1}$, and the subgraph induced by $V \setminus V^2$ is a forest. See Figure 2a for an instance of the 2-core of a graph and the trees hanging from it. Figure 2b visualizes the 3-core as a subset of the 2-core.

As explained above, the largest connected component $\mathscr{C}_{(1)}$ of $\mathrm{CM}_n(\boldsymbol{d})$ can be decomposed into two disjoint subsets of vertices: the 2-core $\mathscr{C}_{(1)}^2$, and a forest of vertex-disjoint trees hanging from the 2-core. Informally speaking, the 2-core is the denser part of the graph. Therefore, at a high level, splitting the 2-core into two parts, each containing a positive proportion of vertices, is in general 'costly', and would lead to formation of a 'huge' number of cross edges. Thus the optimal strategy might be to 'peel off' the hanging trees, since moving each hanging tree to some other partition would form precisely one cross-edge. But in that case also, we show that the number of vertices in each of the hanging trees are small (essentially finite), and hence in order to move $\Theta(n)$ vertices to some other partition, $\Theta(n)$ trees must be cut, and thus, $\Theta(n)$ cross edges will be created.

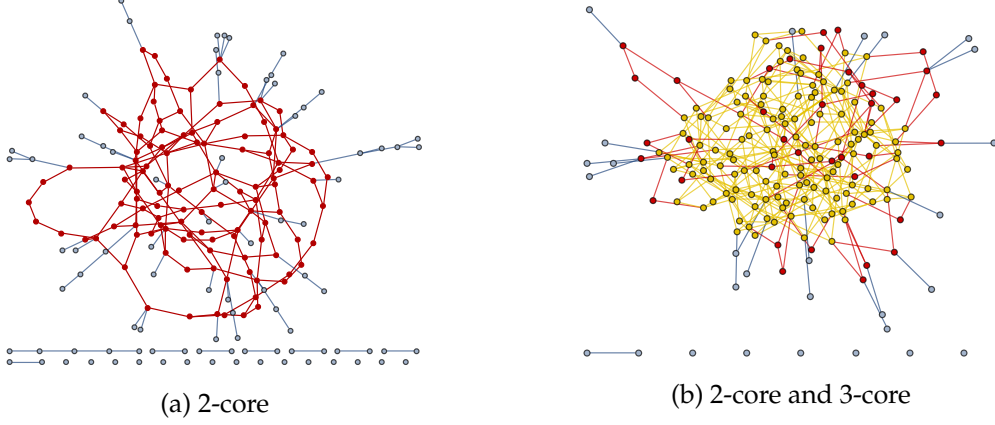|          |                   |
|:--------:|:-----------------:|
| (a) 2-core | (b) 2-core and 3-core |

Figure 2: (a) The highlighted (red) 2-core and the trees hanging from it. (b) The yellow part highlights the 3-core, which is contained in the 2-core (union of red and yellow parts).

To formalize the above heuristics, the proof of Proposition 4.3 breaks into two key steps that we show each holds with high probability:

(i) The hanging trees are not heavy, in the sense that peeling off a small number of them cannot separate out a large number of vertices. This is formalized in Lemma 4.4.

(ii) The 2-core does not have an $(\varepsilon, \delta)$-cut, which is stated in Lemma 4.5.

Denote by $\mathcal{T}_h$, the set of all trees attached with the 2-core of $\mathscr{C}_{(1)}$, i.e., $T \in \mathcal{T}_h$ if and only if the subgraph in $\mathscr{C}_{(1)}$ induced by $T$ is a tree, $T \cap \mathscr{C}_{(1)}^2 = \emptyset$, and there exists only one vertex $v_T \in \mathscr{C}_{(1)}^2$ that shares an edge with some vertex in $T$. With a little abuse of notation we will write $T$ also to denote the set of vertices in $T$. We always assume that each tree $T \in \mathcal{T}_h$ is rooted at the unique point $w_T$ such that $(v_T, w_T)$ is an edge and $v_T \in \mathscr{C}_{(1)}^2$.

**Lemma 4.4** (Hanging trees are not heavy). *For any $\varepsilon > 0$, there exists $\delta = \delta_1(\varepsilon) > 0$, such that with high probability, for any collection $\mathcal{T} \subseteq \mathcal{T}_h$ of $\delta n$ trees contain at most $\varepsilon n$ vertices in total.*

**Lemma 4.5** (No $(\varepsilon, \delta)$-cut in the 2-core). *For any $\varepsilon > 0$, there exists $\delta = \delta_2(\varepsilon) > 0$ such that with high probability, $\mathscr{C}_{(1)}^2$ does not have any $(\varepsilon, \delta)$-cut.*

The above two lemmas are proved at the end of the subsection. Now we prove Proposition 4.3 using Lemmas 4.4 and 4.5.

*Proof of Proposition 4.3.* Fix any $\varepsilon > 0$. Choose $\delta = \min\{\delta_1(\varepsilon/2), \delta_2(\varepsilon/2)\}$, where $\delta_1(\varepsilon)$ and $\delta_2(\varepsilon)$ are as in Lemmas 4.4 and 4.5, respectively.

We now claim that for this choice of $\delta$, there is no $(\varepsilon, \delta)$-cut in $\mathscr{C}_{(1)}$. Indeed, existence of an $(\varepsilon, \delta)$-cut in $\mathscr{C}_{(1)}$ implies that there exists $\delta n$ edges, whose removal splits $\mathscr{C}_{(1)}$ into two parts, both containing at least $\varepsilon n$ vertices. Observe that due to the choice of $\delta$, removal of any set of $\delta n$ edges can separate out at most $\varepsilon n/2$ vertices belonging to $\cup_{T \in \mathcal{T}_h}\{T\}$, and at most $\varepsilon n/2$ vertices belonging to $\mathscr{C}_{(1)}^2$, with high probability, and the proof is complete. $\square$

### 4.2.1 Hanging trees are not heavy

*Proof of Lemma 4.4.* The proof consists of two main steps. The first step establishes that $r$ can be chosen large enough, so that with high probability, the total number of vertices at

depth more than $r$ in all hanging trees combined, is arbitrarily 'small'. This is formalized in Lemma 4.6.

**Lemma 4.6.** *For any $\varepsilon > 0$, there exists $r = r(\varepsilon) > 0$ such that with high probability*

$$\left| \mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}^2_{(1)}, r] \right| < \varepsilon n.$$

In the second step we show the following property of the underlying degree sequence:

**Lemma 4.7.** *Under Assumption 1.b, given any $\varepsilon, r > 0$, there exists $\delta = \delta(\varepsilon, r) > 0$, such that for all $n \geq 1$, the sum of degrees of the $r$-neighborhood of any $\delta n$ vertices is at most $\varepsilon n$, i.e. $\sum_{u \in \mathcal{N}[U,r]} d_i < \varepsilon n$ uniformly over all subsets $U \subseteq [n]$ such that $|U| < \delta n$.*

Given Lemmas 4.6 and 4.7, the proof of Lemma 4.4 can now be completed. Consider the following equivalent re-statement of Lemma 4.4:

> For any $\varepsilon, \beta > 0$, there exists $\delta = \delta(\varepsilon) > 0$ and $n_0 = n_0(\varepsilon, \beta)$, for which, the probability that there exists a subset $\mathcal{T} \subseteq \mathcal{T}_h$ with $|\mathcal{T}| < \delta n$ and $|\bigcup_{T \in \mathcal{T}} \{T\}| \geq \varepsilon n$, is at most $\beta$ for all $n \geq n_0$.

To show the above statement, fix any $\varepsilon, \beta > 0$. Using Lemma 4.6, choose $r = r(\varepsilon/2)$ and $n_1 = n_1(\varepsilon/2, \beta)$, such that for all $n \geq n_1$,

$$\mathbb{P}\left( \left| \mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}^2_{(1)}, r] \right| \geq \frac{\varepsilon n}{2} \right) < \beta.$$

Also, appealing to Lemma 4.7, we choose $\delta = \delta(\varepsilon/2, r)$ and $n_2 = n_2(\varepsilon, r)$, such that for all $n \geq n_2$, $\nexists \, U \subseteq V$ with $|U| < \delta n$ and $\sum_{u \in \mathcal{N}[U,r]} d_i \geq \varepsilon n/2$. Now observe that if there exists a subset $\mathcal{T} \subseteq \mathcal{T}_h$ with $|\mathcal{T}| < \delta n$ and $|\bigcup_{T \in \mathcal{T}} \{T\}| \geq \varepsilon n$, then

$$\text{either} \quad \left| \bigcup_{T \in \mathcal{T}} \mathcal{N}[v_T, r] \right| > \frac{\varepsilon n}{2} \quad \text{or} \quad \left| \bigcup_{T \in \mathcal{T}} \{T\} \setminus \bigcup_{T \in \mathcal{T}} \mathcal{N}[v_T, r] \right| > \frac{\varepsilon n}{2},$$

where $v_T$ is the unique vertex in $T$ that has a neighboring vertex in $\mathscr{C}^2_{(1)}$. Choose $n_0 = \max\{n_1, n_2\}$ so that, for all $n \geq n_0$, the probability of the first event is 0, and that of the latter event is at most $\beta$, which concludes the proof. $\square$

It remains to prove Lemmas 4.6 and 4.7. We start with Lemma 4.7.

