

Expanders in Power Law Graphs

by

Anton Cherniavskiy

M.Sc., Kyiv Polytechnic Institute, Ukraine, 2013

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
School of Computing Science
Faculty of Applied Sciences

© Anton Cherniavskiy 2018
SIMON FRASER UNIVERSITY
Fall 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

Approval

Name: Anton Cherniavskiy
Degree: Master of Science (Computing Science)
Title: Expanders in Power Law Graphs
Examining Committee: **Chair:** Binay Bhattacharya
Professor

Valentine Kabanets
Senior Supervisor
Associate Professor

Andrei Bulatov
Supervisor
Professor

Bojan Mohar
Internal Examiner
Professor
Department of Mathematics

Date Defended: October 22, 2018

Abstract

Random power-law graphs on n vertices can be defined in different ways. One model we study describes graphs where the expected number of vertices of degree x is proportional to a power law $1/x^\beta$, for constant $\beta > 0$. In another model, the exact degree sequence follows the power-law distribution and each vertex i has degree pn/i^β , for $0 < p \leq 1$ and $\beta \geq 0$.

We show that for these models, power-law graphs contain “large” edge and vertex expanders. Those are graphs in which all subsets of vertices up to a certain size have, respectively, many outgoing edges or large vertex boundary. We also explore the trade-offs between expansion of the subsets and their maximum size.

Our findings agree with and complement known results about the presence of linear size expanders in Erdős-Rényi graphs and about the connected components of power-law graphs.

Keywords: random power-law graphs; expanders; expansion property; diameter

Dedication

This work is wholeheartedly dedicated to my dear parents and family.

They always provide me with unconditional love and lift my spirits.

Acknowledgements

First and foremost, I would like to express the deepest gratitude to my senior supervisor Valentine Kabanets. I recognize that this work would be impossible without him giving me this opportunity and providing invaluable guidance along the way. Valentine also introduced me to complexity theory beyond the P versus NP problem and helped in exploring its deep connections to many other fields. Now I have a feeling of great respect for this area and people who work in it.

I am thankful to my supervisor Andrei Bulatov for his support and helpful advices on how to refine this thesis. In addition, I would like to thank Bojan Mohar for agreeing to examine this thesis, as well as Binay Bhattacharya for being the defence chair.

And to everyone who met me with hospitality and kindness — I appreciate that.

Table of Contents

Approval	ii
Abstract	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Related Work	2
1.2 Contribution of the Thesis	4
1.3 Our Methods	5
1.4 Thesis Structure	5
2 Preliminaries	6
2.1 Probability Theory and Inequalities	6
2.1.1 Union Bound	6
2.1.2 Chernoff Bounds	6
2.1.3 Combinations	7
2.1.4 Stirling's Approximation and the Number of Perfect Matchings . . .	7
2.1.5 The Upper Bounds for the Sum of a Special Series	7
2.2 Generalized Harmonic Numbers and Riemann Zeta Function	10
2.3 Gilbert and Erdős-Rényi Random Graphs	11
2.4 Expander Graphs	12
2.4.1 Combinatorial Expansion	12
2.4.2 Spectral Expansion	13
2.4.3 Connections Between Different Types of Expansion	13

2.4.4	Construction of Expanders	13
2.5	Expansion of Random Regular Graphs	14
2.5.1	Vertex Expansion of Leftregular Bipartite Graphs	14
2.5.2	Edge Expansion of Regular Graphs	14
2.5.3	Vertex Expansion of Regular Graphs	15
2.6	Edge Expansion of Random Graphs With Given Degrees	16
2.7	Vertex Expanders in Locally Sparse Graphs and $G(n, p)$	18
3	Random Power-Law Graphs	20
3.1	Models of Graphs	20
3.1.1	Coin Toss Model	20
3.1.2	m Edges Model	21
3.1.3	Permutation Model	21
3.1.4	Full Graph Selection Model	21
3.2	Coin Toss Power-Law Model	21
3.2.1	Maximum Degree, Expected Size, and the Number of Edges	22
3.2.2	Expected Volume and Average Degree	23
3.3	Permutation Power-Law Model	23
3.4	k-CNF Power-Law Model	24
3.4.1	Noncentral Hypergeometric Distribution	24
3.5	“Octopus” Power-Law Graph Model	26
3.5.1	The Second-Order Average Degree of “Octopus” Graphs	27
3.5.2	Diameter of “Octopus” Graphs	27
4	Expanders in Power-Law Graphs	28
4.1	Coin Toss Model	28
4.2	Permutation Model	37
4.2.1	Error from Rounding Degrees	42
4.3	“Octopus” Graphs	42
4.4	Diameter of Vertex Expanders	44
5	Comparison of the Graph Models	46
5.1	Overview of the Expansion Properties	46
5.1.1	Previously Known Results	46
5.1.2	Our Results	46
5.2	Uniform Random Graphs	47
5.3	Connected Components and the Coin Toss Model	47
5.4	Coin Toss and Permutation Models	47
6	Conclusions	48

Bibliography	49
Appendix A Miscellaneous	53
A.1 Approximations	53
A.1.1 Euler-Maclaurin Formula	53
A.1.2 Lambert Function	54
A.2 Finding Expanding Subsets In Locally Sparse Graphs	54

List of Tables

Table 1.1	Comparison of power-law graph models	4
Table 1.2	Consistency of sizes between our results and the other papers	5
Table 3.1	Properties of power-law graphs [ACL01]	22
Table 3.2	Properties of the “octopus” graphs [CL04]	26

List of Figures

Figure 1.1	Power-law functions	2
Figure 1.2	Example of a power-law graph and its degree distribution	2
Figure 1.3	13-Paley graph is an example of an expander	3
Figure 2.1	Riemann zeta function	11

Chapter 1

Introduction

A power-law distribution is widespread and has been studied for almost a century [Lot26]. Its natural examples range from Pareto principle [Par97] to the distribution of species within genera of plants [Yul25], population of cities [Zip49], and magnitude of earthquakes [GR54]. As for computing science, many industrial SAT (boolean satisfiability) instances that reflect real-world problems were found to follow a power law [ABL09a]. Another notable example is Barabási-Albert model [BA99]: it describes preferential attachment processes known as “the rich get richer”, and implicitly develops a power-law distribution with the exponent 3.

Generally, a power law is a relation $f(x) = ax^b$. A few examples are shown in Figure 1.1. One of its important characteristics is that it is scale-free: $f(c_1x) = (ac_1^b)x^b = c_2f(x)$. For instance, scale-free networks like the Internet topology and social networks are of a special interest: they preserve the overall properties at any scale, resulting in their high resistance to accidental failures [BB03, FFF99].

Power-law graphs are those with either degrees or degree frequencies being proportional to a power law $x^{-\beta}$, where $\beta \geq 0$ is a constant. Figure 1.2 illustrates a power-law graph with $n = 200$ vertices, each vertex i having degree $n^{0.6}/i^{0.4}$.

Graph expansion is another basic concept of this work. It was introduced in 1960s [KB67], but expanders, graphs having a high expansion, were later rediscovered and received their name in 1973 [Pin73]. One example of such a sparse yet well-connected graph would be a Paley graph shown in Figure 1.3.

Expanders were proven to be beneficial for solving routing problems. Discovered routing schemes have robustness and path diversity close to those of the underlying graph [FGRV14], and achieve an optimal congestion in case of power-law graphs [GMS03], all this while using only linear number of edges. Decomposing a graph of arbitrary density into a collection of edge expanders is a base of some divide-and-conquer algorithms [MS17]. Fast convergence of random walks on expanders [Mih89] helps with graph exploration. In addition, expanders often arise when justifying the results like SAT lower bounds [AHI05, ABBO⁺05, PRST16], the hardness of pseudorandom generators [ABSRW04] for various proof systems, and even PCP theorem [Din07].

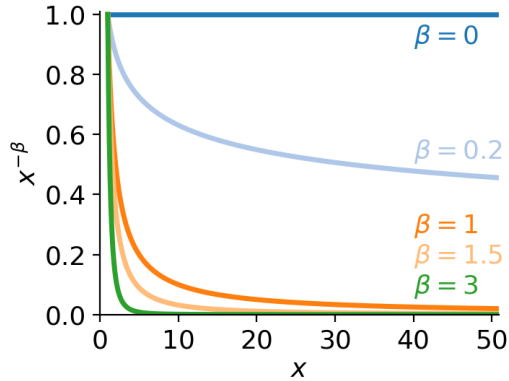


Figure 1.1: Power-law functions

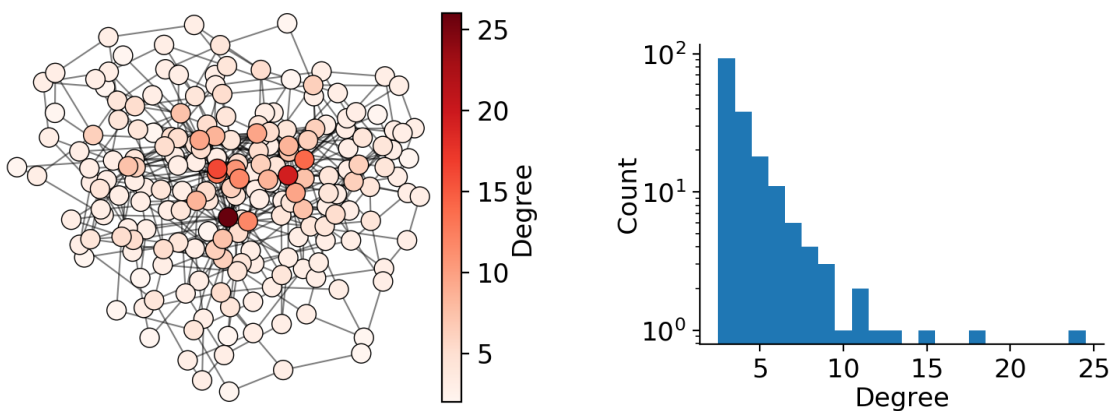


Figure 1.2: Example of a power-law graph and its degree distribution

The original motivation for this research was trying to learn about SAT formulas with power-law structure. Intuitively, such an additional information may lead to faster algorithms for some restricted families of formulas.

Another intriguing idea was about the presence of expanders in the variable incidence graphs (VIG) of SAT formulas. Expanders would represent tight and relatively short connections between the variables. Thus one would typically assume that if there are no large expanders in VIG, it might indicate that satisfiability of the formula could be efficiently decided, e.g., via decomposition.

In this thesis, we work with several models of random power-law graphs, study their properties and trade-offs between the parameters, and show existence of expanding subgraphs for different ranges of the exponent β .

1.1 Related Work

The following results inspired us to look for expanders inside power-law graphs.

Aiello, Chung, and Lu [ACL01] introduced a random graph model for power-law graphs and described different properties, including connectivity and emergence of giant connected

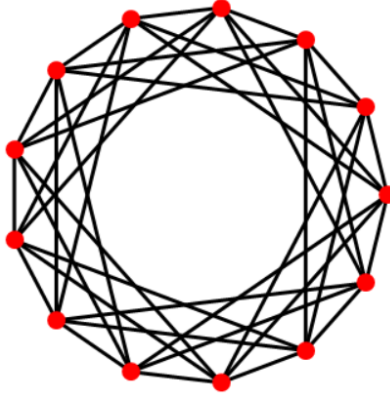


Figure 1.3: 13-Paley graph is an example of an expander

components. This model is asymptotically equivalent to our coin toss model from Section 3.2. The difference is that our model defines the expected degree sequence rather than the exact one, and that is crucial for our proof.

Chung and Lu [CL04] found the average distance in random graphs with given expected degree sequences, both general and power-law with $2 < \beta < 3$. The latter produces the “octopus” graphs described in Section 3.5.

Ansotegui, Bonet, and Levy [ABL09b] presented a power-law model which was shown to fit well industrial SAT instances used in recent international competitions for SAT solvers. They also focused on fitting the SAT instances, i.e., estimating the appropriate distributions that would produce analogous degree sequences. We talk about this model in Section 3.4.

Krivelevich [Kri18] looked for a linearly sized expanders inside graphs that are “locally sparse”, as well as inside random $G(n, p)$ graphs. The paper also contained the algorithm for actually finding the expanding subgraphs.

Finally, Mihail, Saberi, and Tetali [MST06] capitalized on the structure of power-law graphs by displaying that one can discover the nodes via a random walk with lookahead in sublinear time. Considering another work by Mihail [Mih89], this result also hints on possible expansion properties of graphs with paths of constant length replaced by new edges.

As can be seen, there is a substantial amount of research done concerning separately power-law models and expander graphs due to their wide popularity and usefulness. Gkantsidis, Mihail, and Saberi [GMS03] made significant steps in the direction of combining these two topics. They considered random power-law graphs with the exponent $2 < \beta < 3$, degrees of vertices between 3 and $O(\sqrt{n})$, and volume $O(n)$. They also had to slightly modify graph construction used by Aiello et al. [ACL01] in order to ensure certain connectivity properties. These graphs were shown to have conductance $\Theta(1)$, which generalizes the notion of $(n/2, \Theta(1))$ edge expansion.

Table 1.1: Comparison of power-law graph models

Model	Object		Definition
Permutation model	exact	degrees	$\deg(i) = \frac{pn}{i^\beta}$
Model 3 [ACL01]		frequencies of degrees	$ \{i \in V \mid \deg(i) = x\} = \frac{e^\alpha}{x^\beta}$
Coin toss model	expected	frequencies of degrees	$\mathbb{E}_G[\{i \in V \mid \deg(i) = x\}] = \frac{e^\alpha}{x^\beta}$
Model 4 [CL04]		degrees	$\mathbb{E}_G[\deg(i)] = w_i = ci^{-1/(\beta-1)}, \beta > 2$

1.2 Contribution of the Thesis

We defined the models of power-law graphs, which would complement the ones studied earlier. Table 1.1 presents their comparison. The permutation model was chosen so as to fit the uniformly random case when the exponent $\beta = 0$.

The main result is that, under these models, the subgraphs containing the vertices of sufficiently large degrees are edge or vertex expanders w.h.p.

More precisely, in the coin toss model, if $\beta < 1$, actually the whole graph is an edge expander. For $1 \leq \beta \leq 1.6$ we have a linear size expanding subgraph, and for $\beta > 1.6$ the size of the expander is only $\Theta(n^{1/\beta})$. In all the cases edge expansion is close to one half of the expected average degree d , which is linear in the size of the subgraph.

In the permutation model, the case $\beta = 0$ matches the previously known fact that the whole graph has edge and vertex expansion almost $d - 2$. When $\beta > 0$, we generalize the argument to show existence of the edge expanders of size $n/2$ with expansion $d/2$, but the average degree d deteriorates for larger β . Also, if $\beta > 1$, there is an additional constraint $p\zeta(\beta) > 2$ which essentially limits $\beta \leq 1.72$. Meanwhile, vertex expansion of the subgraph with vertices of degree at least d_0 is almost $d_0/2 - 2$ whenever $\beta > 0$.

Obtained results about linear size expanders inside power-law graphs resemble existing result of the same nature for Erdős-Rényi graphs [Kri18]. The size of our expanders is also comparable to the sizes of the largest connected components in power-law graphs [ACL01], i.e., we say that the largest components are not just connected, but “highly connected”. Table 1.2 summarizes these details.

We proceed by showing that the core of the “octopus” graph with the vertices of degree at least $n^{1/\log \log n}$ is an edge expander w.h.p. and it contains a large vertex expander.

As a side result, we prove the logarithmic diameter of vertex expanders with only small expanding subsets of size at most ϵn , for some constant $\epsilon > 0$, as opposed to $\epsilon = 1/2$ in the canonical result.

Table 1.2: Consistency of sizes between our results and the other papers

β	(0; 1.6]	(1.6; 1.72]	(1.72; 3.48)	(3.48; ∞)
The largest components in power-law graphs [ACL01]	the giant component, $\Theta(n)$			$O\left(n^{2/\beta} \log n\right)$
Our edge expanders in coin toss model	$\Theta(n)$	$\Theta\left(n^{1/\beta}\right)$		
Vertex/edge expanders in $G(n, p)$ [Kri18]	$\Theta(n)$			
Our edge expanders in permutation model	$\Theta(n)$	—		

To sum up, our findings provide better understanding of the structure of power-law graphs from some general families for a wide range of parameters. This knowledge can be further used to employ any techniques applicable to expanders on these power-law graphs.

1.3 Our Methods

For the coin toss model, we first obtain the necessary lower bounds on the expected average degree. This is done by approximating the expected size of an arbitrary cut, applying Chernoff concentration bounds, and following the common argument for edge expansion. Then we decide the size of an expanding subgraph and try to keep it linear in the size of the whole graph by choosing an appropriate minimum degree of vertices to be included in this subgraph.

While working with the permutation model, we adopt and generalize the existing approaches from Section 2.5.2 and 2.5.3 for edge and vertex expansion of regular graphs.

Throughout this work, we deal with varying approximations of harmonic numbers, so we have to treat different ranges of the exponent β separately. Lastly, our proof of small diameter of vertex expanders resembles a common technique for graph decomposition.

1.4 Thesis Structure

In Chapter 2 we present all the necessary definitions, approximations, and known results about expanders, on which this research is based. Chapter 3 contains the detailed description of the main models of power-law graphs used in this work. In Chapter 4 we show the existence of expanding subgraphs for the coin toss and permutation models, and analyze the diameter of graphs with vertex expansion of small sets. Finally, we compare the behavior of the random graphs under different models in Chapter 5, including their diameters and sizes of expanders and connected components.

Chapter 2

Preliminaries

Here we explain the basic definitions as well as the notation used in the following chapters.

2.1 Probability Theory and Inequalities

We say an event $E(n)$ over a sample space Ω happens with high probability (w.h.p.) when

$$\lim_{n \rightarrow \infty} \Pr_{\Omega}[E(n)] = 1 \quad (2.1)$$

2.1.1 Union Bound

If Ω is a sample space and E_1, \dots, E_n are events over Ω , then

$$\Pr_{\Omega} \left[\bigcup_{i=1}^n E_i \right] \leq \sum_{i=1}^n \Pr_{\Omega} [E_i] \quad (2.2)$$

2.1.2 Chernoff Bounds

Let $X_1, \dots, X_n \in \{0, 1\}$ be independent random variables, and $X = \sum_{i=1}^n X_i$ with expected value $\mu = \mathbb{E}[X] = \sum_{i=1}^n \Pr[X_i = 1]$.

$$\Pr[X \geq (1 + \delta)\mu] \leq \begin{cases} \exp(-\delta^2\mu/(2 + \delta)) & \text{if } \delta > 1, \\ \exp(-\delta^2\mu/3) & \text{if } 0 < \delta \leq 1. \end{cases} \quad (2.3)$$

$$\Pr[X \leq (1 - \delta)\mu] \leq \exp(-\delta^2\mu/2), \text{ for } 0 < \delta < 1 \quad (2.4)$$

We can combine these two:

$$\Pr[|X - \mu| \geq \delta\mu] \leq 2 \exp(-\delta^2\mu/3), \text{ for } 0 < \delta < 1 \quad (2.5)$$

2.1.3 Combinations

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} = \frac{n!}{k!(n-k)!} \leq \left(\frac{en}{k}\right)^k \quad (2.6)$$

It is easy to show:

$$\frac{n}{k} \leq \frac{n-m}{k-m}, \text{ for } 0 \leq m < k \leq n, \text{ so } \left(\frac{n}{k}\right)^k = \frac{n}{k} \cdots \frac{n}{k} \leq \frac{n}{k} \frac{n-1}{k-1} \cdots \frac{n-k+1}{1} = \binom{n}{k}$$

$$\binom{n}{k} = \frac{n}{k} \frac{n-1}{k-1} \cdots \frac{n-k+1}{1} \leq \frac{n^k}{k!} = \frac{k^k}{k!} \left(\frac{n}{k}\right)^k \leq \left(\frac{en}{k}\right)^k$$

The last step uses the Maclaurin series of the exponential function:

$$e^k = \sum_{n=0}^{\infty} \frac{k^n}{n!} \geq \frac{k^k}{k!} \quad (2.7)$$

2.1.4 Stirling's Approximation and the Number of Perfect Matchings

Let $M(m)$ denote the number of perfect matchings on the set of even size m :

$$M(m) = (m-1)!! = (m-1)(m-3)\dots(3)(1) = \frac{m!}{2^{m/2}(m/2)!} \quad (2.8)$$

We can use Stirling's approximation:

$$m! = (1+o(1))\sqrt{2\pi m}(m/e)^m \quad (2.9)$$

$$M(m) = (1+o(1))\frac{\sqrt{2\pi m}(m/e)^m}{2^{m/2}\sqrt{\pi m}(m/2e)^{m/2}} = (1+o(1))\sqrt{2}(m/e)^{m/2} \quad (2.10)$$

2.1.5 The Upper Bounds for the Sum of a Special Series

The following sum S_n occurs in several proofs. We will need to upper bound it for $n \rightarrow \infty$ and some constants $0 < \alpha \leq 1/2$, $c_1 > 0$, and $c_2 > 0$:

$$S_n = \sum_{s=1}^{\alpha n} (c_1(s/n)^{c_2})^s \quad (2.11)$$

Each subsequent bound will require more rigorous analysis.

Proposition 2.1. $S_n < 1$, when $c_2 \geq 1 + \log c_1$.

Proof. $S_n = \sum_{s=1}^{\alpha n} (c_1(s/n)^{c_2})^s \leq \sum_{s=1}^{\alpha n} (c_1\alpha^{c_2})^s \leq \sum_{s=1}^{\infty} 2^{-c_3 s} \leq \sum_{s=1}^{\infty} 2^{-s} < 1$, where $c_3 = c_2 - \log c_1$ is a constant greater or equal 1:

$$c_3 \geq 1 \iff c_2 \geq 1 + \log c_1 \quad (2.12)$$

$$c_1\alpha^{c_2} \leq c_1 2^{-c_2} = 2^{-c_3} \leq 2^{-1} \quad (2.13)$$

□

Proposition 2.2. $S_n \leq o(1)$ for some sufficiently small α .

Proof. First, we pick α so that $(c_1\alpha^{c_2}) < 1/10$, using some constant $c_3 > 10$:

$$\alpha = (c_1c_3)^{-1/c_2} \quad (2.14)$$

Then we consider small and large values of s separately. Define $0 < \epsilon < 1$ such that

$$\epsilon < (1 - \epsilon)c_2 \iff \epsilon < \frac{c_2}{1 + c_2} \quad (2.15)$$

$$S' = \sum_{s=1}^{n^\epsilon} (c_1(s/n)^{c_2})^s \leq \sum_{s=1}^{n^\epsilon} \left(\frac{c_1}{n^{(1-\epsilon)c_2}} \right)^s \leq n^\epsilon \frac{c_1}{n^{(1-\epsilon)c_2}} = o(1) \quad (2.16)$$

$$S'' = \sum_{s=n^\epsilon+1}^{\alpha n} (c_1(s/n)^{c_2})^s < \sum_{s=n^\epsilon+1}^{\alpha n} 10^{-s} \leq \frac{\alpha n}{10^{n^\epsilon}} = o(1) \quad (2.17)$$

$$S_n = S' + S'' \leq o(1). \quad \square$$

Proposition 2.3. $S_n \leq o(1)$ for any $\alpha \leq 1/2$ and $c_2 > \max\{1, \log c_1\}$.

Moreover, c_1 and c_2 can be superconstants in terms of n , as long as $c_1 = o(n^{c_2-1})$.

Proof. The argument is identical to the known one [GMS03].

For $s \geq 1$, let's define $f(s) = (c_1(s/n)^{c_2})^s$, then $S_n = \sum_{s=1}^{\alpha n} f(s)$. The first two derivatives will help us understand the behavior of $f(s)$.

First, the simple case is when $c_1 = c_2 = 1$:

$$\begin{aligned} \frac{d}{ds} \left(\frac{s}{n} \right)^s &= \frac{d}{ds} \exp \left(s \ln \frac{s}{n} \right) = \left(\frac{s}{n} \right)^s \left(\ln \frac{s}{n} + s \frac{1/n}{s/n} \right) = \left(\frac{s}{n} \right)^s \left(\ln \frac{s}{n} + 1 \right) \\ \frac{d^2}{ds^2} \left(\frac{s}{n} \right)^s &= \left(\frac{s}{n} \right)^s \left(\frac{1}{s} + \left(1 + \ln \frac{s}{n} \right)^2 \right) > 0 \end{aligned}$$

We can see that $\left(\frac{s}{n} \right)^s$ is concave up for any $s \geq 1$ and attains its minimum at $s = n/e$.

Then we perform similar steps for $f(s)$:

$$\begin{aligned}
f'(s) &= \frac{d}{ds} \left(c_1 \left(\frac{s}{n} \right)^{c_2} \right)^s = \frac{d}{ds} \exp \left(s \ln \left(c_1 \left(\frac{s}{n} \right)^{c_2} \right) \right) = \\
&= f(s) \left(\ln \left(c_1 \left(\frac{s}{n} \right)^{c_2} \right) + s \frac{c_1 c_2 (s/n)^{c_2-1} (1/n)}{c_1 (s/n)^{c_2}} \right) = \\
&= f(s) \left(\ln \left(c_1 \left(\frac{s}{n} \right)^{c_2} \right) + c_2 \right) = \\
&= f(s) c_2 \left(\ln \frac{c_1^{1/c_2} s}{n} + 1 \right) = \\
&= f(s) c_2 \left(\ln \frac{s}{n} + \frac{\ln c_1}{c_2} + 1 \right)
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
f''(s) &= \frac{d^2}{ds^2} \left(\frac{s}{n} \right)^s = \left(\frac{s}{n} \right)^s \left(\frac{1}{s} + \left(1 + \ln \frac{s}{n} \right)^2 \right) = \\
&= \frac{c_2}{s} f(s) + c_2 \left(\ln \frac{s}{n} + 1 \right) f'(s) + (\ln c_1) f'(s) = \\
&= \frac{c_2}{s} f(s) + c_2 \left(\ln \frac{s}{n} + \frac{\ln c_1}{c_2} + 1 \right) f'(s) = \\
&= \frac{c_2}{s} f(s) + \frac{(f'(s))^2}{f(s)}
\end{aligned} \tag{2.19}$$

Define s_0 to be an extreme point of $f(s)$:

$$f'(s_0) = 0 \tag{2.20}$$

$$\ln \frac{c_1^{1/c_2} s}{n} = -1 \tag{2.21}$$

$$s_0 = \frac{n}{e c_1^{1/c_2}} \tag{2.22}$$

And verify that $f''(s) > 0$ for any $s \geq 1$:

$$\frac{c_2}{s} f(s) + \frac{(f'(s))^2}{f(s)} > 0 \tag{2.23}$$

$$\frac{c_2}{s} + \left(\frac{f'(s)}{f(s)} \right)^2 > 0 \tag{2.24}$$

$$\begin{cases} f'(s) < 0 & \text{if } s < s_0, \\ f'(s) = 0 & \text{if } s = s_0, \\ f'(s) > 0 & \text{if } s > s_0. \end{cases} \tag{2.25}$$

$$f''(s) > 0, \text{ for } s \in [1; \alpha n] \tag{2.26}$$

Therefore, $f(s)$ is concave up on the whole interval $[1; \alpha n]$ with these endpoints:

$$f(1) = \frac{c_1}{n^{c_2}} \quad (2.27)$$

$$f(\alpha n) = (c_1 \alpha^{c_2})^{\alpha n} \leq (c_1 2^{-c_2})^{\alpha n} \quad (2.28)$$

$f(1) = o(1/n)$ when $c_2 > 1$, and $f(\alpha n) = o(1/n)$ when $c_1 2^{-c_2} < 1 \iff c_2 > \log c_1$.

Additionally, both c_1 and c_2 might be superconstants in terms of n , if this holds:

$$c_1 = o\left(n^{c_2-1}\right) \quad (2.29)$$

To sum up, given $c_2 > \max\{1, \log c_1\}$,

$$S_n = \sum_{s=1}^{\alpha n} f(s) \leq \sum_{s=1}^{\alpha n} \max\{f(1), f(\alpha n)\} = o\left(\frac{\alpha n}{n}\right) = o(1). \quad \square$$

Note that if c_2 in Proposition 2.3 is not large enough, there is no advantage in terms of the size of expanding sets compared to Proposition 2.2:

$$c_1 \alpha^{c_2} < 1 \iff \alpha < c_1^{-1/c_2} \quad (2.30)$$

2.2 Generalized Harmonic Numbers and Riemann Zeta Function

Generalized harmonic numbers are defined as

$$H_{n,m} = \sum_{k=1}^n \frac{1}{k^m} \quad (2.31)$$

We will extensively use these approximations for large n :

$$H_{n,\beta} \approx \begin{cases} \frac{n^{1-\beta}}{1-\beta} & \text{if } \beta < 1, \\ \ln n & \text{if } \beta = 1, \\ \zeta(\beta) & \text{if } \beta > 1. \end{cases} \quad (2.32)$$

The zeta function was introduced in 1737 by Euler, later Riemann allowed its argument to be a complex number. *Riemann zeta function* is defined for all $s \in \mathbb{C}$ with $\Re(s) > 1$, see Figure 2.1:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (2.33)$$

For $x \in \mathbb{R}$, we know that $\zeta(x) \approx 1 + 2^{-x}$ when $x \gg 1$, and $\lim_{x \rightarrow \infty} \zeta(x) = 1$.

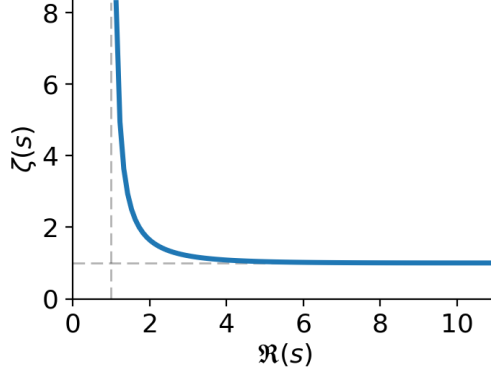


Figure 2.1: Riemann zeta function for $s \in \mathbb{C}$ with $\Re(s) > 1$ and $\Im(s) = 0$

For $m > 1$ as $n \rightarrow \infty$, applying the Euler-Maclaurin formula (A.1) to the remainders $\sum_{n>N} n^{-s}$ of the sum (2.33) provides the following connection:

$$\zeta(s) = H_{N,s} + \frac{1}{(s-1)N^{s-1}} + \frac{1}{2N^s} + O\left(\frac{1}{N^{s+1}}\right), \quad \text{for some } N \gg 1 \quad (2.34)$$

$$H_{n,m} = \zeta(m) - \frac{1}{(m-1)n^{m-1}} - \frac{1}{2n^m} - O\left(\frac{1}{n^{m+1}}\right) = \zeta(m) - o(1) \quad (2.35)$$

2.3 Gilbert and Erdős-Rényi Random Graphs

The most commonly used models of uniform random graphs are $G(n, p)$ and $G(n, m)$ introduced by Gilbert [Gil59] and by Erdős and Rényi [ER59].

$G(n, p)$ contains all graphs where each pair of vertices is connected by an edge with probability p . We denote the maximum number of edges $N = \binom{n}{2}$, then the expected number of edges in such graphs is pN .

$$\Pr_{G \sim G(n,p)} [G \text{ has } m \text{ edges}] = \binom{N}{m} p^m (1-p)^{N-m} \quad (2.36)$$

$$\Pr_{G \sim G(n,p)} [\deg(v) = k] = \binom{n-1}{k} p^k (1-p)^{n-k-1}, \quad \text{for } v \in V(G) \quad (2.37)$$

On the other hand, $G(n, m)$ describes the graphs with n vertices and exactly m edges. They may be generated by successively selecting m pairs of not yet connected vertices, so that each of $\binom{N}{m}$ possible graphs is chosen uniformly at random.

Erdős and Rényi [ER60] depict the following structural properties of $G(n, p)$ graphs. When $p = o(1/n)$, the graph is simply a disjoint union of trees. Then consider the case $p = c/n$, where c is some constant. If $c < 1$, the size of the largest component is just $O(\log n)$,

at the critical point $c = 1$ its size is $\Theta(n^{2/3})$, and a unique giant component of size $\Theta(n)$ emerges when $c > 1$. Finally, the graphs are disconnected w.h.p. when $p < (1 - \epsilon) \log n/n$ and become connected w.h.p. when $p > (1 + \epsilon) \log n/n$, for any $\epsilon > 0$ [ER59].

As per Chung and Lu [CL01, CL04], the diameter of $G(n, p)$ is $\Theta(\log n)$ when $p = c/n$, for some constant $c > 1$, and $(1 + o(1)) \frac{\log n}{\log np}$ if $p = \omega(1/n)$. The diameter of a disconnected graph is assumed to be the maximum diameter of its connected components.

If we choose $m = pN$, $G(n, p)$ and $G(n, m)$ models are almost identical for large n and $p > \log n/n$, but some important properties might be different for smaller p . For example, let $p = 3/n$ and so $m = \frac{3}{2}(n - 1)$. In this case $G(n, p)$ is disconnected w.h.p., while $G(n, m)$ yields a connected graph with high expansion [Mah09] as almost 3-regular.

Independence of edges of $G(n, p)$ makes it similar to our coin toss model for power-law graphs, and $G(n, m)$ corresponds more to m edges model and permutation model, which will be defined in Chapter 3.

2.4 Expander Graphs

2.4.1 Combinatorial Expansion

First, we look at the combinatorial notions of expansion.