*Proof of Lemma 4.7.* Fix any $\varepsilon > 0$. We first verify the case when $r = 1$ and prove this lemma by induction. Due to Assumption 1.b, $K = K(\varepsilon) > 0$ can be chosen such that for all $n \geq 1$,

$$\frac{1}{n} \sum_{i \in [n]} d_i \mathbb{1}\{d_i > K\} < \frac{\varepsilon}{2}. \tag{4.4}$$

Take $\delta = \varepsilon/(2K)$, and fix any $V \subseteq [n]$ with $|V| < \delta n$. Then,

$$\frac{1}{n} \sum_{i \in V} d_i \leq \frac{1}{n} \sum_{i \in V} d_i \mathbb{1}\{d_i \leq K\} + \frac{1}{n} \sum_{i \in [n]} d_i \mathbb{1}\{d_i > K\} < \varepsilon. \tag{4.5}$$

Now suppose that Lemma 4.7 holds for some $r > 0$. Choose $\delta_1 = \delta(\varepsilon, 1)$, and $\delta = \delta(\delta_1, r)$. Notice that for any $U \subseteq [n]$ with $|U| < \delta n$, $|\mathcal{N}[U, r]| < \delta_1 n$, and thus, $\sum_{i \in \mathcal{N}[U, r+1]} d_i = \sum_{i \in \mathcal{N}[\mathcal{N}[U,r],1]} d_i < \varepsilon n$. $\square$

11

To prove Lemma 4.6 we require a detailed understanding of the local neighborhood structure of $\mathrm{CM}_n(\boldsymbol{d})$. For the ease of readeability, we start with a heuristic road-map of the arguments. Observe that for any fixed $r > 0$, and given any random observation $G$ of $\mathrm{CM}_n(\boldsymbol{d})$, $\left|\mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right| = n\mathbb{P}\left(V_n \in \mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}_{(1)}^2, r] \mid G\right)$. Therefore, it is enough to show that for any $\varepsilon > 0$, $r = r(\varepsilon)$ can be chosen large enough, such that

$$\mathbb{P}\left(V_n \in \mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}_{(1)}^2, r] \mid G\right) \xrightarrow{\mathbb{P}} \varepsilon' < \varepsilon, \quad \text{as} \quad n \to \infty.$$

But the primary challenge is that it is difficult to obtain the latter probability. For this reason, we will use the local event approximation technique that has been a key to study sparse random graphs [5, 7, 21, 22, 23, 30]. In particular, our results for the configuration model mirror the ones proved in [5] in the context of inhomogeneous random graphs. Roughly speaking, the key idea is based upon two observations:

(i) The local neighborhood of a typical vertex resembles a branching process, i.e., with high probability, the breadth-first-search (BFS) exploration starting from $V_n$ up to suitable depth can be coupled with a branching process. This is formally stated in Proposition 4.8.

(ii) Looking at the local neighborhood of $V_n$ up to suitable distance, it can be determined whether $V_n$ is in the 2-core. More specifically, the event that $V_n$ is within the $r$ neighborhood of the 2-core, is asymptotically 'equivalent' to the event that for some $L_n \to \infty$, there exists two vertex disjoint paths of length $L_n$ from a vertex within the $r$ neighborhood of $V_n$. This fact is later formalized in Lemma 4.9.

The proof follows once we have these ingredients in place. First we start by introducing some notations. Denote by $\mathcal{X}$ the branching process with initial distribution $D$ and progeny distribution $D^* - 1$, where $D$ is the limiting degree distribution as in Assumption 1, and $D^*$ is the size-biased distribution of $D$, i.e.,

$$\mathbb{P}(D^* = j) = \frac{j\,\mathbb{P}(D = j)}{\mathbb{E}[D]}, \quad j \geq 1.$$

Note that the survival probability of $\mathcal{X}$ is given by $\eta$, as in (2.2) (cf. [29]). The number of offspring of $\mathcal{X}$ in generation $l$ is denoted by $\mathcal{Z}_l$, and the number of vertices at distance $l$ in the breadth-first neighborhood exploration tree (i.e. the BFS tree) starting from vertex $v$ is denoted by $Z_l(v)$. Furthermore, define the following events:

(a) $\mathrm{TC}_r(v)$: the vertex $v$ is within distance $r$ of the 2-core of $\mathscr{C}_{(1)}$,

(b) $\mathrm{LTC}_r(v, L)$: there exists a vertex $v'$ at distance $t$ of $v$, $t \leq r$, with two vertex disjoint paths of length $L$ starting at $v'$ which join $v$ to the vertices at distance $t + L$ from $v$.

(c) $\mathrm{DS}_r$: the branching process $\mathcal{X}$ has a progeny within the first $r$ generations that has two children, both of which survive till infinity.

(d) $\mathrm{LDS}_r(L)$: the branching process $\mathcal{X}$ has a progeny within the first $r$ generations that has two children surviving further $L$ generations.

As explained in the proof sketch above, the following proposition couples the local neighborhood of a random vertex with the branching process $\mathcal{X}$.

**Proposition 4.8** ([21, Proposition 5.4]). *Let* $\{\mathcal{Z}_l^1\}_{l\geq 1}, \{\mathcal{Z}_l^2\}_{l\geq 1}$ *be two independent copies of* $\{\mathcal{Z}_l\}_{l\geq 1}$, *and* $V_n, W_n$ *be two independent typical vertices of* $\mathrm{CM}_n(\boldsymbol{d})$. *There exists* $(L_n)_{n\geq 1}$ *such that* $L_n \to \infty$, *and a coupling* $\left((\hat{Z}_l^1, \hat{Z}_l^2), (\hat{\mathcal{Z}}_l^1, \hat{\mathcal{Z}}_l^2)\right)_{l=1}^{L_n}$ *of* $\left((Z_l(V_n), Z_l(W_n)), (\mathcal{Z}_l^1, \mathcal{Z}_l^2)\right)_{l=1}^{L_n}$ *such that*

$$\lim_{n\to\infty} \mathbb{P}\left(\exists\, l \leq L_n : (\hat{Z}_l^1, \hat{Z}_l^2) \neq (\hat{\mathcal{Z}}_l^1, \hat{\mathcal{Z}}_l^2)\right) = 0. \tag{4.6}$$

The next lemma shows that for any $(L_n)_{n\geq 1}$ that increases to infinity at a rate slower than $\log(n)$, the two events $\mathrm{TC}_r(V_n)$ and $\mathrm{LTC}_r(V_n, L_n)$ are equivalent.

**Lemma 4.9.** *Let* $(L_n)_{n\geq 1}$ *be such that* $L_n \to \infty$ *and* $L_n/\log(n) \to 0$ *as* $n \to \infty$. *Then, for any fixed* $r \geq 1$,

$$\lim_{n\to\infty} \mathbb{P}\left(\mathrm{TC}_r(V_n)\,\Delta\,\mathrm{LTC}_r(V_n, L_n)\right) = 0.$$

We defer the proof of Lemma 4.9 until Section 4.3, and complete the proof of Lemma 4.6 using Lemma 4.9.

*Proof of Lemma 4.6.* Fix any $r > 0$. Observe that for any $L_n$ such that $L_n \to \infty$,

$$\lim_{n\to\infty} \mathbb{P}\left(\mathrm{LDS}_r(L_n)\right) = \mathbb{P}\left(\mathrm{DS}_r\right).$$

Furthermore, choose $L_n^{(1)}$ according to Proposition 4.8, and $L_n^{(2)}$ such that Lemma 4.9 holds. Therefore, for $L_n = \min\{L_n^{(1)}, L_n^{(2)}\}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\mathrm{TC}_r(V_n)\right) = \lim_{n\to\infty} \mathbb{P}\left(\mathrm{LTC}_r(V_n, L_n)\right) = \lim_{n\to\infty} \mathbb{P}\left(\mathrm{LDS}_r(L_n)\right) = \mathbb{P}\left(\mathrm{DS}_r\right). \tag{4.7}$$

Also,

$$\frac{\left|\mathcal{N}[\mathscr{C}_{(1)}^2, r]\right|}{n} = \mathbb{P}\left(\mathrm{TC}_r(V_n)\mid\mathrm{CM}_n(\boldsymbol{d})\right) \implies \frac{1}{n}\mathbb{E}\left[|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|\right] = \mathbb{P}\left(\mathrm{TC}_r(V_n)\right),$$

and hence using (4.7), we get

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|\right] = \lim_{n\to\infty} \mathbb{P}\left(\mathrm{TC}_r(V_n)\right) = \mathbb{P}\left(\mathrm{DS}_r\right). \tag{4.8}$$

To find $\mathrm{Var}\left(|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|\right)$, consider two vertices $V_n, W_n$ chosen uniformly at random independently of the graph and independently of each other. Again, note that

$$\frac{|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|^2}{n^2} = \mathbb{P}\left(V_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r], W_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r]\mid\mathrm{CM}_n(\boldsymbol{d})\right).$$

Thus,

$$\begin{aligned}
\frac{1}{n^2}\mathbb{E}\left[|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|^2\right] &= \mathbb{P}\left(V_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r], W_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right)\\
&= \mathbb{P}\left(\mathrm{TC}_r(V_n) \cap \mathrm{TC}_r(W_n)\right).
\end{aligned} \tag{4.9}$$

Recall from Proposition 4.8 that with high probability, the $L_n$ neighborhoods of $V_n, W_n$ can be coupled with two independent copies of $\mathcal{X}$. Hence,

$$\begin{aligned}
\mathbb{P}\left(\mathrm{TC}_r(V_n) \cap \mathrm{TC}_r(W_n)\right) &= \mathbb{P}\left(\mathrm{LTC}_r(V_n, L_n) \cap \mathrm{LTC}_r(W_n, L_n)\right) + o(1)\\
&= \mathbb{P}\left(\mathrm{LTC}_r(V_n, L_n)\right)\mathbb{P}\left(\mathrm{LTC}_r(W_n, L_n)\right) + o(1)\\
&= \mathbb{P}\left(V_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right)^2 + o(1),
\end{aligned}$$

13

and it follows that

$$\frac{1}{n^2}\mathbb{E}\left[|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|^2\right] = \mathbb{P}\left(V_n \in \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right)^2 + o(1) = \mathbb{P}\left(\mathrm{DS}_r\right)^2 + o(1).$$

Therefore,

$$\frac{1}{n^2}\mathrm{Var}\left(|\mathcal{N}[\mathscr{C}_{(1)}^2, r]|\right) \to 0, \quad \text{as} \quad n \to \infty. \tag{4.10}$$

Using Chebyshev's inequality, (4.8) and (4.10) yields for any fixed $r \geq 1$,

$$\frac{|\mathcal{N}[\mathscr{C}_{(1)}^2, r)]|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}\left(\mathrm{DS}_r\right). \tag{4.11}$$

Further notice that

$$\lim_{r \to \infty} \mathbb{P}\left(\mathrm{DS}_r\right) = \mathbb{P}\left(\mathcal{X}\text{ survives}\right) = \eta. \tag{4.12}$$

Using Theorem 2.1, (4.11) yields

$$\frac{\left|\mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right|}{n} \xrightarrow{\mathbb{P}} \eta - \mathbb{P}\left(\mathrm{DS}_r\right) > 0. \tag{4.13}$$

Now, $\mathbb{P}(\mathrm{DS}_r) \nearrow \eta$ as $r \to \infty$. Thus, for any $\varepsilon > 0$, we can choose $r_0 = r_0(\varepsilon)$ such that $\eta - \mathbb{P}\left(\mathrm{DS}_r\right) < \varepsilon$ for all $r \geq r_0$. Hence, with high probability, $\left|\mathscr{C}_{(1)} \setminus \mathcal{N}[\mathscr{C}_{(1)}^2, r]\right| < \varepsilon n.$ □

### 4.2.2   2-core is well-connected

*Proof of Lemma 4.5.* In this proof we leverage the first moment method argument as used in [5]. Condition on the degree sequence $\tilde{d} = (\tilde{d}_1, \cdots, \tilde{d}_n)$ of $\mathscr{C}_{(1)}^2$. Let $n_2 = |\mathscr{C}_{(1)}^2|$ and let $m_2$ be the number of edges in the 2-core.

Recall that $\mathscr{C}_{(1)}^2$ can be obtained from $\mathscr{C}_{(1)}$ by sequentially deleting the vertices of degree one until all the vertices in the deleted subgraph have degree at least two. Thus, two *paired* half-edges are deleted at each step, and conditional on the deleted half-edges the perfect matching on the rest of the half-edges remains a uniform perfect matching. In particular, $\mathscr{C}_{(1)}^2$ is distributed as a configuration model conditioned on the degree sequence $\tilde{d}$ (cf. [28, Section 3]). Furthermore, we will need the following estimate for the number of degree three vertices in the two core:

**Claim 2.** *Let $N_j$ denote the number vertices in the 2-core having degree $j$. Denote by $\rho_j$ the probability that the root of the branching process $\mathcal{X}$ has exactly $j$ neighbors that survive. Then, as $n \to \infty$, $N_j/n \xrightarrow{\mathbb{P}} \rho_j$.*

*Proof.* The proof follows using similar arguments as in the proof of Lemma 4.6. Note that $\mathbb{E}[N_j]/n = \mathbb{P}(V_n \in \mathscr{C}_{(1)}^2, \text{ and } D_n = j)$, where $V_n$ is a typical vertex, and $D_n$ is the degree of $V_n$. Let $\mathrm{TS}_j(V_n)$ denote the event that $\{V_n \in \mathscr{C}_{(1)}^2, \text{ and } D_n = j\}$ and $\mathrm{LTS}_j(V_n)$ denote the (localized) event that there are $j$ disjoint non self-intersecting paths starting from $V_n$ of length $L_n$, where $L_n \to \infty$ such that Proposition 4.8 holds. The essentially same arguments as in the proof of Lemma 4.9 (see Section 4.3) can be followed to show that, for $L_n \to \infty$ and $L_n = o(\log(n))$ ,

$$\mathbb{P}(\mathrm{TS}_j(V_n) \, \Delta \, \mathrm{LTS}_j(V_n, L_n)) \to 0. \tag{4.14}$$

Moreover, an application of Proposition 4.8 and an argument identical to (4.9) again yields $\mathrm{Var}\left(N_j\right) = o(n^2)$ and the proof follows. ⌟

Having proved the local event approximation in Subsection 4.3, the rest of the proof is similar to [5], and will be sketched briefly for completeness.

For any subset $A \subset \mathscr{C}^2_{(1)}$, we define $\bar{A} = \mathscr{C}^2_{(1)} \setminus A$. Further, recall that for $A \subset \mathscr{C}^2_{(1)}$, we denote the half-edges incident to the vertices in $A$ by $S(A)$. For a set of half-edges $S$, denote by $p(S; \tilde{\boldsymbol{d}})$ the probability that the half-edges of $S$ are paired among each other in $\mathscr{C}^2_{(1)}$, conditional on $\tilde{\boldsymbol{d}}$. Using the fact that the half-edges of $\mathscr{C}^2_{(1)}$ form a uniform perfect matching conditional on the degrees, we obtain

$$p(S; \tilde{\boldsymbol{d}}) = \frac{(|S| - 1)!!(2m_2 - |S| - 1)!!}{(2m_2 - 1)!!} \leq \frac{1}{\binom{m_2}{|S|/2}}. \tag{4.15}$$

A partition $A, \bar{A}$ of $\mathscr{C}^2_{(1)}$ is called $(\varepsilon, \delta)$-*bad* if $|A|, |\bar{A}| \geq \varepsilon n$, and there is a subset $S \subset S(A)$ with $|S(A) \setminus S| \leq \delta n$ such that all the half-edges in $S$ are paired with each other during the random matching of the half-edges. Let $\Gamma_n$ denote the number of bad partitions of $\mathscr{C}^2_{(1)}$. Thus,

$$\mathbb{E}_{\tilde{\boldsymbol{d}}}[\Gamma_n] \leq \sum_{A \subset \mathscr{C}^2_{(1)} : |A|, |\bar{A}| \geq \varepsilon n} \sum_{S \subset S(A) : |S(A) \setminus S| \leq \delta n} p(S; \tilde{\boldsymbol{d}}), \tag{4.16}$$

where $\mathbb{E}_{\tilde{\boldsymbol{d}}}[\cdot]$ denotes the conditional expectation given the degree sequence $\tilde{\boldsymbol{d}}$. We need to show that for all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$, such that $\mathbb{E}_{\tilde{\boldsymbol{d}}}[\Gamma_n] \to 0$, as $n \to \infty$.