Definition 2.4. A graph G is called $(\alpha n, \gamma)$ *edge expander* if for some $\gamma > 0$

$$\min_{\substack{S \subset V \\ 0 < |S| \leq \alpha n}} \frac{|\partial S|}{|S|} \geq \gamma, \quad (2.38)$$

where the left-hand side expression with $\alpha = 1/2$ is *the Cheeger constant* $h(G)$, and

$$\partial S = e(S, V \setminus S) = \{(u, v) \in E \mid u \in S, v \in V \setminus S\}. \quad (2.39)$$

Definition 2.5. Analogously, $(\alpha n, \gamma)$ *vertex expander* satisfies the following condition:

$$\min_{\substack{S \subset V \\ 0 < |S| \leq \alpha n}} \frac{|N_G(S)|}{|S|} \geq \gamma, \quad (2.40)$$

where $N_G(S)$ is the external neighborhood of S in G , i.e.,

$$N_G(S) = \{v \in V \setminus S \mid \exists u \in S, (u, v) \in E\}. \quad (2.41)$$

Definition 2.6. $(\alpha n, \gamma)$ *unique-neighbor expander* is a graph satisfying

$$\min_{\substack{S \subset V \\ 0 < |S| \leq \alpha n}} \frac{|\{v \in V \setminus S \mid v \text{ is adjacent to exactly one } u \in S\}|}{|S|} \geq \gamma. \quad (2.42)$$

2.4.2 Spectral Expansion

Let the vertex v have $\deg(v)$ and define $T = \text{diag}(\deg(v))$. The Laplacian of G without loops or multiple edges is $\mathcal{L} = T^{-\frac{1}{2}}LT^{-\frac{1}{2}}$. If G is d -regular, then $\mathcal{L} = I - A/d$ [Chu97].

$$L_{u,v} = \begin{cases} \deg(v) & \text{if } u = v, \\ -1 & \text{if } (u, v) \in E(G), \\ 0 & \text{otherwise.} \end{cases} \quad (2.43)$$

$$\mathcal{L}_{u,v} = \begin{cases} 1 & \text{if } u = v \text{ and } \deg(v) \neq 0, \\ -\frac{1}{\sqrt{\deg(u)\deg(v)}} & \text{if } (u, v) \in E(G), \\ 0 & \text{otherwise.} \end{cases} \quad (2.44)$$

The spectrum of G are the eigenvalues of \mathcal{L} $\{\lambda_1, \dots, \lambda_n\}$, $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Eigenvalue 0 has eigenvector $T^{1/2}\mathbf{1}$.

Let $g : V \rightarrow \mathbb{R}$ be an arbitrary function, $g = T^{1/2}f$ for some f , and both g and f are viewed as column vectors. The Rayleigh quotient is

$$\frac{\langle g, \mathcal{L}g \rangle}{\langle g, g \rangle} = \frac{\sum_{u \sim v} (f(u) - f(v))^2}{\sum_v f(v)^2 \deg(v)} \quad (2.45)$$

and it is related to the spectral expansion of G , also known as the spectral gap,

$$\lambda_G = \lambda_2 = \inf_{f \perp T^{\frac{1}{2}}\mathbf{1}} \frac{\sum_{u \sim v} (f(u) - f(v))^2}{\sum_v f(v)^2 \deg(v)} \quad (2.46)$$

2.4.3 Connections Between Different Types of Expansion

Trivially, vertex expansion γ_1 implies edge expansion γ_1 . For d -regular graphs, edge expansion γ_2 implies vertex expansion γ_2/d .

The Cheeger inequalities [Che69, HLW06, Chu07] provide the link between combinatorial and spectral expansion:

$$\frac{h^2(G)}{2} \stackrel{(1)}{\leq} \lambda_G \stackrel{(2)}{\leq} 2h(G), \quad (2.47)$$

$$\text{or equivalently } \frac{\lambda_G}{2} \leq h(G) \leq \sqrt{2\lambda_G} \quad (2.48)$$

2.4.4 Construction of Expanders

There are many ways to generate expander graphs: algebraic constructions and zig-zag product [HLW06, Vad12], lifts [AL06], splicers and selectors [FGRV14], local edge flips [ABL⁺16], random coverings of a fixed good expander [Pud15], and so on.

In Chapter 4 we show yet another way — random power-law graphs contain expanding induced subgraphs with vertices of large degrees.

2.5 Expansion of Random Regular Graphs

Theorem 2.7 (Theorem 4.16 [HLW06]). $\forall d \geq 3 \forall \delta > 0 \exists \gamma, \alpha > 0$ such that a random d -regular graph on n vertices is $(\alpha n, \gamma)$ edge expander and $(\alpha n, \gamma)$ vertex expander w.h.p., $\gamma = d - 2 - \delta$.

2.5.1 Vertex Expansion of Leftregular Bipartite Graphs

Here G is assumed to be a bipartite multigraph with n vertices on each side.

Definition 2.8. A graph G is called d -leftregular if each left-vertex is connected to exactly d vertices from the right. In this case expansion is checked for all subsets S of left-vertices.

Theorem 2.9 ([Vad12]). $\forall d \geq 2 \exists \alpha > 0 \forall n$: a uniformly random d -leftregular G is $(\alpha n, d - 2)$ vertex expander w.h.p.

Proof. For a subset S of a fixed size $s \leq \alpha n$ from the left side, $N_G(S)$ is a set of at most sd vertices from the right labeled v_1, v_2, \dots, v_{sd} . For each successively chosen v_i from $N_G(S)$,

$$\Pr_{G,S}[v_i \in \{v_1, \dots, v_{i-1}\}] = \frac{|\{v_1, \dots, v_{i-1}\}|}{n} \leq \frac{i-1}{n} \leq \frac{sd}{n}$$

$$\Pr_{G,S}[|N_G(S)| \leq (d-2)s] \leq \Pr_{G,S}[\text{at least } 2s \text{ repeats in } N_G(S)] \leq \binom{sd}{2s} \left(\frac{sd}{n}\right)^{2s}$$

$$\Pr_G[\exists \text{ non-expanding } S] \leq \binom{n}{s} \binom{sd}{2s} \left(\frac{sd}{n}\right)^{2s} \leq \left(\frac{en}{s}\right)^s \left(\frac{esd}{2s}\right)^{2s} \left(\frac{sd}{n}\right)^{2s} = \left(\frac{e^3 d^4 s}{4n}\right)^s$$

Denote $c = (e^3 d^4)/4$ and choose $\alpha = e^{-3} d^{-4}$, so that $c\alpha = 1/4$.

$$\Pr_G[G \text{ is not } (\alpha n, d-2) \text{ vertex expander}] \leq \sum_{s=1}^{\alpha n} \left(c \frac{s}{n}\right)^s = o(1). \quad \square$$

2.5.2 Edge Expansion of Regular Graphs

Theorem 2.10 ([Mag06]). $\forall d \geq 3 \forall \delta > 0 \exists \gamma, \alpha > 0 \forall \text{even } n$: a random d -regular graph $G = (V_G, E_G)$ is $(\alpha n, \gamma)$ edge expander w.h.p., $\gamma = d - 2 - \delta$.

Proof. Consider a new graph $H = (V_H, E_H)$ with dn vertices and E_H being a uniformly random perfect matching. Partition V_H into sets S_1, \dots, S_n of size d , and identify the vertices from each S_i with a single vertex i in G .

Let $S \subseteq V_G = \{1, \dots, n\}$ be such that $|S| = s \leq \alpha n$. We can assume $S = \{1, \dots, s\}$ and it is identified with $S_1 \cup \dots \cup S_s$. Then to pick S we simply choose ds vertices from H .

To upper bound $|\partial(S)|$ we can lower bound the number of edges inside S .

$$\Pr_{G,S}[|e(S, V_G \setminus S)| < \gamma s] = \Pr_{G,S}\left[|e(S, S)| \geq \frac{(d-\gamma)s}{2}\right] \leq \frac{\binom{dn/2}{(d-\gamma)s/2} \binom{dn-(d-\gamma)s}{\gamma s}}{\binom{dn}{ds}}$$

$$\begin{aligned}
\Pr_G[G \text{ is not } (\alpha n, \gamma) \text{ edge expander}] &\leq \sum_{s=1}^{\alpha n} \binom{n}{s} \frac{\left(e \frac{dn}{(d-\gamma)s}\right)^{(d-\gamma)s/2} \left(e \frac{dn-(d-\gamma)s}{\gamma s}\right)^{\gamma s}}{\left(\frac{n}{s}\right)^{ds}} = \\
&= \sum_{s=1}^{\alpha n} \binom{n}{s} \frac{(edn)^{(d-\gamma)s/2} e^{\gamma s} (dn - (d-\gamma)s)^{\gamma s} s^{ds}}{((d-\gamma)s)^{(d-\gamma)s/2} (\gamma s)^{\gamma s} n^{ds}} \leq \\
&\leq \sum_{s=1}^{\alpha n} \binom{n}{s} \frac{(ed)^{ds/2} (edn)^{\gamma s} s^{ds}}{(d-\gamma)^{(d-\gamma)s/2} (\gamma s)^{\gamma s} n^{ds}} \binom{n}{s}^{(d-\gamma)s/2} = \\
&= \sum_{s=1}^{\alpha n} \binom{n}{s} \frac{(ed)^{ds/2+\gamma s}}{(d-\gamma)^{(d-\gamma)s/2} \gamma^{\gamma s}} \left(\frac{s}{n}\right)^{(d-\gamma)s/2} \leq \\
&\leq \sum_{s=1}^{\alpha n} \frac{(ed)^{ds/2+\gamma s}}{(d-\gamma)^{(d-\gamma)s/2}} \left(\frac{e}{\gamma^\gamma}\right)^s \left(\frac{s}{n}\right)^{\left(\frac{d-\gamma}{2}-1\right)s} = \\
&= \sum_{s=1}^{\alpha n} (c_1 (s/n)^{c_2})^s,
\end{aligned}$$

where $c_1 = \frac{e(ed)^{d/2+\gamma}}{\gamma^\gamma (d-\gamma)^{(d-\gamma)/2}} \geq 0$ and $c_2 = \frac{d-\gamma}{2} - 1$.

We need $c_2 > 0$, which requires

$$\gamma < d - 2, \tag{2.49}$$

then by Proposition 2.2,

$$\Pr_G[G \text{ is not } (\alpha n, \gamma) \text{ edge expander}] = o(1). \quad \square$$

We will prove one of our results, Theorem 4.5, in a similar fashion.

Theorem 2.11 ([Bol88]). *Let $d \geq 3$ and $0 < \epsilon < 1$ be such that*

$$(1 - \epsilon) \log(1 - \epsilon) + (1 + \epsilon) \log(1 + \epsilon) > 4/d. \tag{2.50}$$

Then any random d -regular graph G has $h(G) \geq (1 - \epsilon)d/2$ w.h.p.

The proof of this result uses the number of matchings between endpoints of edges not crossing the cut $(S, V \setminus S)$.

Note that these two theorems sacrifice different parameters. While Theorem 2.10 shows how to achieve almost optimal expansion close to $d - 2$, only sufficiently small subsets are expanding. On the contrary, Theorem 2.11 talks about the Cheeger constant $h(G)$, so the subsets may be as large as $n/2$, but the expansion might be as small as $d/2$.

2.5.3 Vertex Expansion of Regular Graphs

Theorem 2.12 ([Rao12]). *$\forall d \geq 3 \forall \delta > 0 \exists \gamma, \alpha > 0 \forall \text{even } n : \text{ a random } d\text{-regular graph } G = (V, E) \text{ is } (\alpha n, \gamma) \text{ vertex expander w.h.p., } \gamma = d/2 - 2 - \delta.$*

Proof. Let E be the union of d uniformly random perfect matchings.

For any subset S of size $|S| = s \leq \alpha n$, and any subset T of $(1 + \gamma)s$ vertices we will bound the probability that $N_G(S) \subseteq T$ under one of the perfect matchings.

Any fixed perfect matching $E' \subset E$ might be viewed as connecting one pair of unmatched vertices sequentially, until every vertex has a match. Let E_i denote the event that i 'th vertex in S is matched to some vertex in T under E' . Of course, E_i is well defined only for $i \leq \lfloor \frac{s}{2} \rfloor$.

For the sake of simplicity, let's assume s is even.

$$\begin{aligned} \Pr_{G,S,T,E'} \left[\bigwedge_{i=1}^{\lfloor s/2 \rfloor} E_i \right] &= \prod_{i=1}^{\lfloor s/2 \rfloor} \Pr_{G,S,T,E'} \left[E_i \mid \bigwedge_{k=1}^{i-1} E_k \right] \leq \prod_{i=1}^{\lfloor s/2 \rfloor} \Pr_{G,S,T,E'} [E_i] \leq \left(\frac{(1+\gamma)s}{n} \right)^{\lfloor s/2 \rfloor} \\ \Pr_{G,S,T} [N_G(S) \subseteq T] &\leq \left(\frac{(1+\gamma)s}{n} \right)^{sd/2} \\ \Pr_G [\exists \text{ non-expanding } S \text{ of size } s] &\leq \binom{n}{s} \binom{n}{(1+\gamma)s} \left(\frac{(1+\gamma)s}{n} \right)^{sd/2} \leq \\ &\leq \left(\frac{en}{s} \right)^s \left(\frac{en}{(1+\gamma)s} \right)^{(1+\gamma)s} \left(\frac{(1+\gamma)s}{n} \right)^{sd/2} = \left(\frac{(en)^{2+\gamma} ((1+\gamma)s)^{d/2}}{n^{d/2} (1+\gamma)^{1+\gamma} s^{2+\gamma}} \right)^s = \\ &= \left(e^{2+\gamma} (1+\gamma)^{d/2-1-\gamma} \left(\frac{s}{n} \right)^{d/2-2-\gamma} \right)^s = \\ &= (c_1 (s/n)^{c_2})^s. \end{aligned}$$

We need $c_2 > 0$, which requires $\gamma < d/2 - 2$. For sufficiently small α ,

$$\Pr_G [G \text{ is not } (\alpha n, \gamma) \text{ vertex expander}] \leq \sum_{s=1}^{\alpha n} (c_1 (s/n)^{c_2})^s = o(1). \quad \square$$

2.6 Edge Expansion of Random Graphs With Given Degrees

Gkantsidis et al. [GMS03] deal with the power-law graphs and their model is based on the one from Aiello et al. [ACL01]. However, this following result holds for random graphs with general degree sequences.

Theorem 2.13 (Lemma 3.2 [GMS03]). *Let (w_1, \dots, w_n) be a sequence of integers and*

$$w_1 \geq w_2 \geq \dots \geq w_n \geq d_{\min} = 3 \quad (2.51)$$

$$D = \sum_{i=1}^n w_i \quad (2.52)$$

Let $G = (V, E)$ be a random graph with a degree sequence (w_1, \dots, w_n) generated according to the permutation model from Section 3.1.3. Then there is a constant $\gamma > 0$ satisfying

$$\begin{cases} \gamma < 1 - 2/d_{\min} \\ \gamma \leq 1 - 8/(3d_{\min}) \\ \gamma \leq 0.0175 \end{cases} \quad (2.53)$$

and w.h.p. the conductance of G is $\min_{\substack{S \subset V \\ 0 < \text{vol}(S) \leq D/2}} \frac{|e(S, V \setminus S)|}{\text{vol}(S)} \geq \gamma$, where $\text{vol}(S) = \sum_{i \in S} w_i$.

Proof. For a constant $\gamma > 0$ a set S with $\text{vol}(S) \leq D/2$ is called “bad” if $\frac{|e(S, V \setminus S)|}{\text{vol}(S)} < \gamma$.