We first derive a lower bound on $\binom{m_2}{|S|/2}$. Observe that each vertex in $\mathscr{C}^2_{(1)}$ has degree at least 2, and thus, $|S(A)|/2 \geq |A|$ and $m_2 - |S(A)|/2 \geq |\bar{A}|$, where for the second inequality we have used the fact that $2m_2 - |S(A)| = S(\bar{A})$. Note that for a supercritical $\mathrm{CM}_n(\boldsymbol{d})$, $\mathbb{E}[D(D - 2)] > 0$, and Assumption 1.c thus implies $\mathbb{P}(D \geq 3) > 0$. Therefore, by Claim 2, there exists $\varepsilon_1 > 0$ such that the proportion of degree 3 vertices in $\mathscr{C}^2_{(1)}$ is at least $\varepsilon_1 n$ with high probability. Fix such an $\varepsilon_1 > 0$, and let $\mathcal{A}_n$ denote the event that the proportion of degree 3 vertices in $\mathscr{C}^2_{(1)}$ is at least $\varepsilon_1 n$. Note that on $\mathcal{A}_n$, one of the parts among $A$ and $\bar{A}$ contains at least $\varepsilon_1 n/2$ degree three vertices. Consequently, either $|S(A)|/2 \geq |A| + \varepsilon_1 n/4$ or $m_2 - |S(A)|/2 \geq |\bar{A}| + \varepsilon_1 n/4$ on $\mathcal{A}_n$. Using these bounds, and the fact that $|A|, |\bar{A}| \geq \varepsilon n$, it follows that

$$\binom{m_2}{|S(A)|/2} \geq \exp(4an) \binom{n_2}{|A|},$$

for some $a > 0$ chosen as a function of $\varepsilon, \varepsilon_1$. Moreover, for any partitions $A, \bar{A}$ we have $|S(A) \setminus S| \leq \delta n$, and $\delta$ can be chosen small enough such that

$$\binom{m_2}{|S|/2} \geq \exp(3an) \binom{n_2}{|A|},$$

which gives us the requisite lower bound.

To derive an upper bound on the number of possible choices for $A$ and $S$ in (4.16), we note that given $|A| = a_0$, there are $\binom{n_2}{a_0}$ ways of choosing $A$. Also, given $A$, there are at most $\binom{2m_2}{\delta n}$ choices for $S(A) \setminus S$ such that $|S(A) \setminus S| \leq \delta n$. Plugging these estimates back into (4.16) yields

$$\mathbb{E}_{\tilde{\boldsymbol{d}}}[\Gamma_n] \leq \sum_{\varepsilon n \leq a_0 \leq n_2 - \varepsilon n} \binom{n_2}{a_0} m_2 \binom{2m_2}{\delta n} \binom{n_2}{a_0}^{-1} \exp(-3an) \quad \text{on } \mathcal{A}_n.$$

15

Thus, for a small enough choice of $\delta > 0$, it follows that

$$\mathbb{E}_{\tilde{\boldsymbol{d}}}[\Gamma_n] \to 0 \quad \text{on } \mathcal{A}_n, \quad \text{and} \quad \mathbb{P}(\mathcal{A}_n^c) \to 0.$$

This completes the proof. $\qquad\qquad\square$

## 4.3 Approximation of typical local neighborhoods

We prove Lemma 4.9 in this section. A component $\mathscr{C}_{(i)}$ for $i \geq 2$ (i.e., except the largest component) will be called an *intermediate component* if $|\mathscr{C}_{(i)}| > L_n$ for some $L_n \to \infty$. We need to study some structural properties of the intermediate components that will play a key role in establishing Lemma 4.9. For any $L > 0$, define

$$Q_n(L) := \sum_{i \geq 2} |\mathscr{C}_{(i)}| \mathbb{1}\{\mathscr{C}_{(i)} \geq L\}. \tag{4.17}$$

The next lemma roughly states that the proportion of vertices that belong to some intermediate component is negligible with high probability.

**Lemma 4.10** (Small number of vertices in intermediate components). *For any $\varepsilon > 0$, there exists $K = K(\varepsilon)$, such that*

$$\limsup_{n \to \infty} \mathbb{P}\left(V_n \notin \mathscr{C}_{(1)}, |\mathscr{C}(V_n)| > K\right) < \varepsilon. \tag{4.18}$$

*Consequently, for any $L_n$ such that $L_n \to \infty$, as $n \to \infty$,*

$$\frac{\mathbb{E}\left[Q_n(L_n)\right]}{n} \to 0. \tag{4.19}$$

*Proof.* Fix any $K \geq 1$. Recall from Theorem 2.1 that

$$\mathbb{P}\left(V_n \in \mathscr{C}_{(1)}\right) = \mathbb{P}\left(\mathscr{C}(V_n) = \mathscr{C}_{(1)}\right) \to \mathbb{P}\left(|\mathcal{X}| = \infty\right). \tag{4.20}$$

Now, based on the information about the $K$-neighborhood of $V_n$, it can be exactly determined whether the event $\{|\mathscr{C}(V_n)| > K\}$ has occurred or not. Therefore, using Proposition 4.8, we have

$$\mathbb{P}\left(|\mathscr{C}(V_n)| \leq K\right) = \mathbb{P}\left(|\mathcal{X}| \leq K\right) + o(1). \tag{4.21}$$

Combining (4.20) and (4.21) yields

$$\mathbb{P}\left(V_n \notin \mathscr{C}_{(1)}, |\mathscr{C}(V_n)| > K\right) = \mathbb{P}\left(|\mathcal{X}| \in (K, \infty)\right) + o(1), \tag{4.22}$$

and hence (4.18) follows. To see (4.19), notice that by (4.18),

$$\frac{1}{n}\mathbb{E}\left[Q_n(L_n)\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i \geq 2} |\mathscr{C}_{(i)}| \mathbb{1}\{|\mathscr{C}_{(i)}| > L_n\}\right] = \mathbb{P}\left(V_n \notin \mathscr{C}_{(1)}, |\mathscr{C}(V_n)| > L_n\right) \to 0.$$

$\qquad\qquad\square$

Let us now introduce the following novel construction of the configuration model, that will allow us to relate it to the graph after deletion of one vertex. This will be crucial for the completion of the proof of Lemma 4.9.
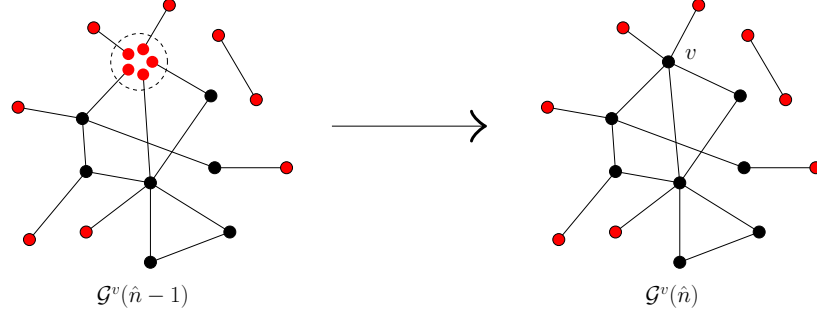
Figure 3: The red vertices are the ones that have not yet been assigned any index. At $\hat{n}^{\text{th}}$ step, five of the unlabeled degree one vertices are selected, and the vertex $v$ is formed.

**Algorithm 1.** Consider a given degree sequence $\boldsymbol{d}$ on vertex set $[n]$.

(S1) Initially there are $\ell_n$ degree one vertices labeled $v(1), \ldots, v(\ell_n)$, each with an attached half-edge. Construct a uniform matching of these $\ell_n$ half-edges. Denote the corresponding graph by $\mathcal{G}(0)$, and set $V(0) = \emptyset$. Also, take any permutation of the index set $\{i \in [n] : d_i > 1\}$ of all vertices of degree more than one, and denote it by $\{\sigma_1, \sigma_2, \ldots, \sigma_{\hat{n}}\}$, where $\hat{n} = n - n_1 - n_0$.

(S2) At step $t + 1$, $0 \leq t \leq \hat{n} - 1$, choose $d_{\sigma_t}$ degree one vertices from the graph $\mathcal{G}(t)$ uniformly at random independently of the perfect matching, and coalesce them into a single vertex with index $\sigma_t$. Let $\mathcal{G}(t+1)$ be the new modified graph, and set $V(t+1) = V(t) \cup \{\sigma_t\}$. See Figure 3 for an illustration of this step.

(S3) After $\hat{n}^{\text{th}}$ step, when all indices $i$ with $d_i > 1$ are exhausted, label all the degree one, and isolated vertices at random, independently of (S1) and (S2).