We will show that for some γ , $\Pr_G[\exists \text{ bad } S] \leq o(1)$.

$$\begin{aligned} \Pr_G[\exists \text{ bad } S] &= \sum_{k=d_{\min}}^{D/2} \Pr_G[\exists \text{ bad } S, \text{vol}(S) = k] \leq \\ &\leq \sum_{k=d_{\min}}^{D/2} \binom{D/d_{\min}}{k/d_{\min}} \Pr_G[\text{a fixed } S, \text{vol}(S) = k, \text{ is bad}]. \end{aligned}$$

Let A denote a set of k mini-vertices corresponding to S , and \bar{A} to be $(D - k)$ mini-vertices that correspond to $V \setminus S$. Let $B_A \subset A$ be a set of mini-vertices matched to the ones from \bar{A} , and similarly $B_{\bar{A}} \subset \bar{A}$ is matched to some mini-vertices from A . S is “bad” when $|B_A| = |B_{\bar{A}}| < \gamma k$. For each cardinality from 0 to γk , there are at most $\binom{k}{\gamma k}$ and $\binom{D - k}{\gamma k}$ ways to choose B_A and $B_{\bar{A}}$ respectively. Then we consider a random perfect matchings with all mini-vertices from $A \setminus B_A$ matched inside $A \setminus B_A$, and the same for $\bar{A} \setminus B_{\bar{A}}$ and $B_A \cup B_{\bar{A}}$.

$$\Pr_G[\exists \text{ bad } S] \leq \sum_{k=d_{\min}}^{D/2} \binom{D/d_{\min}}{k/d_{\min}} \gamma^k \binom{k}{\gamma k} \binom{D - k}{\gamma k} \frac{M(2\gamma k) M(k - \gamma k) M(D - k - \gamma k)}{M(D)}$$

We will use the approximation (2.10) for the number of perfect matchings on m vertices, where c_1 and c_2 are some positive constants:

$$c_1(m/e)^{m/2} < M(m) < c_2(m/e)^{m/2} \quad (2.54)$$

$$\begin{aligned} \gamma^k \frac{M(2\gamma k) M(k - \gamma k) M(D - k - \gamma k)}{M(D)} &< \\ &< c_4 \gamma^k \left(\frac{2\gamma k}{e}\right)^{\gamma k} \left(\frac{k - \gamma k}{e}\right)^{(k - \gamma k)/2} \left(\frac{D - k - \gamma k}{e}\right)^{(D - k - \gamma k)/2} \left(\frac{e}{D}\right)^{D/2} = \\ &= c_4 \gamma^k \frac{(2\gamma k)^{\gamma k} (k - \gamma k)^{(k - \gamma k)/2} (D - k - \gamma k)^{(D - k - \gamma k)/2}}{D^{D/2}} < \\ &< c_4 \gamma^k (2\gamma)^{\gamma k} \frac{k^{(k + \gamma k)/2} (D - k)^{(D - k - \gamma k)/2}}{D^{D/2}}. \end{aligned}$$

We apply Stirling’s approximation (2.9), for some constant $c_3 > 0$.

$$\begin{aligned} \binom{D/d_{\min}}{k/d_{\min}} &= \frac{(D/d_{\min})!}{(k/d_{\min})!((D - k)/d_{\min})!} < \\ &< \Theta(1) \left(\frac{D}{d_{\min}}\right)^{D/d_{\min} + 1/2} \left(\frac{d_{\min}}{k}\right)^{k/d_{\min} + 1/2} \left(\frac{d_{\min}}{D - k}\right)^{(D - k)/d_{\min} + 1/2} \leq \\ &\leq c_3 \left(\frac{D^D}{k^k (D - k)^{D - k}}\right)^{1/d_{\min}}. \end{aligned}$$

Finally, we use $\binom{n}{m} \leq (en/m)^m$.

$$\binom{k}{\gamma k} \binom{D - k}{\gamma k} \leq \left(\frac{e}{\gamma}\right)^{2\gamma k} \left(\frac{D - k}{k}\right)^{\gamma k}$$

Let’s combine all of the above.

$$\Pr_G[\exists \text{ bad } S] < \sum_{k=d_{\min}}^{D/2} c_5 \gamma k \frac{2^{\gamma k} e^{2\gamma k}}{\gamma^{\gamma k}} \left(\frac{k}{D}\right)^{k((1-\gamma)/2-1/d_{\min})}$$

Define $\alpha = \frac{2e^2}{\gamma}$, $\beta = \frac{1-\gamma}{2} - \frac{1}{d_{\min}}$, and

$$G(k) = c_5 \gamma k \alpha^{\gamma k} \left(\frac{k}{D}\right)^{\beta k} \quad (2.55)$$

We require $\beta > 0$, which means $d_{\min} > \frac{2}{1-\gamma}$, so $d_{\min} \geq 3$. Also $\gamma < 1 - 2/d_{\min}$.

$$\frac{dG(k)}{dk} = \left(\frac{1}{k} + \gamma \ln \alpha + \beta \ln \frac{k}{D} + \beta\right) G(k) \quad (2.56)$$

$$\frac{d^2G(k)}{dk^2} = \left(-\frac{1}{k^2} + \frac{\beta}{k} + \left(\frac{dG(k)}{dk}\right)^2\right) G(k) \quad (2.57)$$

$dG(k)/dk < 0$ for $k = 3d_{\min}$ when D is large.

$d^2G(k)/dk^2 > 0$ for $k \geq 3d_{\min}$ and $\gamma \leq 1 - 8/(3d_{\min})$.

Clearly, $\sum_{k=d_{\min}}^{3d_{\min}} G(k) = o(1)$.

To upper bound the sum for $3d_{\min} \leq k \leq D/2$, we can use $\max\{G(3d_{\min}), G(D/2)\}$.

$$G(3d_{\min}) = c_6 D^{-3d_{\min}\beta} = o\left(\frac{1}{D}\right) \quad (2.58)$$

$$G(D/2) = c_7 D \left(\frac{\alpha^\gamma}{2^\beta}\right)^{D/2} = o\left(\frac{1}{D}\right) \quad (2.59)$$

for constants c_6, c_7 , and

$$\begin{cases} 3d_{\min}\beta > 1 \\ \alpha^\gamma < 2^\beta \end{cases} \implies \begin{cases} \gamma \leq 1 - 8/(3d_{\min}) \\ \gamma\left(\frac{1}{2} + 2\log_2 e - \log_2 \gamma\right) < \frac{1}{2} - \frac{1}{d_{\min}} \end{cases} \quad (2.60)$$

If $d_{\min} = 3$, the last inequality holds whenever $\gamma \leq 0.0175$.

$$\Pr_G[\exists \text{ bad } S] < \sum_{k=d_{\min}}^{3d_{\min}} G(k) + \sum_{k=3d_{\min}}^{D/2} G(k) \leq o(1). \quad \square$$

2.7 Vertex Expanders in Locally Sparse Graphs and $G(n, p)$

For a graph $G = (V, E)$ on n vertices and a subset $W \subseteq V$ set $\text{vol}(W) = \sum_{v \in W} \deg(v)$.

We say that the set $S \subseteq V$ touches an edge, if it contains at least one of its endpoints, and S spans an edge, if it contains both endpoints.

Theorem 2.14 ([Kri18]). *Let $c_1 > c_2 > 0, 0 < \alpha < 1, \Delta > 0$, and $G = (V, E), \Delta(G) \leq \Delta, \frac{|E|}{|V|} \geq c_1$. Then there exist a constant $\gamma = \gamma(c_1, c_2, \alpha, \Delta) > 0$ and $\text{poly}(n)$ -time algorithm that finds either $W \subset V$ of size $|W| \leq \alpha n$ spanning at least $c_2|W|$ edges, or an induced $(|W|/2, \gamma)$ vertex expander on $|W| \geq \alpha n$ vertices.*

Proof outline. The idea is to begin with G and iteratively decrease the number of vertices, while maintaining the density. See Algorithm 1 for details. \square

Proposition 2.15 ([Kri18]). *Let $c_1 > c_2 > 1$ be reals. Define $\alpha = \left(\frac{c_2}{5c_1}\right)^{c_2/(c_2-1)}$. Let G be a random graph drawn from the probability distribution $G\left(n, \frac{c_1}{n}\right)$. Then w.h.p. every set of $k \leq \alpha n$ vertices of G spans fewer than $c_2 k$ edges.*

Proposition 2.16 ([Kri18]). *For every $c > 0$ and all sufficiently small $\delta > 0$ the following holds. Let G be a random graph drawn from the probability distribution $G\left(n, \frac{c}{n}\right)$. Then w.h.p. every set of $\frac{\delta}{\ln 1/\delta} n$ vertices of G touches fewer than δn edges.*

Theorem 2.17 (Linear size expanders in $G(n, p)$ graphs [Kri18]). $\forall \epsilon > 0 \exists \gamma > 0$: a random graph $G \sim G\left(n, \frac{1+\epsilon}{n}\right)$ contains w.h.p. an induced bounded degree $(n'/2, \gamma)$ vertex expander on $n' \geq \gamma n$ vertices.

Proof outline. A known fact about $G\left(n, \frac{1+\epsilon}{n}\right)$ is the existence of a giant connected component with “enough” edges. Removing a constant fraction of vertices having the highest degree, we get the bounded maximum degree by Proposition 2.16. Then Proposition 2.15 guarantees the local sparsity, and we apply Theorem 2.14. \square

Chapter 3

Random Power-Law Graphs

3.1 Models of Graphs

We will work with models of graphs that define either degree sequences or frequencies of degrees. In the latter case one can easily obtain the degree sequence as well, e.g., the frequencies $(x_1, x_2, x_3) = (2, 1, 4)$ would be transformed into the degrees $(1, 1, 2, 3, 3, 3, 3)$.

Given a degree sequence (w_1, \dots, w_n) , there are several ways to generate corresponding random graphs $G = (V, E)$ [Hop08]. Depending on the particular model, this will be either exact (fixed) or expected degree sequence.

All subsequent models possibly result in pseudographs — multigraphs with self-loops. Therefore, the degree sequences are only required to be pseudographic: degrees must be non-negative, and in case of exact degrees, they should add up to an even number [Hak62].

3.1.1 Coin Toss Model

For each pair of vertices i and j , we include the edge (i, j) in the graph with probability

$$p_{ij} = \Pr_G[(i, j) \in E] = \begin{cases} \frac{w_i w_j}{\sum_{k \in V} w_k} & \text{if } i \neq j, \\ \frac{w_i^2}{2 \sum_{k \in V} w_k} & \text{otherwise.} \end{cases} \quad (3.1)$$

Given expression is a valid probability, if we have

$$\max_k w_k^2 < \sum_{k \in V} w_k \quad (3.2)$$

This assumption also implies that the sequence is graphic, i.e., realizable by some graph.

The standard convention is that self-loop edges contribute 2 to the degrees of the nodes. Considering this, we show that the model is well-defined:

$$\mathbb{E}_G[\deg(i)] = \sum_{j \neq i} p_{ij} + 2p_{ii} = \frac{\sum_{j \neq i} w_i w_j + w_i^2}{\sum_{j \in V} w_j} = w_i \quad (3.3)$$

Alternatively, one can generate a graph by including some number of edges between i and j according to a Poisson process with mean $\lambda = \frac{w_i w_j}{\sum_{k \in V} w_k}$. High degree vertices would then form a clique.

In this model all edges are independent, and (w_1, \dots, w_n) are the expected degrees.

3.1.2 m Edges Model

We generate $m = \frac{1}{2} \sum_{k \in V} w_k$ edges uniformly at random by successively selecting pairs of vertices with probability proportional to their degrees. In this case w_i 's are also expected degrees, however, we get the exact number of edges, which are no longer independent.

3.1.3 Permutation Model

We create a sequence with w_i mini-vertices for each vertex i , randomly permute it, and connect each consecutive non-intersecting pair of mini-vertices with an edge. Equivalently, one can choose a uniformly random perfect matching on this sequence. Then we put an edge between vertices i and j if at least one pair of mini-vertices corresponding to both of them is adjacent.

This model guarantees the exact number of edges and they are not independent, but the degree sequence is now exact.

3.1.4 Full Graph Selection Model

We randomly pick a graph G from a family of all graphs with given exact degree sequence, but it is usually not practical.

In this chapter we will examine the properties of power-law graphs under different models.

3.2 Coin Toss Power-Law Model

We are given the expected degree sequence (w_1, \dots, w_n) , where $w_v = \mathbb{E}[\deg(v)]$.

$$\Pr_G[(u, v) \in E] = \frac{w_u w_v}{\text{Vol}(G)}, \quad \text{assuming } \max_v w_v^2 < \text{Vol}(G) \quad (3.4)$$

$$\text{Vol}(G) = \mathbb{E}_G[\text{vol}(G)] = \mathbb{E}_G \left[\sum_{v \in V} \deg(v) \right] = \sum_{v \in V} w_v \quad (3.5)$$

Let the expected number of vertices with degree x be equal to y , which follows the power-law distribution for some $\alpha = \alpha(n) > 0$ and constant $\beta > 0$. Note that α and β are the only open parameters in this model.

$$\mathbb{E}_G[|\{v \in V \mid \deg(v) = x\}|] = y = \frac{e^\alpha}{x^\beta}, \quad \text{where } x \in \{1, \dots, d_{max}\}. \quad (3.6)$$

Table 3.1: Properties of power-law graphs [ACL01], $\beta_0 \approx 3.4785$

β	(0; 1)	1	(1; 2)	2	(2; β_0)	(β_0 ; ∞)
the largest component	n	the giant component, $\Theta(n)$				$O(n^{2/\beta} \log n)$
the 2nd largest component	—	$O(1)$	$O\left(\frac{\log n}{\log \log n}\right)$	$O(\log n)$		
$ V = n$	$\frac{e^{\alpha/\beta}}{1-\beta}$	αe^α	$\zeta(\beta)e^\alpha$			
$ E $	$\frac{1}{2} \frac{e^{2\alpha/\beta}}{2-\beta}$		$\frac{1}{4} \alpha e^\alpha$	$\frac{1}{2} \zeta(\beta-1)e^\alpha$		

Our model is asymptotically identical to Aiello et al. [ACL01], whose properties are summarized in Table 3.1. The only distinction is in defining the expected frequencies instead of the exact ones.

3.2.1 Maximum Degree, Expected Size, and the Number of Edges

Like Aiello et al. [ACL01], we ignore isolated vertices, $y \geq 1$ and $0 \leq \ln y = \alpha - \beta \ln x$, consequently, we may deduce

$$d_{max} = e^{\alpha/\beta} \quad (3.7)$$

The size of the graph and the expected number of edges are

$$n = \sum_{x=1}^{e^{\alpha/\beta}} \frac{e^\alpha}{x^\beta} \quad \mathbb{E}_G[|E|] = \frac{\text{Vol}(G)}{2} = \frac{1}{2} \sum_{x=1}^{e^{\alpha/\beta}} x \frac{e^\alpha}{x^\beta} \quad (3.8)$$

To get some intuition about the density of G , we consider the special case when $d_{max} = n$.

$$n = \sum_{x=1}^n \frac{e^\alpha}{x^\beta} = e^\alpha H_{n,\beta} \quad e^\alpha = \frac{n}{H_{n,\beta}} \quad (3.9)$$

$$\mathbb{E}_G[m] = \frac{1}{2} \mathbb{E}_G\left[\sum_{v \in V} \deg(v)\right] = \frac{1}{2} \sum_{x=1}^n x \frac{e^\alpha}{x^\beta} = \frac{e^\alpha}{2} \sum_{x=1}^n x^{1-\beta} = \frac{e^\alpha}{2} H_{n,\beta-1} = \frac{n}{2} \frac{H_{n,\beta-1}}{H_{n,\beta}} \quad (3.10)$$

$$H_{n,-1} = \frac{n(n+1)}{2}, \quad H_{n,0} = n, \quad H_{n,1} = \log n + O(1), \quad H_{n,\beta} = O(1), \text{ if } \beta > 1.$$

$$\mathbb{E}_G[m] = \begin{cases} \Theta(n^2) & \text{if } \beta = 0, \\ \Theta(n^2 / \log n) & \text{if } \beta = 1, \\ \Theta(n \log n) & \text{if } \beta = 2, \\ \Theta(n) & \text{if } \beta > 2. \end{cases} \quad (3.11)$$

3.2.2 Expected Volume and Average Degree

When $n \rightarrow \infty$, the expected volume $\text{Vol}(G)$ is well-defined only for $\beta > 2$.

$$\text{Vol}(G) \approx \int_1^\infty x \frac{e^\alpha}{x^\beta} dx = \frac{e^\alpha}{2-\beta} \left[x^{-\beta+2} \right]_1^\infty = \frac{e^\alpha}{\beta-2} \quad (3.12)$$

On the other hand, considering $d_{max} = e^{\alpha/\beta}$ we get

$$\text{Vol}(G) \approx \frac{e^\alpha}{2-\beta} \left(\left(e^{\alpha/\beta} \right)^{-\beta+2} - 1 \right) = \frac{1}{2-\beta} \left(e^{2\alpha/\beta} - e^\alpha \right) \quad (3.13)$$

The expected average degree is

$$d = \frac{\text{Vol}(G)}{n} \quad (3.14)$$

The second-order average degree is

$$\tilde{d} = \frac{\sum_{v \in V} w_v^2}{\text{Vol}(G)} \approx \begin{cases} d \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} & \text{if } \beta > 3 \\ \frac{1}{2} d \ln \frac{2d_{max}}{d} & \text{if } \beta = 3 \\ d^{\beta-2} \frac{(\beta-2)^{\beta-1} d_{max}^{\beta-\beta}}{(\beta-1)^{\beta-2} (3-\beta)} & \text{if } 2 < \beta < 3 \end{cases} \quad (3.15)$$

3.3 Permutation Power-Law Model

In the permutation model each vertex $v \in \{1, \dots, n\}$ has exactly w_v neighbors chosen with probability proportional to their degrees,

$$\deg(v) = w_v = \frac{pn}{v^\beta}, \quad 0 < p \leq 1. \quad (3.16)$$

If $\beta = 0$, the coin toss model with expected degree sequence (w_1, \dots, w_n) would be identical to $G(n, p)$ with self-loops:

$$w_v = pn \quad \text{vol}(G) = pn^2 \quad |E| = \frac{pn^2}{2}$$

$$\Pr_G[(u, v) \in E] = \frac{w_u w_v}{\sum_i w_i} = \frac{(pn)^2}{pn^2} = p$$

Note that the degrees w_i should be integers so we have to round down (3.16). The issue of emerging error is addressed in Section 4.2.1.