Note that the vertex index assignment process is independent of the initial perfect matching, and therefore, at any time step $t$, $\mathcal{G}(t)$ is a configuration model given its degree sequence. The algorithm, thus indeed produces a configuration model with degree sequence $\boldsymbol{d}$ in the end. This is formally stated in the next lemma.

**Lemma 4.11.** *For all $t \geq 0$, $\mathcal{G}(t)$ is a configuration model given its degree sequence. In particular, the final graph is distributed as $\mathrm{CM}_n(\boldsymbol{d})$.*

**Remark 4.** In Algorithm 1, the indices corresponding to the vertices with degrees at most one are assigned at the final step (S3). It is worthwhile to note that this is not strictly necessary in order for the algorithm to work. In particular, since the uniform matching is created independent of the index assignments, any assignment ordering produces $\mathrm{CM}_n(\boldsymbol{d})$ in the end. In the proof of Lemma 4.9 below, however, we will require the stated order of indexing the vertices.

Fix any vertex $v$ of degree at least 2, and any permutation $\{\sigma_1, \sigma_2, \ldots, \sigma_{\hat{n}}\}$ of the set $\{i \in [n] : d_i > 1\}$ such that $\sigma_{\hat{n}} = v$. Denote the sequence of graphs constructed in Algorithm 1 by $\{\mathcal{G}^v(t)\}_{t \geq 0}$, i.e., $\mathcal{G}^v(t)$ denotes the graph at the $t^{\text{th}}$ step. The $(\hat{n} - 1)^{\text{th}}$ and $\hat{n}^{\text{th}}$ steps of the algorithm are schematically presented in Figure 3. Now, we complete the proof of Lemma 4.9. We will use the following fact:

**Lemma 4.12.** *For any degree sequence satisfying Assumptions 1.a, and 1.b, the maximum degree $d_{\max}$ is $o(\sqrt{n})$.*

*Proof.* For each fixed $K \geq 1$, $\mathbb{E}[D_n^2 \mathbb{1}\{D_n \leq K\}] \to \mathbb{E}[D^2 \mathbb{1}\{D \leq K\}]$, and consequently, $\mathbb{E}[D_n^2 \mathbb{1}\{D_n > K\}] \to \mathbb{E}[D^2 \mathbb{1}\{D > K\}]$. Thus,

$$\lim_{K \to \infty} \lim_{n \to \infty} \mathbb{E}[D_n^2 \mathbb{1}\{D_n > K\}] = \lim_{K \to \infty} \mathbb{E}[D^2 \mathbb{1}\{D > K\}] = 0.$$

Moreover, observe that

$$\limsup_{n \to \infty} \frac{d_{\max}^2}{n} \leq \limsup_{n \to \infty} \left[ \frac{1}{n} \sum_{i: d_i > K} d_i^2 + \frac{K^2}{n} \right] = \limsup_{n \to \infty} \mathbb{E}\left[ D_n^2 \mathbb{1}\{D_n > K\} \right].$$

Since the left side of the above inequality does not depend on $K$, it follows that

$$\limsup_{n \to \infty} \frac{d_{\max}^2}{n} \leq \limsup_{K \to \infty} \limsup_{n \to \infty} \mathbb{E}\left[ D_n^2 \mathbb{1}\{D_n > K\} \right] = 0.$$

$\square$

*Proof of Lemma 4.9.* Fix $r > 0$, and $L_n$ such that $L_n \to \infty$ and $L_n / \log(n) \to 0$ as $n \to \infty$. The proof is split into two steps: we show that (i) $\mathbb{P}\left( \mathrm{TC}_r(V_n) \setminus \mathrm{LTC}_r(V_n, L_n) \right) \to 0$, and (ii) $\mathbb{P}\left( \mathrm{LTC}_r(V_n, L_n) \setminus \mathrm{TC}_r(V_n) \right) \to 0$.

(i) Define the event $\mathrm{C}(v, r, L)$ that the vertex $v$ is within $r$ distance from a cycle of length at most $L$. Then note that

$$\mathrm{TC}_r(v) \setminus \mathrm{LTC}_r(v, L) \subseteq \mathrm{C}(v, r + L, 2L).$$

Indeed, suppose that $\mathrm{TC}_r(v) \setminus \mathrm{LTC}_r(v, L)$ holds. Let $v_1 \in \mathscr{C}_{(1)}^2$ be such that there exists a path $\mathcal{P}_1$ of length at most $r$ from $v$ to $v_1$ (take $v_1 = v$ if $v \in \mathscr{C}_{(1)}^2$). Now, since $v_1$ is in the two-core, there exists at least two vertex-disjoint paths (disjoint from $\mathcal{P}_1$) starting from $v_1$, and because $\mathrm{LTC}_r(v, L)$ does not happen, any two such paths must either meet each other, or one of them intersects itself within distance $L$ from $v'$. In either cases a cycle of length at most $2L$ is created that is joined to $v$ via a path of length at most $r + L$, and therefore $\mathrm{C}(v, r + L, 2L)$ must hold.

**Claim 3.** *Let $L_n = o(\log(n))$. As $n \to \infty$, $\mathbb{P}\left( \mathrm{C}(V_n, L_n, L_n) \right) \to 0$.*

*Proof.* In the proof we will make use of the path counting techniques introduced in [26]. Define $\ell_n' := \ell_n - 4L_n + 1$. Note that due to Assumption 1.b, a constant $\kappa > 1$ can be chosen such that

$$\frac{1}{\ell_n'} \sum_{i \in [n]} d_i(d_i - 1) \leq \kappa \quad \text{for all} \quad n \geq 1. \tag{4.23}$$

The event $\mathrm{C}(V_n, L_n, L_n)$ implies that there is a path $(V_n, x_1, x_2, \ldots, x_l)$ of length $l \leq L_n$, and $x_l$ belongs to a cycle $(x_l, x_{l+1}, \ldots, x_{l+m-1})$ of length $m \leq L_n$, where the $x_i$'s are distinct. Fix some $V_n = v$. Then the number of structures with a path $(v, x_1, x_2, \ldots, x_l)$ and a cycle $(x_l, x_{l+1}, \ldots, x_{l+m-1})$ is given by

$$\left[ d_v \left( \prod_{i=1}^{l-1} d_{x_i}(d_{x_i} - 1) \right) d_{x_l} \right] \times \left[ (d_{x_l} - 1) \left( \prod_{i=l+1}^{l+m-1} d_{x_i}(d_{x_i} - 1) \right) (d_{x_l} - 2) \right],$$

18

where the first term in the product is due to the number of ways the path can be formed, and the second is due to the cycle. Furthermore, each of these specific configurations has probability $[(\ell_n - 1)(\ell_n - 3) \dots (\ell_n - 2l - 2m + 1)]^{-1}$. Therefore,

$$
\begin{aligned}
\mathbb{P}\left(\mathrm{C}(V_n, L_n, L_n) \mid V_n = v\right) & \\
&\leq \sum_{l,m \leq L_n} \sum_{x_1,\dots,x_{l+m-1}} \frac{(d_{x_1} - 2)d_v \prod_{i=1}^{l+m-1} d_{x_i}(d_{x_i} - 1)}{(\ell_n - 1)(\ell_n - 3) \dots (\ell_n - 2l - 2m + 1)} \\
&\leq \sum_{l,m \leq L_n} \frac{1}{(\ell_n')^{m+l}} \left(\sum_{i \in [n]} d_i(d_i - 1)\right)^{l+m-2} d_v \sum_{i \in [n]} d_i(d_i - 1)(d_i - 2) \quad (4.24) \\
&\leq \sum_{l,m \leq L_n} \left(\frac{1}{\ell_n'} \sum_{i \in [n]} d_i(d_i - 1)\right)^{l+m-2} d_v \frac{d_{\max}}{\ell_n'} \frac{1}{\ell_n'} \sum_{i \in [n]} d_i(d_i - 1) \\
&\leq \sum_{l,m \leq L_n} \left(\frac{1}{\ell_n'} \sum_{i \in [n]} d_i(d_i - 1)\right)^{l+m-1} d_v \frac{d_{\max}}{\ell_n'} \quad \leq K \frac{d_v \kappa^{2L_n}}{\sqrt{n}}
\end{aligned}
$$

for some constant $K > 0$ where in the final step we have used (4.23) and Lemma 4.12. Therefore,

$$
\mathbb{P}\left(\mathrm{C}(V_n, L_n, L_n)\right) \leq \frac{K \kappa^{2L_n}}{\sqrt{n}} \mathbb{E}\left[D_n\right] = K \mathbb{E}\left[D_n\right] \exp\left(2L_n \log(\kappa) - \frac{1}{2} \log(n)\right) \to 0, \quad (4.25)
$$

by Assumption 1.b, and the fact that $L_n = o(\log(n))$. ⌟

Therefore, for any fixed $r \geq 1$, $\mathbb{P}\left(\mathrm{TC}_r(V_n) \setminus \mathrm{LTC}_r(V_n, L_n)\right) \leq \mathbb{P}\left(\mathrm{C}(V_n, r + L_n, 2L_n)\right) \leq \mathbb{P}\left(\mathrm{C}(V_n, 2L_n, 2L_n)\right) \to 0$, and the proof of part (i) is complete.