3.4 k-CNF Power-Law Model

Generation of CNF f having m clauses with k literals each is done by adding variable $i \in \{1, \dots, n\}$ with a random sign to the current clause $C \in \{C_1, \dots, C_m\}$ at a position t , $1 \leq t \leq k$, with probability $p_i = \frac{w_i}{\sum_{j=1}^n w_j}$, where $w_i = i^{-\beta}$ [ABL09b].

If any resulting clause is a tautology (it contains both some variable and its negation) or simplifiable (there is the same literal at multiple positions), then it is discarded and regenerated from scratch. Essentially it means that we are choosing variables for each clause without replacement.

The *variable incidence graph (VIG)* $G(f) = (V, E)$ is defined on $V = \{1, \dots, n\}$, and edge (i, j) is added to E whenever both i and j appear in the same clause. The signs of the literals are ignored in VIG, so we will ignore them as well, writing “ $i \in C$ ” instead of “ $i \in C \vee -i \in C$ ”.

$$\Pr_{G(f)} [(i, j) \in E] = 1 - \prod_C (1 - \Pr[i \in C \wedge j \in C]), \text{ for } i \neq j \quad (3.17)$$

For simplicity, let's assume $k = 3$.

$$\Pr[i \in C \wedge j \in C] = \sum_{\substack{1 \leq x \leq n \\ x \notin \{i, j\}}} (\Pr[C = (i, j, x)] + \Pr[C = (j, i, x)] + \dots + \Pr[C = (x, j, i)])$$

$$p_x \leq 1 - p_i - p_j$$

$$\Pr[C = (i, j, x)] = p_i \frac{p_j}{1 - p_i} \frac{p_x}{1 - p_i - p_j}$$

$$\Pr[C = (i, x, j)] = p_i \frac{p_x}{1 - p_i} \frac{p_j}{1 - p_i - p_x}, \text{ and so on.}$$

3.4.1 Noncentral Hypergeometric Distribution

Sampling of elements from different classes with weights without replacement is captured by *Wallenius' noncentral hypergeometric distribution*.

We have $\sum_{i=1}^c m_i$ objects total from c classes, m_i is the number of objects of class i . n objects are sampled without replacement. The probability that an object from class i is sampled is proportional to w_i .

If at step t we have sampled $(x_{1,t}, \dots, x_{c,t})$ objects, i.e., exactly $x_{i,t}$ objects from class i , then the probability that the next draw gives an object from class i is

$$\frac{(m_i - x_{i,t})w_i}{\sum_{j=1}^c (m_j - x_{j,t})w_j} \quad (3.18)$$

The multivariate case:

$$\text{mwnchypg}(\mathbf{x}; n, \mathbf{m}, \mathbf{w}) = \left(\prod_{i=1}^c \binom{m_i}{x_i} \right) \left(\int_0^1 \prod_{i=1}^c (1 - t^{w_i/d})^{x_i} dt \right) \quad (3.19)$$

$$d = \mathbf{w}(\mathbf{m} - \mathbf{x}) = \sum_{i=1}^c w_i(m_i - x_i) \quad (3.20)$$

The probability function can be calculated using a variety of methods: recursive calculation, binomial expansion methods, Taylor expansion methods, continued fraction expansion, and numerical integration [Fog08]. At the same time direct formal analysis does not look promising. Instead, we can use the fact that the sequence of expected degrees in VIG $G(f)$ also follows power-law distribution.

Theorem 3.1 (Theorem 1 [ABL09b]). *In the k -CNF power-law model with continuous probability distribution $\phi(x, \beta) = (1 - \beta)x^{-\beta}$, when $n \rightarrow \infty$,*

$$\Pr_f[\text{variable } i \text{ has } k \text{ occurrences}] \propto k^{-\alpha}, \quad \text{where } \alpha = 1/\beta + 1. \quad (3.21)$$

Let x_i be the number of occurrences/clauses with variable i .

$$\Pr[x_i = k] = ck^{-\alpha} \quad (3.22)$$

Events “occurrence of i in clause C ” are independent for all clauses, so the actual number of occurrences will be close to its expectation.

$$\mathbb{E}[x_i] = \sum_{k=1}^m k \Pr[x_i = k] = \sum_{k=1}^m ck^{1-\alpha} = cH_{m, \alpha-1}$$

For any pair (i, j) , if $x_i + x_j \geq m$, then $\Pr[(i, j) \in E] = 1$ by the pigeonhole principle.

Otherwise $x_i + x_j < m$ and

$$\begin{aligned} \Pr_{G(f)}[(i, j) \in E] &= \Pr[\exists C : i \in C \wedge j \in C] = 1 - \frac{\binom{m-x_i}{x_j}}{\binom{m}{x_j}} = \\ &= 1 - \frac{(m-x_i)!}{(m-x_i-x_j)!} \frac{(m-x_j)!}{m!} = \\ &= 1 - \prod_{k=0}^{x_j-1} \frac{m-x_i-k}{m-k}. \end{aligned}$$

Table 3.2: Properties of the “octopus” graphs [CL04], β' is the actual exponent

β	(2; 3)	3	(3; ∞)
$\beta' = 1/(\beta - 1)$	(1/2; 1)	1/2	(0; 1/2)
average distance	$O(\log \log n)$	$\Theta(\log n / \log \log n)$	$\Theta(\log n / \log \tilde{d})$
diameter	$\Theta(\log n)$		

3.5 “Octopus” Power-Law Graph Model

The “octopus” graph model was defined by Chung and Lu [CL04]:

$$\mathbb{E}_G[\deg(i)] = w_i = ci^{-1/(\beta-1)} \quad (3.23)$$

$$i_0 \leq i < n + i_0 \quad (3.24)$$

$$c = \frac{\beta - 2}{\beta - 1} dn^{\frac{1}{\beta-1}} \quad (3.25)$$

$$i_0 = n \left(\frac{d(\beta - 2)}{d_{max}(\beta - 1)} \right)^{\beta-1} \quad (3.26)$$

The known results about the average distance and the diameter of resulting graphs are collected in Table 3.2. The generation process is the same as in Section 3.1.1.

We consider the same specific family of such graphs, as Chung and Lu [CL04]:

$$2 < \beta < 3 \implies \frac{1}{2} \leq \frac{1}{\beta - 1} \leq 1 \quad (3.27)$$

$$d > 1 \quad (3.28)$$

$$d_{max} \gg d \quad (3.29)$$

$$\log d_{max} \gg \log n / \log \log n \quad (3.30)$$

The graphs are generated like in the coin toss model and (3.2) is assumed to be true.

Note that d_{max} and d are simply parameters of this model.

The maximum degree $w_{i_0} = \frac{\beta - 2}{\beta - 1} dn^{\frac{1}{\beta-1}} i_0^{-1/(\beta-1)} = d_{max}$.

And the average degree is $\frac{1}{n} \sum_{k=i_0}^{n+i_0} w_i \approx \frac{1}{n} \sum_{k=1}^n w_i = \frac{1}{n} \sum_{k=1}^n \frac{\beta - 2}{\beta - 1} dn^{\frac{1}{\beta-1}} k^{-1/(\beta-1)} =$

$$\begin{aligned} &= \frac{\beta - 2}{\beta - 1} \frac{d}{n^{(\beta-2)/(\beta-1)}} \sum_{k=1}^n k^{-1/(\beta-1)} \approx \\ &\approx \frac{\beta - 2}{\beta - 1} \frac{d}{n^{(\beta-2)/(\beta-1)}} \frac{n^{1-1/(\beta-1)}}{1 - 1/(\beta - 1)} = d. \end{aligned}$$

3.5.1 The Second-Order Average Degree of “Octopus” Graphs

Chung and Lu [CL04] only state the final results, so we provide the derivations here for completeness.

$$\begin{aligned}\tilde{d} &= \frac{1}{dn} \sum_{k=1}^n w_i^2 = \frac{1}{dn} \sum_{k=1}^n \left(\frac{\beta-2}{\beta-1} \right)^2 d^2 n^{\frac{2}{\beta-1}} k^{-2/(\beta-1)} = \\ &= \left(\frac{\beta-2}{\beta-1} \right)^2 \frac{d}{n^{1-2/(\beta-1)}} \sum_{k=1}^n k^{-2/(\beta-1)}\end{aligned}$$

Case $\beta > 3$:

$$\tilde{d} \approx \left(\frac{\beta-2}{\beta-1} \right)^2 \frac{d}{n^{1-2/(\beta-1)}} \frac{n^{(\beta-3)/(\beta-1)}}{(\beta-3)/(\beta-1)} = \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d$$

Case $\beta = 3$:

$$\begin{aligned}i_0 &= n \left(\frac{d}{2d_{max}} \right)^2 \\ \tilde{d} &\approx \frac{d}{4} \sum_{k=i_0}^n k^{-1} \approx \frac{d}{4} (\ln n - \ln i_0) = \frac{d}{4} (\ln n - \ln n - 2 \ln \frac{d}{2d_{max}}) = \frac{d}{2} \ln \frac{2d_{max}}{d}\end{aligned}$$

Case $2 < \beta < 3$:

$$\begin{aligned}\tilde{d} &\approx \left(\frac{\beta-2}{\beta-1} \right)^2 \frac{d}{n^{(\beta-3)/(\beta-1)}} \sum_{k=i_0}^n k^{-2/(\beta-1)} \approx \\ &\approx \left(\frac{\beta-2}{\beta-1} \right)^2 \frac{d}{n^{(\beta-3)/(\beta-1)}} \left(\frac{n^{(\beta-3)/(\beta-1)}}{(\beta-3)/(\beta-1)} - \frac{i_0^{(\beta-3)/(\beta-1)}}{(\beta-3)/(\beta-1)} \right) \approx \\ &\approx - \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} d \left(\frac{d(\beta-2)}{d_{max}(\beta-1)} \right)^{\beta-3} = \frac{(\beta-2)^{\beta-1}}{(\beta-1)^{\beta-2}(3-\beta)} d^{\beta-2} d_{max}^{3-\beta}.\end{aligned}$$

3.5.2 Diameter of “Octopus” Graphs

Theorem 3.2 (Theorem 2.6 [CL04]). *Suppose a power-law random graph with exponent β has average degree $d > 1$ and maximum degree $d_{max} \gg n^{1/\log \log n}$. If $2 < \beta < 3$, its diameter is $\Theta(\log n)$ w.h.p.*

Proof outline. The core of the “octopus” graph is defined to contain all vertices of degree at least $n^{1/\log \log n}$. Combining the following claims gives $O(\log n)$ diameter of the graph w.h.p.

Claim 1: The diameter of the core is $O(\log \log n)$ w.h.p.

Claim 2: Almost all vertices with degree at least $\log n$ are within distance $O(\log \log n)$ from the core.

Claim 3: Each vertex in the giant connected component is within distance $O(\log n)$ from a vertex of degree at least $O(\log n)$. \square

Such a specific structure explains the name of these graphs.

Chapter 4

Expanders in Power-Law Graphs

From now on, we will work with an induced subgraph $H = (V_H, E_H)$ of G , obtained by retaining only vertices of degree at least d_0 . This definition is both intuitive and easy to work with. We denote the size $|V_H| = n'$.

If H is an expander, then outputting H is as simple as checking the degree of each vertex during the process of generation of the graph G . Connecting large spectral gap with high combinatorial expansion would allow one to eliminate error of outputting non-expander, but it is still an open problem for our models.

4.1 Coin Toss Model

Theorem 4.1 (Existence of an edge expander). $\exists d_0 \forall c < 1/3 \forall \delta > 0 \exists \gamma, c_1 = c_1(\beta, c) > 0$: let $G = (V, E)$ be a random power-law graph and $H = (V_H, E_H)$ be its induced subgraph on $|V_H| = n'$ vertices, such that $V_H = \{v \in V \mid \deg(v) \geq d_0\}$.

Then if $0 < \beta < 1$, the whole graph G is $(n/2, \gamma)$ edge expander w.h.p.

Otherwise, the subgraph H is $(n'/2, \gamma)$ edge expander w.h.p.

$$d_{max} = \max_{v \in V} \deg(v) \tag{4.1}$$

$$d_0 = \begin{cases} 1 & \text{if } 0 < \beta < 1, \\ d_{max}/\sqrt{n} & \text{if } \beta = 1, \\ \frac{d_{max}}{n^{1/\beta}} & \text{if } 1 < \beta \leq 1.6, \\ c d_{max} & \text{if } \beta > 1.6. \end{cases} \tag{4.2}$$

$$n' = \begin{cases} n & \text{if } 0 < \beta < 1, \\ n/2 & \text{if } \beta = 1, \\ \frac{n}{(\beta-1)\zeta(\beta)^{1/\beta}} & \text{if } 1 < \beta \leq 1.6, \\ \frac{c^{-\beta+1}-1}{\beta-1} (n/\zeta(\beta))^{1/\beta} = \Theta(n^{1/\beta}) & \text{if } \beta > 1.6. \end{cases} \quad (4.3)$$

$$c_1 = \begin{cases} \frac{(1-\beta)^2}{2-\beta} & \text{if } 0 < \beta < 1, \\ \frac{1}{1-c} & \text{if } \beta = 1, \\ \frac{(\ln 1/c)^2}{(\beta-1)^2} \frac{1-c^{2-\beta}}{(c^{-\beta+1}-1)^2} & \text{if } 1 < \beta < 2 \text{ or } \beta > 2, \\ \frac{\ln 1/c}{(1/c-1)^2} & \text{if } \beta = 2. \end{cases} \quad (4.4)$$

In all cases the average degree d and the expansion γ are as follows:

$$d = c_1 n' \quad (4.5)$$

$$\gamma = \frac{d}{2} - \delta \quad (4.6)$$

The proof of the theorem is on page 36.

Lemma 4.2. $\forall \delta > 0 \exists \gamma$: If the expected average degree of a graph H on n' vertices is $d > 10 \ln n'$, then H is $(n'/2, \gamma)$ edge expander w.h.p., $\gamma = d/2 - \delta$.

Proof. Given an arbitrary cut (S, T) : $e(S, T) = \{(u, v) \in E_H \mid u \in S, v \in T\}$. Using (3.4),

$$\frac{\mathbb{E}[|e(S, T)|]}{G} = \frac{\text{Vol}(S) \text{Vol}(T)}{\text{Vol}(H)} \quad (4.7)$$

The volume of an arbitrary subset $S \subset V_H$ of a fixed size $|S| = s \leq n'/2$ is $\text{Vol}(S) = sd$, where $d = \text{Vol}(H)/n'$ is the expected average degree of H .

Define a random variable X_S to be the size of a cut $(S, V_H \setminus S)$ with expected value μ .

$$X_S = |e(S, V_H \setminus S)| \quad (4.8)$$

$$\mu = \frac{\mathbb{E}[X_S]}{G} = \frac{\text{Vol}(S) \text{Vol}(V_H \setminus S)}{\text{Vol}(H)} = \text{Vol}(S) \left(1 - \frac{s}{n'}\right) \quad (4.9)$$

$$\frac{\mu}{s} = d \left(1 - \frac{s}{n'}\right) \quad (4.10)$$

$$\frac{d}{2} \leq \frac{\mu}{s} \leq d \quad (4.11)$$

We will use the Chernoff bound for the lower tail (2.4) for some $0 < \lambda < 1$.

$$\begin{aligned}
\gamma s &= (1 - \lambda)\mu \\
0 < \lambda &= 1 - \gamma s/\mu \\
\gamma &< \mu/s
\end{aligned} \tag{4.12}$$

Therefore, the only requirement is $\gamma < d/2$. Now, as for the expansion of H ,

$$\begin{aligned}
\Pr_{G,S}[S \text{ is non-expanding}] &= \Pr_{G,S}[X_S \leq \gamma s] \leq \exp(-\lambda^2 \mu/2) = \exp(-(\mu - \gamma s)^2/2\mu) = \\
&= \exp(-\mu/2 + \gamma s - \gamma^2 s^2/2\mu) \leq \exp(\gamma s - \mu/2).
\end{aligned}$$

$$\begin{aligned}
\Pr_G[H \text{ is not } (n'/2, \gamma) \text{ edge expander}] &\leq \sum_{s=1}^{n'/2} \binom{n'}{s} \Pr_{G,S}[S \text{ is non-expanding}] \leq \\
&\leq \sum_{s=1}^{n'/2} \left(\frac{e n'}{s} \exp\left(-\frac{(\mu/s - \gamma)^2}{2\mu/s}\right) \right)^s = \sum_{s=1}^{n'/2} \exp\left(\left(1 + \ln \frac{n'}{s} - \frac{(\mu/s - \gamma)^2}{2\mu/s}\right) s\right) \leq \\
&\leq \sum_{s=1}^{n'/2} \exp\left(\left(1 - \ln \frac{s}{n'} - \frac{d}{2}(1 - s/n') + \gamma - \frac{\gamma^2}{2d(1 - s/n')}\right) s\right) \leq \\
&\leq \sum_{s=1}^{n'/2} \exp\left(\left(1 + \ln n' - \frac{d}{4} + \gamma\right) s\right) \leq \\
&\leq \sum_{s=1}^{n'/2} \left(\frac{1}{n'^{(1+\sigma)}}\right)^s \leq \frac{n'/2}{n'^{(1+\sigma)}} = o(1).
\end{aligned}$$

The last line is obtained if we satisfy this inequality for some small constant $\sigma > 0$:

$$\begin{aligned}
1 + \ln n' - \frac{d}{4} + \gamma &\leq -(1 + \sigma) \ln n' \\
0 < \gamma &\leq \frac{d}{4} - (2 + \sigma) \ln n' - 1 \\
d &> 4((2 + \sigma) \ln n' + 1) \\
d &> 10 \ln n'
\end{aligned}$$

□

Lemma 4.3 (Size and volume of the subgraph H). *If $d_0 = c d_{max}$ for $0 < c < 1$, then*

$$n' = \begin{cases} (1 - c^{1-\beta}) n & \text{if } \beta < 1, \\ \frac{\ln 1/c}{c^{-\beta+1} - 1} n \approx (\ln 1/c) \frac{n}{\ln n} & \text{if } \beta = 1, \\ \frac{c^{-\beta+1} - 1}{\beta - 1} (n/\zeta(\beta))^{1/\beta} & \text{if } \beta > 1. \end{cases} \quad (4.13)$$

$$\text{Vol}(H) = \begin{cases} \frac{1 - c^{2-\beta}}{2 - \beta} e^{2\alpha/\beta} & \text{if } \beta < 2, \\ (\ln 1/c) e^\alpha & \text{if } \beta = 2, \\ \frac{c^{-\beta+2} - 1}{\beta - 2} e^{2\alpha/\beta} + \frac{c^{-\beta+1} - 1}{2} e^{\alpha/\beta} & \text{if } \beta > 2. \end{cases} \quad (4.14)$$

$$d = \frac{\text{Vol}(H)}{n'} = \Theta(n') \quad (4.15)$$

Proof. Consider (3.8) and $n' = \sum_{x=d_0}^{d_{max}} \frac{e^\alpha}{x^\beta}$, $\text{Vol}(H) = \sum_{x=d_0}^{d_{max}} x \frac{e^\alpha}{x^\beta}$.

Case $\beta < 1$: similarly to (3.13),

$$n = \frac{e^{\alpha/\beta}}{1 - \beta} \quad n' = \frac{e^\alpha}{1 - \beta} (d_{max}^{1-\beta} - d_0^{1-\beta}) = \frac{1 - c^{1-\beta}}{1 - \beta} e^{\alpha/\beta}$$

Case $\beta = 1$: $d_{max} = e^\alpha$.

$$n = (\ln d_{max}) e^\alpha = \alpha e^\alpha \quad n' = (\ln d_{max} - \ln d_0) e^\alpha = (\ln 1/c) e^\alpha$$

Case $\beta > 1$:

$$n = e^\alpha H_{d_{max}, \beta}$$

$$\begin{aligned} n' &= e^\alpha (H_{d_{max}, \beta} - H_{c d_{max}, \beta}) = \left(1 - \frac{H_{c d_{max}, \beta}}{H_{d_{max}, \beta}} \right) n = \\ &= \left(1 - \frac{\zeta(\beta) - \frac{1}{(\beta-1)(c d_{max})^{\beta-1}} - \frac{1}{2(c d_{max})^\beta} - O\left(\frac{1}{(c d_{max})^{\beta+1}}\right)}{\zeta(\beta) - \frac{1}{(\beta-1)d_{max}^{\beta-1}} - \frac{1}{2d_{max}^\beta} - O\left(\frac{1}{d_{max}^{\beta+1}}\right)} \right) n = \\ &= \left(\frac{\frac{c^{-\beta+1}-1}{(\beta-1)d_{max}^{\beta-1}} + \frac{c^{-\beta}-1}{2d_{max}^\beta} + O\left(\frac{1}{d_{max}^{\beta+1}}\right)}{\zeta(\beta) - \frac{1}{(\beta-1)d_{max}^{\beta-1}} - \frac{1}{2d_{max}^\beta} - O\left(\frac{1}{d_{max}^{\beta+1}}\right)} \right) n = \\ &= \left(\frac{c^{-\beta+1}-1}{(\beta-1)d_{max}^{\beta-1}} + \frac{c^{-\beta}-1}{2d_{max}^\beta} + O\left(\frac{1}{d_{max}^{\beta+1}}\right) \right) e^\alpha = \\ &= \frac{c^{-\beta+1}-1}{\beta-1} d_{max} + \frac{c^{-\beta}-1}{2} + O\left(\frac{1}{d_{max}}\right) \approx \\ &\approx \frac{c^{-\beta+1}-1}{\beta-1} e^{\alpha/\beta} \approx \frac{c^{-\beta+1}-1}{\beta-1} \left(\frac{n}{\zeta(\beta)} \right)^{1/\beta}. \end{aligned}$$

Now for the expected volume, **case** $\beta < 2$:

$$\text{Vol}(H) = \frac{e^\alpha}{2-\beta} \left(d_{max}^{2-\beta} - d_0^{2-\beta} \right) = \frac{1-c^{2-\beta}}{2-\beta} e^{2\alpha/\beta}$$

Case $\beta = 2$: $d_{max} = e^{\alpha/2}$.

$$\text{Vol}(H) = (\ln d_{max} - \ln d_0) e^\alpha = (\ln 1/c) e^\alpha$$

Case $\beta > 2$:

$$\begin{aligned} \text{Vol}(H) &= e^\alpha (H_{d_{max}, \beta-1} - H_{c d_{max}, \beta-1}) = \\ &= \left(-\frac{1}{(\beta-2)d_{max}^{\beta-2}} - \frac{1}{2d_{max}^{\beta-1}} + \frac{1}{(\beta-2)(c d_{max})^{\beta-2}} + \frac{1}{2(c d_{max})^{\beta-1}} + O\left(\frac{1}{d_{max}^\beta}\right) \right) e^\alpha = \\ &= \left(\frac{c^{-\beta+2} - 1}{(\beta-2)d_{max}^{\beta-2}} + \frac{c^{-\beta+1} - 1}{2d_{max}^{\beta-1}} + O\left(\frac{1}{d_{max}^\beta}\right) \right) e^\alpha = \\ &= \frac{c^{-\beta+2} - 1}{\beta-2} d_{max}^2 + \frac{c^{-\beta+1} - 1}{2} d_{max} + O(1) \approx \\ &\approx \frac{c^{-\beta+2} - 1}{\beta-2} e^{2\alpha/\beta} + \frac{c^{-\beta+1} - 1}{2} e^{\alpha/\beta} \end{aligned}$$

Finally, the average degree is always $\Theta(n')$.

$$d = \frac{\text{Vol}(H)}{n'} = \begin{cases} \frac{1-\beta}{2-\beta} \frac{1-c^{2-\beta}}{1-c^{1-\beta}} e^{\alpha/\beta} & \text{if } \beta < 1, \\ \frac{1-c}{\ln 1/c} e^\alpha & \text{if } \beta = 1, \\ \frac{\beta-1}{2-\beta} \frac{1-c^{2-\beta}}{c^{-\beta+1}-1} e^{\alpha/\beta} & \text{if } 1 < \beta < 2, \\ \frac{\ln 1/c}{1/c-1} e^{\alpha/2} & \text{if } \beta = 2, \\ \frac{\beta-1}{\beta-2} \frac{c^{-\beta+2}-1}{c^{-\beta+1}-1} e^{\alpha/\beta} + O(1) & \text{if } \beta > 2. \end{cases}$$

□

Considering H independently from $G \setminus H$, we need to account for $e(H, G \setminus H)$.

$$\text{Vol}(G) = \begin{cases} \Theta(n^2) & \text{if } \beta < 1, \\ \Theta\left(\left(\frac{n}{\log n}\right)^2\right) & \text{if } \beta = 1, \\ \Theta(n^{2/\beta}) & \text{if } 1 < \beta < 2, \\ \Theta(n \log n) & \text{if } \beta = 2, \\ \Theta(n) & \text{if } \beta > 2. \end{cases} \quad (4.16)$$

$$\text{Vol}(H) = \sum_{v \in H} w_v = \Theta(n'^2) = \begin{cases} \Theta(n^2) & \text{if } \beta < 1, \\ \Theta\left(\left(\frac{n}{\log n}\right)^2\right) & \text{if } \beta = 1, \\ \Theta(n^{2/\beta}) & \text{if } \beta > 1. \end{cases} \quad (4.17)$$

$$2|e(H)| = \text{Vol}(H) - |e(H, G \setminus H)| \quad (4.18)$$

$$|e(H, G \setminus H)| = \frac{\text{Vol}(H) \text{Vol}(G \setminus H)}{\text{Vol}(G)} \quad (4.19)$$

$$|e(H)| = \frac{\text{Vol}(H)^2}{2 \text{Vol}(G)} \quad (4.20)$$

$$\begin{aligned} |e(G)| &= |e(H)| + |e(H, G \setminus H)| + |e(G \setminus H)| = \\ &= \frac{\text{Vol}(H)^2/2 + \text{Vol}(H) \text{Vol}(G \setminus H) + \text{Vol}(G \setminus H)^2/2}{\text{Vol}(G)} = \\ &= \frac{(\text{Vol}(H) + \text{Vol}(G \setminus H))^2}{2 \text{Vol}(G)} = \frac{\text{Vol}(G)^2}{2 \text{Vol}(G)} = \frac{\text{Vol}(G)}{2} \end{aligned} \quad (4.21)$$

How much smaller would $2e(H)$ be than $\text{Vol}(H)$ that we use?

$$2e(H) = \frac{\text{Vol}(H)^2}{\text{Vol}(G)} = \text{Vol}(H) x \quad (4.22)$$

$$x = \frac{\text{Vol}(H)}{\text{Vol}(G)} = \begin{cases} \Theta(1) & \text{if } \beta < 2, \\ \frac{1}{\log n} & \text{if } \beta = 2, \\ \frac{1}{n^{1-2/\beta}} & \text{if } \beta > 2. \end{cases} \quad (4.23)$$

It means the larger β , the fewer edges are between high-degree vertices in H than edges from H to low-degree vertices in $G \setminus H$.

In order to get larger subgraph H for $\beta \geq 1$, we need to include more vertices in H .

Lemma 4.4 (Linear size subgraph H). *When $\beta = 1$, leaving the vertices of degree at least $d_0 = d_{\max}/\sqrt{n}$ results in the subgraph H of size $n' = n/2$. And when $1 < \beta \leq 1.6$, we pick $d_0 = \frac{d_{\max}}{n^{1/\beta}}$ to get H of size $n' = \Theta(n)$.*

Proof. **Case $\beta = 1$:** $d_{max} = e^\alpha$. In this case $n = \alpha e^\alpha$ and $\alpha \approx \ln n$ by (A.4).

$$\begin{aligned} d_0 &= d_{max} e^{-x} \\ \ln d_0 &= \ln d_{max} - x \\ n' &= (\ln d_{max} - \ln d_0) \frac{n}{\alpha} = \frac{x n}{\alpha} \end{aligned}$$

$$\begin{aligned} x = \frac{\ln n}{2} : \quad d_0 &= \frac{d_{max}}{\sqrt{n}}, \quad n' \approx \frac{n}{2} \\ x = c \ln n : \quad d_0 &= \frac{d_{max}}{n^c}, \quad n' \approx cn \end{aligned}$$

$$\begin{aligned} \text{Vol}(H) &= e^\alpha (d_{max} - d_{max} n^{-c}) = e^{2\alpha} (1 - n^{-c}) \\ d = \frac{\text{Vol}(H)}{n'} &= \frac{e^\alpha}{c\alpha} (1 - n^{-c}) = \frac{n'}{c^2 \alpha^2} (1 - (n'/c)^{-c}) \approx \frac{1}{c^2} (1 - (n'/c)^{-c}) \frac{n'}{(\ln(n'/c))^2} \end{aligned}$$

$$\text{If } c = 1/2, \quad d \approx 4 \left(1 - \frac{1}{\sqrt{2n'}}\right) \frac{n'}{(\ln(2n'))^2} = \omega(\ln n').$$

Case $\beta > 1$:

$$\begin{aligned} n' &= \left(1 - \frac{H_{x,\beta}}{H_{d_{max},\beta}}\right) n \\ 1 - \frac{H_{x,\beta}}{H_{d_{max},\beta}} &= c \\ (1-c)H_{d_{max},\beta} &= H_{x,\beta} \\ (1-c) \sum_{k=1}^{d_{max}} \frac{1}{k^\beta} &= \sum_{k=1}^x \frac{1}{k^\beta} \quad (1-c) \sum_{k=x+1}^{d_{max}} \frac{1}{k^\beta} = c \sum_{k=1}^x \frac{1}{k^\beta} \end{aligned}$$

$$\text{Known special case: } H_{2x,2} = \frac{1}{2} \left(\zeta(2) + \frac{1}{2} (H_{x,2} + H_{x-\frac{1}{2},2}) \right)$$

$$\begin{aligned} (1-c) &\left(\zeta(\beta) - \frac{1}{(\beta-1)d_{max}^{\beta-1}} - \frac{1}{2d_{max}^\beta} - O\left(\frac{1}{d_{max}^{\beta+1}}\right) \right) = \\ &= \zeta(\beta) - \frac{1}{(\beta-1)x^{\beta-1}} - \frac{1}{2x^\beta} - O\left(\frac{1}{x^{\beta+1}}\right) \\ \frac{1}{(\beta-1)x^{\beta-1}} + \frac{1}{2x^\beta} &- \frac{1-c}{(\beta-1)d_{max}^{\beta-1}} - \frac{1-c}{2d_{max}^\beta} + O\left(\frac{1}{x^{\beta+1}}\right) = c \zeta(\beta) \end{aligned}$$

$$\begin{aligned}
(1-c)H_{d_{max},\beta} &= H_{x,\beta} \\
(1-c) \sum_{k=1}^{d_{max}} \frac{1}{k^\beta} &= \sum_{k=1}^x \frac{1}{k^\beta} \\
\zeta(\beta) - H_{a,\beta} &= \int_a^\infty \frac{1}{t^\beta} dt + \frac{1}{2a^\beta} - \sum_{i=2}^k \frac{b_i}{i!} f^{(i-1)}(a) - \int_a^\infty \frac{B_k(\{1-t\})}{k!} f^{(k)}(t) dt = \\
&= \frac{1}{(\beta-1)a^{\beta-1}} + \frac{1}{2a^\beta} - \frac{-\beta}{12a^{\beta+1}} - O\left(\frac{1}{a^{\beta+1}}\right)
\end{aligned}$$

$$\begin{aligned}
(1-c)H_{d_{max},\beta} &= (1-c) \left(\zeta(\beta) - \frac{1}{(\beta-1)d_{max}^{\beta-1}} - \frac{1}{2d_{max}^\beta} + O\left(\frac{1}{d_{max}^{\beta+1}}\right) \right) \\
H_{d_0,\beta} &= \zeta(\beta) - \frac{1}{(\beta-1)d_0^{\beta-1}} - \frac{1}{2d_0^\beta} + O\left(\frac{1}{d_0^{\beta+1}}\right)
\end{aligned}$$

If $d_0 = xd_{max}$,

$$\begin{aligned}
d_{max} &= e^{\alpha/\beta} \\
n &= e^\alpha H_{d_{max},\beta} \approx e^\alpha \zeta(\beta) = d_{max}^\beta \zeta(\beta)
\end{aligned}$$

$$\begin{aligned}
n' &= e^\alpha (H_{d_{max},\beta} - H_{xd_{max},\beta}) = \\
&= e^\alpha \left(\frac{x^{-\beta+1} - 1}{(\beta-1)d_{max}^{\beta-1}} + \frac{x^{-\beta} - 1}{2d_{max}^\beta} + O\left(\frac{1}{d_{max}^{\beta+1}}\right) \right) = \\
&= \frac{x^{-\beta+1} - 1}{(\beta-1)} d_{max} + \frac{x^{-\beta} - 1}{2} + O\left(\frac{1}{d_{max}}\right) \approx \frac{x^{-\beta+1} - 1}{(\beta-1)} \left(\frac{n}{\zeta(\beta)}\right)^{1/\beta}
\end{aligned}$$

We now choose $x = n^y$.

$$\begin{aligned}
n' &\approx \frac{n^{1/\beta - y(\beta-1)}}{(\beta-1)\zeta(\beta)^{1/\beta}} \\
1/\beta - y(\beta-1) &= 1 \\
y &= \frac{1/\beta - 1}{\beta-1} = -1/\beta
\end{aligned}$$

$$d_0 = \frac{d_{max}}{n^{1/\beta}} \implies n' \approx \frac{n}{(\beta-1)\zeta(\beta)^{1/\beta}} = \Theta(n).$$

We note that $(\beta-1)\zeta(\beta)^{1/\beta} > 1$ requires $\beta \leq 1.6$.