(ii) We prove this part for the case when $r = 0$. The proof of the general case is included at the end. Fix any vertex $v \in [n]$, and condition on $V_n = v$. If $d_v \leq 1$ or $v \notin \mathscr{C}_{(1)}$, then $\mathbb{P}\left(\mathrm{LTC}_0(v, L_n)\right) = \mathbb{P}\left(\mathrm{TC}_0(v)\right) = 0$. So, without loss of generality assume that $d_v > 1$ and $v \in \mathscr{C}_{(1)}$. Recall the construction in Algorithm 1 and the definition of the graph $\mathcal{G}^v(t)$. Note that, if $\mathrm{LTC}_0(v, L_n) \setminus \mathrm{TC}_0(v)$ happens, then there are two vertex-disjoint paths in $\mathrm{CM}_n(\boldsymbol{d})$ starting from $v$, which have length at least $L_n$, but they do not meet each other. Furthermore, the event $\mathrm{LTC}_0(v, L_n) \setminus \mathrm{TC}_0(v)$ is determined by the graph $\mathcal{G}^v(\hat{n})$. Define the event $E(v)$ that, while creating the vertex with index $v$ at time $\hat{n}$, one of the degree one vertices in one of the intermediate components of $\mathcal{G}^v(\hat{n} - 1)$ was chosen. Observe that

$$
\mathrm{LTC}_0(v, L_n) \setminus \mathrm{TC}_0(v) \subseteq E(v).
$$

Let $Q_n^v(L_n)$ denote the total number of vertices in the intermediate components of size more than $L_n$, in the graph $\mathcal{G}^v(\hat{n} - 1)$. Using Lemmas 4.11 and 4.12, it follows that $\mathcal{G}^v(\hat{n} - 1)$ is a configuration model given its degree sequence that satisfy Assumption 1.

**Claim 4.** $\frac{1}{n} \max_{v \in [n]} \mathbb{E}\left[Q_n^v(L_n)\right] \to 0$, as $n \to \infty$.

*Proof.* Note that an application of Lemma 4.10 directly implies that $\mathbb{E}\left[Q_n^v(L_n)\right] \to 0$, for any fixed $v$. Let $(d_i^v)$ denote the degree sequence of $\mathcal{G}^v(\hat{n} - 1)$ and let $\nu_n^v := \sum_i d_i^v(d_i^v - 1) / \sum_i d_i^v$. Observe that (i) $\sum_i d_i^v = \ell_n$, (ii) $\sum_i d_i^v(d_i^v - 1) = \sum_{i \in [n]} d_i(d_i - 1) + O(d_{\max}^2)$. Therefore, we get (iii) $\max_{v \in [n]} |\nu_n^v - \nu_n| \to 0$ as $n \to \infty$. Now, while approximating the breadth-first exploration of $\mathcal{G}^v(\hat{n} - 1)$ by a suitable branching process in (4.6), one can in fact obtain error

estimates that are uniform over $v$. This is a consequence of the precise bounds stated in [21, Lemma 5.6], that are used as the main ingredient for the proof of [21, Proposition 5.4]. Therefore, while proving (4.21) for the graph $\mathcal{G}^v(\hat{n}-1)$, one can use (i) and (iii) above to get error estimates that are uniform in $v$. Thus, the claim follows. ⌟

Finally, we bound the probability of the event $E(V_n)$. Note that, in $\mathcal{G}^v(\hat{n}-1)$ there are $n_1 + d_v - 1$ degree one vertices. Therefore, conditional on $\mathcal{G}^v(\hat{n}-1)$, the vertex $v$ is created at step $\hat{n}$ by choosing $d_v$ vertices from a set of $n_1 + d_v - 1$ vertices, and $E(v)$ occurs if at least one of those degree one vertices is from an intermediate component (for which there are at most $Q_n^v(L_n)$ choices). Thus,

$$\mathbb{P}\left(E(v)\right) \leq \frac{d_v}{n_1 + d_v - 1} \mathbb{E}\left[Q_n^v(L_n)\right]$$

Again, by Assumption 1, there exists a constant $K > 0$, such that $n_1 + d_v - 1 \geq \ell_n/K$ for all large $n$. Hence,

$$\begin{aligned}
\mathbb{P}\left(E(V_n)\right) = \frac{1}{n}\sum_{v\in[n]}\mathbb{P}\left(E(v)\right) &\leq \sum_{v\in[n]}\frac{d_v}{n_1 + d_v - 1}\frac{1}{n}\mathbb{E}\left[Q_n^v(L_n)\right] \\
&\leq \frac{K}{n}\left(\max_{v\in[n]}\mathbb{E}\left[Q_n^v(L_n)\right]\right)\sum_{v\in[n]}\frac{d_v}{\ell_n} \leq \frac{K}{n}\max_{v\in[n]}\mathbb{E}\left[Q_n^v(L_n)\right] \to 0,
\end{aligned}$$

(4.26)

where the last step follows from Claim 4. Thus it follows that

$$\mathbb{P}\left(\mathrm{LTC}_0(V_n, L_n) \setminus \mathrm{TC}_0(V_n)\right) = o(1).$$

(4.27)

To see the general case for $d \geq 1$, note that (4.27) implies $\mathbb{E}[\#\{v \in [n] : \mathrm{LTC}_0(v, L) \setminus \mathrm{TC}_0(v) \text{ occurs}\}]/n \to 0$. Using Lemma 4.7, it now follows that the fraction of vertices which are within the $d$ neighborhood of a vertex $v'$ for which $\mathrm{LTC}_0(v', L) \setminus \mathrm{TC}_0(v')$ occurs converges to zero in $L^1$. Therefore, $\mathbb{P}(\mathrm{LTC}_d(V_n, L) \setminus \mathrm{TC}_d(V_n)) = o(1)$, and the proof is complete. □

# 5   Proof for the Max-Cut

We prove Theorem 3.2 in this section. The proof for the sub-/supercritical cases are provided in Sections 5.1 and 5.2, respectively.

## 5.1   Subcritical case

The idea in the subcritical regime is to count the number of cycles. This idea has also been adopted in the proof of [10, Theorem 19] for Erdős-Rényi random graphs. Observe that the bipartite components (components with no cycles or only cycles of even length) contribute all of their edges to the Max-Cut. To analyze the non-bipartite components we first observe in Lemma 5.1 that all the components of a subcritical $\mathrm{CM}_n(\boldsymbol{d})$ are unicyclic (contains only one cycle) with high probability.

**Lemma 5.1** ([20, Theorem 1.2 (b)]). *For subcritical $\mathrm{CM}_n(\boldsymbol{d})$ satisfying* Assumption 1, *the probability that there exists a component with more than one cycle tends to zero as $n \to \infty$.*

Observe that the Max-Cut leaves precisely one edge uncut in each of these unicyclic, non-bipartite components. Therefore, the number of uncut edges in the Max-Cut is with high probability equal to the number of cycles of odd length that the graph contains. Now, the asymptotic number of cycles of length $k$ in $\mathrm{CM}_n(\boldsymbol{d})$, for any fixed $k \geq 1$, is derived in [6, Theorem 2.18], and is stated in the following lemma. Let $C_k^n$ denote the number of cycles of length $k$ in $\mathrm{CM}_n(\boldsymbol{d})$ (a cycle of length one denotes a loop and of length two denotes a multiple edge).

**Lemma 5.2** ([6, Theorem 2.18]). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying* Assumption 1. *Then, for any* $K \geq 1$, *as* $n \to \infty$,

$$(C_k^n)_{k \in [K]} \xrightarrow{d} (X_k)_{k \in [K]}, \tag{5.1}$$

*where* $X_k \sim \mathrm{Poisson}(\nu^k/2k)$, *independently for* $k \in [K]$.

The next lemma proves that with high probability, there are no cycles of growing length. This will be used to show that asymptotically, the total number of odd-length cycles is equal to the sum of the number of all cycles of finite and odd length.