Finally, we get the expected average degree d for different ranges of β :

$$d = \frac{\text{Vol}(H)}{n'} = \frac{H_{d_{max},\beta-1} - H_{d_0,\beta-1}}{H_{d_{max},\beta} - H_{d_0,\beta}}$$

$$1 < \beta < 2: \quad d = \frac{d_{max}^{2-\beta} - d_0^{2-\beta}}{(2-\beta)\Theta(1)} = \frac{1 - n^{-(2-\beta)/\beta}}{(2-\beta)\Theta(1)} d_{max}^{2-\beta} = \Theta(d_{max}^{2-\beta}) = \Theta(n'^{(2-\beta)/\beta})$$

$$\beta = 2: \quad d = \frac{\ln d_{max} - \ln d_0}{\Theta(1)} = \frac{\frac{1}{\beta} \ln n}{\Theta(1)} = \Theta(\ln n')$$

$$\beta > 2: \quad d = \Theta(1)$$

We can see that Lemma 4.2 is applicable for $1 \leq \beta \leq 1.6$ because $d = \omega(\ln n')$. \square

Proof of Theorem 4.1. For $1 \leq \beta \leq 1.6$ we use Lemma 4.4 and Lemma 4.2 and we are done. Otherwise, we get n' and $\text{Vol}(H)$ from Lemma 4.3, which covers all possible values of β .

Case $0 < \beta < 1$:

$$\text{Vol}(S) = s \frac{1-\beta}{2-\beta} \frac{1-c^{2-\beta}}{1-c^{1-\beta}} e^{\alpha/\beta} = \frac{(1-\beta)^2}{2-\beta} \frac{1-c^{2-\beta}}{(1-c^{1-\beta})^2} sn'$$

Let $c_1 = \frac{(1-\beta)^2}{2-\beta} \frac{1-c^{2-\beta}}{(1-c^{1-\beta})^2} > 0$, then $\text{Vol}(S) = c_1 sn'$.

$$\Pr_{G,S}[S \text{ is non-expanding}] \leq \exp(\gamma s - \mu/2) = \exp((\gamma - c_1(n' - s)/2) s).$$

$$\Pr_G[H \text{ is not } (\epsilon n', \gamma) \text{ edge expander}] \leq \sum_{s=1}^{\epsilon n'} \binom{n'}{s} \Pr_{G,S}[S \text{ is non-expanding}] \leq$$

$$\leq \sum_{s=1}^{\epsilon n'} \left(e^{\gamma+1} \frac{n'}{s} e^{-c_1(n'-s)/2} \right)^s = o(1).$$

When $1 \leq s \leq \epsilon n'$, the last equality holds for any $\epsilon > 0$ satisfying the following condition:

$$e^{\gamma+1} \frac{n'}{s} e^{-c_1(n'-s)/2} < \frac{1}{10}$$

$$e^{\gamma+1} n' e^{-c_1 n'(1-\epsilon)/2} < \frac{1}{10}$$

$$c_1 n'(1-\epsilon)/2 > \ln(10e^{\gamma+1} n')$$

$$\epsilon < 1 - \frac{2}{c_1 n'} \ln(10e^{\gamma+1} n') = 1 - \frac{\ln(c_2 n')}{c_3 n'} = 1 - o(1)$$

We choose $\epsilon = 1/2$.

To satisfy the requirement (4.12) when $1 \leq s \leq n'/2$,

$$\gamma < \mu/s = c_1 \left(1 - \frac{s}{n'}\right) n'$$

$$\gamma = \frac{c_1}{2} n' - \delta, \text{ for any } \delta \geq 0$$

Finally, we restrict $c < 1/3$ to have $c_1 \leq 1$ and $\text{Vol}(S) \leq sn'$.

Note that if $0 < \beta < 1$ and $d_0 = 1$, then $n' = n$ and $\text{Vol}(H) = \text{Vol}(G)$ but the argument still holds, i.e., the whole graph G is an edge expander w.h.p.

Case $\beta = 1$:

$$\text{Vol}(S) = s \frac{1-c}{\ln 1/c} e^\alpha = \frac{1-c}{(\ln 1/c)^2} sn'$$

Let $c_1 = \frac{1-c}{(\ln 1/c)^2} > 0$, then $\text{Vol}(S) = c_1 sn'$.

The rest of the proof is identical to the previous case, but necessarily $c > 0$.

Case $1 < \beta < 2$:

$$\text{Vol}(S) = s \frac{\beta-1}{2-\beta} \frac{1-c^{2-\beta}}{c^{-\beta+1}-1} e^{\alpha/\beta} = \frac{(\beta-1)^2}{2-\beta} \frac{1-c^{2-\beta}}{(c^{-\beta+1}-1)^2} sn'$$

Let $c_1 = \frac{(\beta-1)^2}{2-\beta} \frac{1-c^{2-\beta}}{(c^{-\beta+1}-1)^2} > 0$, then $\text{Vol}(S) = c_1 sn'$.

The case is identical to $\beta = 1$.

Case $\beta = 2$:

$$\text{Vol}(S) = s \frac{\ln 1/c}{1/c-1} e^{\alpha/2} = \frac{\ln 1/c}{(1/c-1)^2} sn'$$

Analogously, $c_1 = \frac{\ln 1/c}{(1/c-1)^2} > 0$ and again $\text{Vol}(S) = c_1 sn'$.

Case $\beta > 2$: similarly to the case $1 < \beta < 2$,

$$\text{Vol}(S) = s \left(\frac{\beta-1}{\beta-2} \frac{c^{-\beta+2}-1}{c^{-\beta+1}-1} e^{\alpha/\beta} + O(1) \right) \approx \frac{(\beta-1)^2}{\beta-2} \frac{c^{-\beta+2}-1}{(c^{-\beta+1}-1)^2} sn'$$

□

4.2 Permutation Model

According to our choice of H , $n' = |V_H|$ is the largest number from $\{1, \dots, n\}$ such that

$$\deg(n') = \frac{pn}{n'^\beta} \geq d_0 \tag{4.24}$$

$$n' \approx (pn/d_0)^{1/\beta} \tag{4.25}$$

$$\text{vol}(H) = \sum_{v=1}^{n'} \frac{pn}{v^\beta} = pn H_{n',\beta} = d_0 n'^\beta H_{n',\beta} \tag{4.26}$$

$$\text{vol}(S) = s \frac{\text{vol}(H)}{n'} \tag{4.27}$$

We can apply Theorem 2.13 to show small constant conductance of H , as long as $d_0 \geq 3$. For edge and vertex expansion we have the following theorems.

Theorem 4.5 (Existence of an edge expander). $\forall 0 < p \leq 1 \exists d_0 = d_0(p, n, \beta) \forall \delta > 0 \exists \gamma, \epsilon > 0$: let G be a random power-law graph where each vertex v has $\deg(v) = pnv^{-\beta}$, and H is its induced subgraph of size n' obtained by retaining vertices of degree at least d_0 .

Then if $\beta = 0$, the whole graph G is $(\epsilon n, \gamma)$ edge expander w.h.p.

Otherwise, H is $(\epsilon n', \gamma)$ edge expander w.h.p.

When $\beta > 1$, the additional requirement is $p\zeta(\beta) > 2$. More roughly, $\zeta(\beta) > 2$ or $\beta \leq 1.72$, and the smaller p , the smaller β for which subsets still expand.

$$n' = \begin{cases} n & \text{if } \beta = 0, \\ n/2 & \text{if } \beta > 0. \end{cases} \quad (4.28)$$

$$d_0 = \begin{cases} pn & \text{if } \beta = 0, \\ 2^\beta pn^{1-\beta} & \text{if } \beta > 0. \end{cases} \quad (4.29)$$

$$d = \begin{cases} d_0 & \text{if } \beta = 0, \\ d_0/(1-\beta) & \text{if } 0 < \beta < 1, \\ d_0 \ln n' & \text{if } \beta = 1, \\ d_0 \zeta(\beta) n'^{\beta-1} & \text{if } \beta > 1. \end{cases} = \begin{cases} pn & \text{if } \beta = 0, \\ \frac{2^\beta p}{1-\beta} n^{1-\beta} & \text{if } 0 < \beta < 1, \\ 2p \ln(n/2) & \text{if } \beta = 1, \\ 2p \zeta(\beta) & \text{if } \beta > 1. \end{cases} \quad (4.30)$$

$$\gamma = \begin{cases} d - 2 - \delta & \text{if } \beta = 0, \\ d/2 & \text{if } \beta > 0. \end{cases} \quad (4.31)$$

Theorem 4.6 (Existence of a vertex expander). $\forall 0 < p \leq 1 \exists d_0 = d_0(p, n, \beta) \forall \delta > 0 \exists \gamma, \epsilon > 0$: let G be a random power-law graph where each vertex v has $\deg(v) = pnv^{-\beta}$, and H is its induced subgraph of size n' obtained by retaining vertices of degree at least d_0 .

Then if $\beta = 0$, the whole graph G is $(\epsilon n, \gamma)$ vertex expander w.h.p.

Otherwise, if $0 < \beta < 1$, H is $(\epsilon n', \gamma)$ vertex expander w.h.p. At the same time, either $\epsilon = 1/2$ and $\gamma = 1 - \delta$, or ϵ is sufficiently small and

$$\gamma = \begin{cases} d/2 - 2 - \delta & \text{if } \beta = 0, \\ d_0/2 - 2 - \delta & \text{if } 0 < \beta < 1. \end{cases} \quad (4.32)$$

The proof of this theorem can be found on page 41.

Proof of Theorem 4.5. Case $\beta = 0$: $\deg(v) = pn$ for each $v \in V$, so we choose $d_0 = pn$, then trivially $H = G$ and $n' = n$. This reduces the proof to known case for regular graphs described in Section 2.5.2.

As for the other cases when $\beta > 0$, we want our subgraph H to have linear size, so we need to choose appropriate d_0 .

$$\begin{aligned} n' &= (pn/d_0)^{1/\beta} = cn \\ d_0 &= \frac{pn}{(cn)^\beta} = \frac{p}{c^\beta} n^{1-\beta} \end{aligned}$$

Fixing $c = 1/2$ gives us

$$\begin{aligned} n' &= n/2 \\ d_0 &= 2^\beta pn^{1-\beta} = 2pn'^{1-\beta} \end{aligned}$$

$$\begin{aligned} \Pr_G[H \text{ is not } (\epsilon n', \gamma) \text{ edge expander}] &= \sum_{s=1}^{\epsilon n'} \binom{n'}{s} \Pr_{G,S}[|e(S, V_H \setminus S)| < \gamma s] = \\ &= \sum_{s=1}^{\epsilon n'} \binom{n'}{s} \Pr_{G,S} \left[|e(S, S)| \geq \frac{\text{vol}(S) - \gamma s}{2} \right] \leq \\ &\leq \sum_{s=1}^{\epsilon n'} \binom{n'}{s} \frac{\binom{\text{vol}(H)/2}{(\text{vol}(S)-\gamma s)/2} \binom{\text{vol}(H) - (\text{vol}(S) - \gamma s)}{\gamma s}}{\binom{\text{vol}(H)}{\text{vol}(S)}} \leq \\ &\leq \sum_{s=1}^{\epsilon n'} \left(\frac{\epsilon n'}{s} \right)^s \frac{\left(e \frac{\text{vol}(H)}{\text{vol}(S) - \gamma s} \right)^{(\text{vol}(S) - \gamma s)/2} \left(e \frac{\text{vol}(H) - (\text{vol}(S) - \gamma s)}{\gamma s} \right)^{\gamma s}}{\binom{\text{vol}(H)}{\text{vol}(S)}} \leq \\ &\leq \sum_{s=1}^{\epsilon n'} \left(\frac{\epsilon n'}{s} \frac{\left(\frac{e \text{vol}(H)}{(\text{vol}(H)/n' - \gamma)s} \right)^{(\text{vol}(H)/n' - \gamma)/2} \left(\frac{e \text{vol}(H)}{\gamma s} \right)^\gamma}{\left(\frac{n'}{s} \right)^{\text{vol}(H)/n'}} \right)^s = \\ &= \sum_{s=1}^{\epsilon n'} \left(\frac{e^{1+\gamma/2}}{\gamma^\gamma} e^{d/2} \left(\frac{1}{d-\gamma} \right)^{(d-\gamma)/2} \left(\frac{\text{vol}(H)}{s} \right)^{(d+\gamma)/2} \left(\frac{s}{n'} \right)^{d-1} \right)^s = \\ &= \sum_{s=1}^{\epsilon n'} \left(\frac{e^{1+\gamma/2}}{\gamma^\gamma} e^{d/2} \frac{d^{(d+\gamma)/2}}{(d-\gamma)^{(d-\gamma)/2}} \left(\frac{s}{n'} \right)^{d/2 - \gamma/2 - 1} \right)^s, \end{aligned}$$

where $d = \text{vol}(H)/n'$ is the average degree of H .

In order to upper bound the probability of H not being an edge expander we need

$$d - \gamma - 2 > 0 \tag{4.33}$$

Let $\gamma = d/2$ so we can simplify further. Note that (4.33) becomes $d > 4$.

$$\begin{aligned} \Pr_G[H \text{ is not } (\epsilon n', \gamma) \text{ edge expander}] &\leq \sum_{s=1}^{\epsilon n'} \left(\frac{e^{1+d/4}}{(d/2)^{d/2}} e^{d/2} \frac{d^{3d/4}}{(d/2)^{d/4}} \left(\frac{s}{n'} \right)^{d/4-1} \right)^s = \\ &= \sum_{s=1}^{\epsilon n'} \left(e(2e)^{3d/4} \left(\frac{s}{n'} \right)^{d/4-1} \right)^s. \end{aligned}$$

Unfortunately, we cannot use Proposition 2.3 with $c_1 = e(2e)^{3d/4}$, $c_2 = d/4 - 1$ because $c_2 < \log c_1$, even though the condition $c_1 = o(n'^{c_2-1})$ is met.

What's left to argue is that in all cases this expression may be upper bounded by $o(1)$, while also maintaining $d > 4$.

Case $0 < \beta < 1$:

$$\begin{aligned}\text{vol}(H) &= \frac{d_0}{1-\beta} n' \\ d &= \frac{d_0}{1-\beta} = \frac{2p}{1-\beta} n'^{1-\beta} = \Theta(n'^{1-\beta})\end{aligned}$$

Clearly, our choice of d_0 satisfies (4.33).

As for the summation, we consider small and large values of s separately.

$$\begin{aligned}\bullet \quad s \leq \sqrt{n'} : & \sum_{s=1}^{\sqrt{n'}} \left(e(2e)^{3d/4} \left(\frac{s}{n'} \right)^{d/4-1} \right)^s \leq e^{\sqrt{n'}} (2e)^{\sqrt{n'} 3d/4} \sum_{s=1}^{\sqrt{n'}} \left(\frac{s}{n'} \right)^{(d/4-1)s} \\ & \leq e^{\sqrt{n'}} (2e)^{\sqrt{n'} 3d/4} \sum_{s=1}^{\sqrt{n'}} \left(\frac{1}{\sqrt{n'}} \right)^{d/4-1} \\ & \leq e^{\sqrt{n'}} (2e)^{\sqrt{n'} 3d/4} \frac{\sqrt{n'}}{n'^{d/4-1}} \\ & = e^{n^{1/2}} (2e)^{\frac{3}{4} \Theta(n^{3/2-\beta})} \frac{1}{n'^{\Theta(n^{1-\beta})/4-3/2}} = o(1) \\ \bullet \quad \text{otherwise:} & \sum_{s=\sqrt{n'}+1}^{\epsilon n'} \left(e(2e)^{3d/4} \left(\frac{s}{n'} \right)^{d/4-1} \right)^s < \sum_{s=\sqrt{n'}+1}^{\epsilon n'} 10^{-s} \leq \frac{\epsilon n'}{10^{\sqrt{n'}}} = o(1)\end{aligned}$$

$$\text{For this we need: } e(2e)^{3d/4} \epsilon^{d/4-1} < \frac{1}{10}, \text{ or } \frac{((2e)^3 \epsilon)^{d/4}}{\epsilon} < \frac{1}{10e}.$$

$$\text{We can choose } \epsilon = \frac{1}{10(2e)^3}, \text{ and we are done because } d = \Theta(n'^{1-\beta}).$$

$$\Pr_G[H \text{ is not } (\epsilon n', \gamma) \text{ edge expander}] \leq o(1).$$

Case $\beta = 1$:

$$\begin{aligned}n' &= pn/d_0 \\ \text{vol}(H) &= d_0 n' \ln n' \\ d &= \frac{\text{vol}(H)}{n'} = d_0 \ln n' = \Theta(n'^{1-\beta} \ln n')\end{aligned}$$

This is identical to the previous case $0 < \beta < 1$.