**Lemma 5.3.** *Consider a subcritical* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying* Assumption 1. *Then,*

$$\lim_{K \to \infty} \lim_{n \to \infty} \mathbb{P}\left(\exists\, k > K : C_k^n \geq 1\right) = 0. \tag{5.2}$$

Lemma 5.3 is proved at the end of this subsection. Now, we prove result for the subcritical Max-Cut by using Lemmas 5.1, 5.2 and 5.3

*Proof of Theorem 3.2 (i).* As mentioned earlier, the Max-Cut leaves precisely one edge uncut in each of the unicyclic, non-bipartite components, and by Lemma 5.1, with high probability, the total number of uncut edges precisely equals to the total number of odd-length cycles. Therefore, recalling that the total number of edges equals $\ell_n/2$, it follows that $\ell_n/2 - \mathsf{MaxCut}(\mathrm{CM}_n(\boldsymbol{d})) = \sum_{k \geq 3, k \text{ is odd}} C_k^n$, with high probability. Hence, Lemmas 5.2 and 5.3 yield, as $n \to \infty$,

$$\frac{\ell_n}{2} - \mathsf{MaxCut}(\mathrm{CM}_n(\boldsymbol{d})) = \sum_{k \geq 3, k \text{ is odd}} C_k^n \xrightarrow{d} X, \quad X \sim \mathrm{Poisson}\left(\sum_{\substack{k \geq 3, \\ k \text{ is odd}}} \frac{\nu^k}{2k}\right).$$

$\square$

*Proof of Lemma 5.3.* For brevity of notation, denote by $M$ the total number of edges, i.e., $M = \ell_n/2$. We find the expected value of $C_k^n$ using again the path-counting techniques. To this end, we first fix $k$ distinct vertices $x_1, \ldots, x_k$ which participate in the cycle in the given order. We denote by $\mathcal{I}_k = \{(x_1, \ldots, x_k) : x_i \neq x_j, \forall i \neq j\}$. For each vertex $x_i$, the two half-edges which participate in the cycle may be chosen in $d_{x_i}(d_{x_i} - 1)$ ways. The number of ways to pair these half-edges is thus $\prod_i d_{x_i}(d_{x_i} - 1)$. For any fixed $R \geq 1$, $2R$ half-edges can be paired among each other in $\binom{2R}{R} R!/2^R$. Therefore,

$$\mathbb{E}[C_k^n] = \sum_{\mathcal{I}_k} \prod_{i=1}^k d_{x_i}(d_{x_i} - 1) \frac{\binom{2M-2k}{M-k} \frac{(M-k)!}{2^{M-k}}}{\binom{2M}{M} \frac{M!}{2^M}}$$

$$\leq \left(\frac{1}{(n)_k} \sum_{\mathcal{I}_k} \prod_{i=1}^k d_{x_i}(d_{x_i} - 1)\right) \frac{2^k (n)_k (M)_k}{(2M)_{2k}} \tag{5.3}$$

$$\leq \left(\frac{1}{n} \sum_{i \in [n]} d_i(d_i - 1)\right)^k \frac{2^k (M)_k}{(n)_k (2M)_{2k}},$$

21

where the last step follows from [18, Theorem 52] (see also the proof of [25, Lemma 5.1]). Now, using Stirling's formula we have,

$$(n)_k = \exp[k\ln(n) - k^2/2n - O(k/n + k^3/n^2)].$$

In analogy with (2.1), we define $\nu_n := \mathbb{E}[D_n(D_n - 1)]/\mathbb{E}[D_n]$. Therefore,

$$\mathbb{E}[\# \text{ cycles in } \mathrm{CM}_n(\boldsymbol{d}) \text{ of lengths in } (K, \sqrt{n})] = \sum_{k=K+1}^{\sqrt{n}} \mathbb{E}[C_k^n]$$

$$\leq \kappa_1 \sum_{k=K+1}^{\sqrt{n}} \nu_n^k \left(\frac{1}{n}\sum_{i\in[n]} d_i\right)^k 2^k \exp\left(k\ln(n) - \frac{k^2}{2n} + k\ln(M) - \frac{k^2}{2M} - 2k\ln(2M) + \frac{k^2}{M}\right)$$

$$\leq \kappa_1 \sum_{k=K+1}^{\sqrt{n}} \nu_n^k \exp\left\{-\frac{k^2}{2}\left(\frac{1}{n} - \frac{1}{M}\right)\right\},$$

where the constant $\kappa_1 > 0$ can be chosen to be independent of $K$. Now, for subcritical $\mathrm{CM}_n(\boldsymbol{d})$, we have $M < n$. To see this, note that by the Cauchy-Schwarz inequality,

$$M = \frac{1}{2}\sum_{i\in[n]} d_i \leq \frac{\sqrt{n}}{2}\sqrt{\sum_{i\in[n]} d_i^2} = \frac{\sqrt{n}}{2}\sqrt{2M(1+\nu_n)}. \tag{5.4}$$

Taking the square on both sides and using the fact that since $\nu_n < 1$, we get

$$M \leq \frac{n}{2}(1+\nu_n) < n.$$

Therefore, $(1/M - 1/n) > 0$, and hence,

$$\max_{k\leq\sqrt{n}} \exp\left\{-\frac{k^2}{2}\left(\frac{1}{n} - \frac{1}{M}\right)\right\} \leq \exp\left\{-\frac{n}{2}\left(\frac{1}{n} - \frac{1}{M}\right)\right\} \leq \kappa_2,$$

where the constant $\kappa_2 > 0$ is independent of $K$. Thus,

$$\mathbb{E}(\# \text{ cycles in } \mathrm{CM}_n(\boldsymbol{d}) \text{ of length in } (K, \sqrt{n})) = \kappa_1\kappa_2 \sum_{k=K+1}^{\infty} \nu_n^k \to 0, \tag{5.5}$$

if we first take $n \to \infty$ and then $K \to \infty$. To count the number of cycles of length $> \sqrt{n}$, note that

$$\mathbb{P}\left(\exists \text{ a cycle of length more than } \sqrt{n}\right) \leq \mathbb{P}\left(\exists i \geq 1 : |\mathscr{C}_i| > \sqrt{n}\right) = \mathbb{P}\left(|\mathscr{C}_{\max}| > \sqrt{n}\right). \tag{5.6}$$

Now, an application of [23, Theorem 1.3] yields that $|\mathscr{C}_{\max}| = O_{\mathbb{P}}(d_{\max}) = o(n^{1/2})$ and therefore the probability in (5.6) tends to 0 as $n \to \infty$. The proof of Lemma 5.3 is now complete by combining (5.5), and (5.6). $\qquad\square$

22

## 5.2 Supercritical case

The proof for the supercritical case (Theorem 3.2 (ii)) uses the first moment method. For a set of vertices $A$, recall that $S(A)$ is the total number of half-edges associated with $A$, and $E(A, A^c)$ is the number of edges between $A$ and $A^c$. Partition the graph $CM_n(\boldsymbol{d})$ in two parts $A$ and $A^c$, where we assume without loss of generality that $S = S(A) \leq \ell_n/2$. In this case,

$$
\begin{aligned}
\mathbb{P}\left(E(A, A^c) = K\right) &\leq \frac{\binom{S}{K} K! (S - K - 1)!! \binom{\ell_n - S}{K} (\ell_n - S - K - 1)!!}{(\ell_n - 1)!!} \\
&= \exp\left(\ell_n f(x_n, y_n)(1 + o(1))\right), \quad \text{where} \\
f(x, y) &= \ln\left(x^x (1 - x)^{1-x} (1 - x - y)^{-(1-x-y)/2} (x - y)^{-(x-y)/2} y^{-y}\right),
\end{aligned}
\tag{5.7}
$$

and $y_n = K/\ell_n$, $x_n = S/\ell_n$. We note that

$$
\frac{\partial f}{\partial x}(x, y) = \ln\left(\frac{x}{1 - x}\right) - \frac{1}{2} \ln\left(\frac{x - y}{1 - x - y}\right).
$$

The fact that $x/(1 - x) > (x - y)/(1 - x - y)$ for any $0 < y < x \leq 1/2$, implies that for any fixed $y$, $\{f(x, y) : x \leq 1/2\}$ is maximized at $x = 1/2$. Now, for any $1 \leq K \leq \ell_n/2$,

$$
\left| \left\{ U \subset V : \sum_{i \in U} d_i = K \right\} \right| \leq 2^n.
\tag{5.8}
$$

Thus, for any constant $c > 0$, a union bound and (5.7) yields

$$
\begin{aligned}
\mathbb{P}\left(\mathsf{MaxCut}(CM_n(\boldsymbol{d})) \geq \frac{\ell_n}{4} + nc\sqrt{\mu}\right) &= \mathbb{P}\left(\exists\, A \subseteq [n], S(A) \leq \frac{\ell_n}{2}, E(A, A^c) \geq \frac{\ell_n}{4} + nc\sqrt{\mu}\right) \\
&\leq \sum_{S=1}^{\ell_n/2} \sum_{A : S(A) = S} \sum_{K = \frac{\ell_n}{4} + nc\sqrt{\mu}}^{S} \mathbb{P}\left(E(A, A^c) = K\right) \\
&\leq (1 + o(1)) 2^n \left(\frac{\ell_n}{2}\right)^2 \max_{\frac{\ell_n}{4} + nc\sqrt{\mu} \leq K \leq \frac{\ell_n}{2}} \exp\left[\ell_n f\left(\frac{1}{2}, \frac{K}{\ell_n}\right)\right].
\end{aligned}
$$

Writing $K/\ell_n = y$, notice that

$$
\begin{aligned}
2^n &\exp[\ell_n f(1/2, y)] \\
&= \exp\left[n \ln(2) + \ell_n \ln\left(\frac{1}{2} y^{-y} \left(\frac{1}{2} - y\right)^{-\frac{1}{2}\left(\frac{1}{2} - y\right)} \left(\frac{1}{2} - y\right)^{-\frac{1}{2}\left(\frac{1}{2} - y\right)}\right)\right] \\
&= \exp\left[(1 + o(1)) \ell_n \ln\left(2^{1/\mu} \left(\frac{1}{2}\left(\frac{1}{2} - y\right)^{-\frac{1}{2} + y} y^{-y}\right)\right)\right],
\end{aligned}
\tag{5.9}
$$

since $\ell_n/n \to \mu$. Therefore, we obtain

$$
\begin{aligned}
\mathbb{P}\left(\mathsf{MaxCut}(CM_n(\boldsymbol{d})) \geq \frac{\ell_n}{4} + nc\sqrt{\mu}\right) \\
\leq \kappa n^2 \max_{\frac{1}{4} + \frac{c}{\sqrt{\mu}} \leq y \leq \frac{1}{2}} \exp\left[\ell_n \ln\left(2^{\frac{1}{\mu} - 1} \left(\frac{1}{2} - y\right)^{-\frac{1}{2} + y} y^{-y}\right)\right],
\end{aligned}
$$

for some constant $\kappa > 0$. Now observe that $\left(\frac{1}{2}-y\right)^{-\frac{1}{2}+y}y^{-y}$ is non-increasing in the interval $(1/4, 1/2]$, and therefore the above maximum is attained at $y = \frac{1}{4} + \frac{c}{\sqrt{\mu}}$. Define

$$c^{\star}(\mu) = \inf_{c>0}\left\{c : \left(\frac{1}{4} - \frac{c}{\sqrt{\mu}}\right)^{-\left(\frac{1}{4} - \frac{c}{\sqrt{\mu}}\right)}\left(\frac{1}{4} + \frac{c}{\sqrt{\mu}}\right)^{-\left(\frac{1}{4} + \frac{c}{\sqrt{\mu}}\right)} < 2^{-\frac{1}{\mu}+1}\right\}.$$

Thus, we can conclude that for any $c > c^{\star}(\mu)$,

$$\mathbb{P}\left(\frac{1}{n}\mathsf{MaxCut}(\mathrm{CM}_n(\boldsymbol{d})) > \frac{\mu}{4} + c\sqrt{\mu}\right) \to 0, \quad \text{as} \quad n \to \infty.$$

# Acknowledgment

# References

[1] Albert, R. and Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97.

[2] Barabaśi, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.

[3] Bertoni, A., Campadelli, P., and Posenato, R. (1997). An upper bound for the maximum cut mean value. In *Graph-Theoretic Concepts in Computer Science: 23rd International Workshop*, pages 78–84, Berlin, Heidelberg. Springer.

[4] Bollobás, B. (2001). *Random Graphs:*. Cambridge University Press, 2 edition.

[5] Bollobás, B., Janson, S., and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, 31(1):3–122.

[6] Bordenave, C. (2012). Lecture notes on random graphs and probabilistic combinatorial optimization.

[7] Britton, T., Janson, S., and Martin-Löf, A. (2007). Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability*, 39(4):922–948.

[8] Chang, K. C. and Du, D. H. C. (1987). Efficient algorithms for layer assignment problem. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 6(1):67–78.

[9] Chen, R.-W., Kajitani, Y., and Chan, S.-P. (1983). A graph-theoretic via minimization algorithm for two-layer printed circuit boards. *IEEE Transactions on Circuits and Systems*, 30(5):284–299.

[10] Coppersmith, D., Gamarnik, D., Hajiaghayi, M., and Sorkin, G. B. (2004). Random MAX SAT, random MAX CUT, and their phase transitions. *Random Structures and Algorithms*, 24(4):502–545.

[11] Daudé, H., Martínez, C., Rasendrahasina, V., and Ravelomanana, V. (2012). The MAX-CUT of sparse random graphs. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 265–271. Society for Industrial and Applied Mathematics.

[12] Díaz, J., Petit, J., and Serna, M. (2002). A survey of graph layout problems. *ACM Computing Surveys (CSUR)*, 34(3):313–356.

[13] Faloutsos, M., Faloutsos, P., and Faloutsos, C. (1999). On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 29(4):251–262.

[14] Feige, U. and Krauthgamer, R. (2002). A polylogarithmic approximation of the minimum bisection. *SIAM Journal on Computing*, 31(4):1090–1118.

[15] Fountoulakis, N. (2007). Percolation on sparse random graphs with given degree sequence. *Internet Mathematics*, 4(1):329–356.

[16] Halperin, E. and Zwick, U. (2001). A unified framework for obtaining improved approximation algorithms for maximum graph bisection problems. In *Proceedings of the 8th International IPCO Conference on Integer Programming and Combinatorial Optimization*, pages 210–225, London, UK. Springer-Verlag.

[17] Hansen, P. and Mladenović, N. (2002). *Developments of Variable Neighborhood Search*, pages 415–439. Springer US.

[18] Hardy, G. H., Littlewood, J. E., and Pólya, G. (1952). *Inequalities*. Cambridge University Press.

[19] Håstad, J. (2001). Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859.

[20] Hatami, H. and Molloy, M. (2012). The scaling window for a random graph with a given degree sequence. *Random Structures and Algorithms*, 41(1):99–123.

[21] Hofstad, R. v. d. (2014). *Random Graphs and Complex Networks.*, volume II.

[22] Hofstad, R. v. d. (2016). *Random Graphs and Complex Networks.*, volume I. *To appear with* Cambridge University Press.

[23] Janson, S. (2008). The largest component in a subcritical random graph with a power law degree distribution. *The Annals of Applied Probability*, 18(4):1651–1668.

[24] Janson, S. (2009a). On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:87–118.

[25] Janson, S. (2009b). The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225.

[26] Janson, S. (2010). Susceptibility of random graphs with given vertex degrees. *Journal of Combinatorics*, 1(3-4):357–387.

[27] Janson, S., Knuth, D. E., Łuczak, T., and Pittel, B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4(3):233–358.

[28] Janson, S. and Luczak, M. J. (2007). A simple solution to the k-core problem. *Random Structures and Algorithms*, 30(1-2):50–62.

[29] Janson, S. and Luczak, M. J. (2009). A new approach to the giant component problem. *Random Structures and Algorithms*, 34(2):197–216.

[30] Janson, S., Łuczak, T., and Rucinski, A. (2000). *Random Graphs.* Wiley, New York.

[31] Jerrum, M. and Sorkin, G. B. (1993). Simulated annealing for graph bisection. In *Proceedings of 34th Annual Symposium on Foundations of Computer Science, 1993*, pages 94–103. IEEE.

[32] Khot, S. (2004). Ruling out PTAS for graph min-bisection, densest subgraph and bipartite clique. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 136–145. IEEE.

[33] Luczak, M. J. and McDiarmid, C. (2001). Bisecting sparse random graphs. *Random Structures and Algorithms*, 18(1):31–38.

[34] Mezard, M., Parisi, G., and Virasoro, M. (1987). *Spin Glass Theory and Beyond, An Introduction to the Replica Method and Its Applications.* World Scientific.

[35] Molloy, M. and Reed, B. (1995). A critical-point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6(2-3):161–179.

[36] Poljak, S. and Tuza, Z. (1995). Maximum cuts and large bipartite subgraphs. *DIMACS Series*, 20:181–244.

[37] Siganos, G., Faloutsos, M., Faloutsos, P., and Faloutsos, C. (2003). Power laws and the AS-level internet topology. *IEEE/ACM Transactions on Networking*, 11(4):514–524.