Case $\beta > 1$:

$$\begin{aligned}\text{vol}(H) &= \zeta(\beta) d_0 n'^{\beta} \\ d_0 &= 2pn^{1-\beta} \\ d &= \zeta(\beta) d_0 n'^{\beta-1} = 2p \zeta(\beta) = O(1)\end{aligned}$$

Here we get an additional requirement $p\zeta(\beta) > 2$, because we need $d > 4$. As $0 < p \leq 1$ and $\zeta(\beta)$ is monotonically decreasing, it serves as an upper bound for β .

Finally, let $c_1 = e(2e)^{3d/4}$, $c_2 = d/4 - 1$ be some positive constants, then for sufficiently small ϵ :

$$\Pr_G^{\epsilon}[H \text{ is not } (\epsilon n', \gamma) \text{ edge expander}] \leq \sum_{s=1}^{\epsilon n'} \left(c_1 \left(\frac{s}{n'} \right)^{c_2} \right)^s = o(1).$$

More precisely, $c_1 \epsilon^{c_2} < 1/10$, or $\epsilon = (c_1 c_3)^{-1/c_2}$ for some $c_3 > 10$. \square

Proof of Theorem 4.6. The approach is identical to Theorem 2.12, but d perfect matchings on n vertices are replaced by a single perfect matching on $\text{vol}(H)$ mini-vertices, which correspond to n' vertices of H .

When $\beta = 0$, G is pn -regular graph, so the original proof with Proposition 2.2 or Proposition 2.3 suffice.

Otherwise, $\beta > 0$. For any $S \subset V_H$ of size $|S| = s \leq \epsilon n'$ and $T \subset V_H$ of size $(1 + \gamma)s$,

$$d = \frac{\text{vol}(H)}{n'} \tag{4.34}$$

$$\text{vol}(S) = sd \tag{4.35}$$

$$\text{vol}(T) = (1 + \gamma)sd \tag{4.36}$$

All mini-vertices from S are matched to mini-vertices in T with probability

$$\Pr_{G,S,T} [N_H(S) \subseteq T] \leq \left(\frac{\text{vol}(T)}{\text{vol}(H)} \right)^{\text{vol}(S)/2}$$

As for the number of sets S of a given size (the same goes for T), $\binom{n'}{s}$ would only work for regular graphs, while $\binom{\text{vol}(H)}{\text{vol}(S)}$ would overcount because mini-vertices corresponding to the same vertex in H are indistinguishable. Instead we use the trick as Gkantsidis et al. [GMS03] by taking the minimum degree d_0 into account.

$$\begin{aligned} \Pr_G^{\epsilon}[\exists \text{ non-expanding } S \text{ of size } s] &\leq \binom{\text{vol}(H)/d_0}{\text{vol}(S)/d_0} \binom{\text{vol}(H)/d_0}{\text{vol}(T)/d_0} \left(\frac{\text{vol}(T)}{\text{vol}(H)} \right)^{\text{vol}(S)/2} \leq \\ &\leq \left(\frac{\epsilon n'}{s} \right)^{sd/d_0} \left(\frac{\epsilon n'}{(1 + \gamma)s} \right)^{(1 + \gamma)sd/d_0} \left(\frac{(1 + \gamma)s}{n'} \right)^{sd/2} = \\ &= \left(\left(\frac{\epsilon n'}{s} \right)^{(2 + \gamma)d/d_0} \left(\frac{1}{1 + \gamma} \right)^{(1 + \gamma)d/d_0} \left(\frac{(1 + \gamma)s}{n'} \right)^{d/2} \right)^s = \\ &= \left(e^{(2 + \gamma)d/d_0} (1 + \gamma)^{d(1/2 - (1 + \gamma)/d_0)} \left(\frac{s}{n'} \right)^{d(1/2 - (2 + \gamma)/d_0)} \right)^s. \end{aligned}$$

In order to proceed we need this exponent to be positive:

$$d(1/2 - (2 + \gamma)/d_0) > 0 \tag{4.37}$$

$$0 \leq \gamma < d_0/2 - 2 \tag{4.38}$$

Note that this requires $d_0 > 4$, while similar theorems 2.12 and 2.13 ask for $d_0 \geq 3$.

When $\beta > 0$, it translates into

$$d_0 = 2^\beta pn^{1-\beta} > 4 \quad (4.39)$$

$$n^{1-\beta} > 2^{2-\beta} \quad (4.40)$$

$$\beta < 1 \quad (4.41)$$

$$\Pr_G[H \text{ is not } (\epsilon n', \gamma) \text{ vertex expander}] \leq \sum_{s=1}^{\epsilon n'} \Pr_G[\exists \text{ non-expanding } S \text{ of size } s] = o(1).$$

This final step may be achieved in two ways. First, we may apply Proposition 2.2, then ϵ must be some small constant. Secondly, it also follows from Proposition 2.3, but we need to verify $c_1 = o(n^{c_2-1})$ and $c_2 > \max\{1, \log c_1\}$.

$$c_1 = e^{(2+\gamma)d/d_0} (1 + \gamma)^{d(1/2 - (1+\gamma)/d_0)}$$

$$\log c_1 = (\log e)(2 + \gamma)d/d_0 + (\log(1 + \gamma))(d/2 - (1 + \gamma)d/d_0)$$

$$c_2 = d/2 - (2 + \gamma)d/d_0$$

$$c_2 - \log c_1 = d/2 - (2 + \gamma)d/d_0 - (\log e)(2 + \gamma)d/d_0 - (\log(1 + \gamma))(d/2 - (1 + \gamma)d/d_0)$$

$$c_2 - \log c_1 > 0, \text{ for any } \gamma < 1,$$

In this case we get $\epsilon = 1/2$, but both conditions are satisfied only when $\gamma < 1$. □

4.2.1 Error from Rounding Degrees

To get the exact degree sequence, one needs to round down the expression (3.16) for $\deg(v)$. In this case the error of $\text{vol}(G)$ is at most n . For our choice of d_0 this error is negligible compared to $\text{vol}(H)$ when $0 \leq \beta \leq 1$, but it might be significant otherwise.

$$\text{vol}(H) = \begin{cases} \Theta(n^2) & \text{if } \beta = 0, \\ \Theta(n^{2-\beta}) & \text{if } 0 < \beta < 1, \\ \Theta(n \log n) & \text{if } \beta = 1, \\ \Theta(n) & \text{if } \beta > 1. \end{cases} \quad (4.42)$$

4.3 “Octopus” Graphs

The core of the “octopus” graph is defined as an induced subgraph with all the vertices of degree at least $d_0 = n^{1/\log \log n}$.

The number of vertices of degree more than d is about $\left(\frac{\beta-2}{\beta-1}\right)^{\beta-1} n$.

$$\begin{aligned} w_x &= \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}} x^{-1/(\beta-1)} > d \\ \frac{\beta-2}{\beta-1} n^{\frac{1}{\beta-1}} &> x^{1/(\beta-1)} \\ x &< \left(\frac{\beta-2}{\beta-1}\right)^{\beta-1} n \end{aligned}$$

Analogously, i_1 is the index of the last vertex included in the core:

$$\begin{aligned} w_x &= \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}} x^{-1/(\beta-1)} \geq n^{1/\log \log n} = d_0 \\ \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}-1/\log \log n} &\geq x^{1/(\beta-1)} \\ x &\leq \left(\frac{\beta-2}{\beta-1} d\right)^{\beta-1} n^{1-\frac{\beta-1}{\log \log n}} = \left(\frac{c}{d_0}\right)^{\beta-1} = i_1 \end{aligned}$$

$i_1 \leq n$ implies a restriction on the average degree:

$$d \leq \frac{\beta-1}{\beta-2} n^{1/\log \log n} \quad (4.43)$$

By our assumptions, $i_1 = i_0 \left(\frac{d_{max}}{n^{1/\log \log n}}\right)^{\beta-1} \gg i_0$.

Let n' denote the size of the core.

$$n' = i_1 - i_0 \quad (4.44)$$

$$\log n' \approx \log i_1 \approx \log n \left(1 - \frac{\beta-1}{\log \log n}\right) \quad (4.45)$$

The average degree of the core is

$$\begin{aligned} \frac{\sum_{k=i_0}^{i_1} w_i}{i_1 - i_0} &\approx \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}} \frac{1}{i_1} \sum_{k=1}^{i_1} k^{-1/(\beta-1)} \approx \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}} \frac{1}{i_1} \frac{i_1^{1-1/(\beta-1)}}{(\beta-2)/(\beta-1)} = \\ &= d \left(\frac{n}{i_1}\right)^{\frac{1}{\beta-1}} = d \left(\left(\frac{\beta-2}{\beta-1} d\right)^{\beta-1} n^{-\frac{\beta-1}{\log \log n}}\right)^{-\frac{1}{\beta-1}} = d \left(\frac{\beta-2}{\beta-1} d\right)^{-1} n^{1/\log \log n} = \\ &= \frac{\beta-1}{\beta-2} n^{1/\log \log n} = \omega(\log n') \end{aligned}$$

As a result, the core is edge expander w.h.p. by Lemma 4.2.

Moreover, this core contains $G(n', p)$ with $p = d_0^2/dn$ [CL04], and if $d = \Theta(1)$, then its expected degrees are

$$n'p = \left(\frac{c}{d_0}\right)^{\beta-1} \frac{d_0^2}{dn} = \Theta\left(d^{\beta-2} d_0^{3-\beta}\right) = \Theta\left(d^{\beta-2} n^{(3-\beta)/\log \log n}\right) = \omega(\log n)$$

Chung and Lu [CL04] use this fact to prove $O(\log \log n)$ diameter of the core.

We can apply Theorem 2.17 on this $G(n', p)$ to get linear size vertex expander of diameter $O(\log n)$ inside of the core. If we could also say that other vertices are not “too far” from this expander, it could serve as an alternative step for the proof of $O(\log n)$ diameter of the whole “octopus” graph.

4.4 Diameter of Vertex Expanders

In this section we focus on $(\epsilon n, \gamma)$ vertex expanders, especially on the case $\epsilon < 1/2$, and the results apply to all graph models.

First, consider an example: a graph G that consists of two disconnected $(n/4, \gamma)$ vertex expanders of size $n/2$ each. Any subset of G of size up to $n/4$ is the union of subsets of these two parts, and it expands proportionally. Therefore, G is $(n/4, \gamma)$ vertex expander as well, and yet it is not connected.

In general, when $\epsilon \leq 1/4$, the connectivity of $(\epsilon n, \gamma)$ vertex expander is not guaranteed.

A well known fact about $(n/2, \gamma)$ vertex expanders is that they have diameter $O(\log n)$ [Rao12, HLW06]. We extend on this by demonstrating that the diameter is still $O(\log n)$ even for small constant $\epsilon < 1/2$, assuming the graph is connected.

Theorem 4.7. *Let $G = (V, E)$ be a connected $(s(n), \gamma)$ vertex expander of size n , where $s(n) \leq n/2$ and $\gamma = \Omega(1)$. Then the diameter of G is $O\left(\frac{n}{s(n)} \log s(n)\right)$.*

Corollary 4.7.1. *The diameter of a connected $(\epsilon n, \gamma)$ vertex expander of size n , where $\epsilon \leq 1/2$ and γ are some positive constants, is $O(\log n)$.*

Similarly, $s(n) = \Theta\left(\frac{n}{\log n}\right)$ gives us diameter $O(\log^2 n)$, and $s(n) = \Theta(\sqrt{n})$ gives $O(\sqrt{n} \log n)$. We now prove an intermediate result needed for the theorem.

Lemma 4.8. *If $G = (V, E)$ is a connected $(s(n), \gamma)$ vertex expander, then there exists a partitioning $\{S_1, \dots, S_k\}$ of V , such that $k = O(n/s(n))$ and each S_i has diameter $O(\log s(n))$.*

Proof. Let $B(v, r) = \{u \in V \mid \text{dist}(v, u) \leq r\}$ denote a ball of radius $r \geq 0$ around vertex v .

For any $v \in V$ and $r \geq 1$, if $|B(v, r-1)| \leq s(n)$, then $|B(v, r)| \geq \min\{s(n), (1+\gamma)^r\}$ by the expansion property of G . As all subsets of size up to $s(n)$ are expanding, there exists

$$r_0 = \left\lceil \log_{1+\gamma} s(n) \right\rceil = \Theta(\log s(n)), \tag{4.46}$$

such that $|B(v, r_0)| \geq s(n)$. The diameter of $B(v, r_0)$ is $O(r_0)$.

We maintain an invariant that all S_i are disjoint and have diameter $O(r_0)$ throughout the following partitioning process.

Begin with arbitrary $v_1 \in V$ and select $S_1 = B(v_1, r_0)$. As long as there is some other vertex v whose $B(v, r_0)$ contains at least $s(n)/2$ of not yet selected vertices, add these new

vertices to the next set S_i . We will stop after $k \leq 2n/s(n)$ steps. Every remaining vertex u must be within distance r_0 from some S_j , so we add each such u to the corresponding S_j . \square

Proof of Theorem 4.7. Lemma 4.8 gives us the partitioning $\{S_1, \dots, S_k\}$ of V . We create a graph G' of size k from G by contracting each S_i into a single vertex and merging multiple edges. Note that G' is connected by construction.

Now consider arbitrary u, v from G . Clearly, the distance between their corresponding vertices in G' is at most k . Let D be the maximum diameter of any S_i . Then $\text{dist}(u, v)$ in G is at most $k(2D + 1)$ which is exactly $O\left(\frac{n}{s(n)} \log s(n)\right)$. \square

Chapter 5

Comparison of the Graph Models

We have briefly compared our findings to some previous papers in Chapter 1. Now we would like to address more correspondences, both quantitative and qualitative.

5.1 Overview of the Expansion Properties

As usual, we denote the size of the subgraphs n' , the minimum degree d_0 and the average degree d , the expansion γ , and $\epsilon, \delta > 0$ are some arbitrary small constants.

5.1.1 Previously Known Results

We saw in Chapter 2 that random d -regular graphs are themselves $(\epsilon n, d - 2 - \delta)$ vertex and edge expanders.

The model from Gkantsidis et al. [GMS03] is similar to our coin toss model and it describes the graphs with a small constant conductance 0.0175, which is a generalization of $(n/2, \gamma)$ edge expansion.

Lastly, $G(n, p)$ graphs contain $(n'/2, \Theta(1))$ vertex expanders of linear size [Kri18].

5.1.2 Our Results

Graphs in the coin toss model contain $(n'/2, d/2 - \delta)$ edge expanders. The size of these subgraphs is $n' = \Theta(n)$ when $\beta \leq 1.6$, but only $n' = \Theta(n^{1/\beta})$ for larger β .

In the permutation model, on the other hand, we were able to prove the existence of $(n'/2, d/2)$ edge expanders if $\beta \leq 1.72$, and also both $(n'/2, 1 - \delta)$ and $(\epsilon n', d_0/2 - 2 - \delta)$ vertex expanders for $0 < \beta < 1$. The trade-off between the maximum size of expanding subsets and the expansion rate is the most apparent in this model. And as we know, any vertex expansion would imply the same edge expansion, but unluckily, we couldn't secure even weak vertex expansion for $\beta > 1.72$.

5.2 Uniform Random Graphs

Although power-law graphs have less uniform structure than $G(n, p)$, we produced evidence that they are still random enough because of the large expanding subgraphs.

$G(n, p)$ graphs have a sharp threshold for connectivity [ER59] and become connected w.h.p. when $p = (1 + \epsilon) \log n/n$, for $\epsilon > 0$, that is, when the expected degrees are $d \approx pn > \log n$. It harmonizes with our Lemma 4.2, which explains that graphs with the expected average degree $d > 10 \log n$ are $(n/2, d/2 - \delta)$ edge expanders, for any small constant $\delta > 0$.

Theorem 2.17 adds to this by saying that $G(n, p)$ graph in supercritical regime with $p = (1 + \epsilon)/n$ w.h.p. contains $(n'/2, \gamma)$ vertex expander on $n' = \Theta(n)$ vertices [Kri18].

The parameters are $c_1 = 1 + \epsilon$, $c_2 = 1 + \epsilon^2/10$, $\alpha = \left(\frac{c_2}{5c_1}\right)^{c_2/(c_2-1)}$, and $\Delta = 4 \ln \frac{1}{\epsilon}$.

In this case it guarantees the vertex expansion

$$\gamma = \frac{c_1 - c_2}{\Delta \log_2 \frac{1}{\alpha}} = \frac{\epsilon - \epsilon^2/10}{\left(4 \ln \frac{1}{\epsilon}\right) \frac{c_2}{c_2-1} \log_2 \frac{5c_1}{c_2}} = \frac{\epsilon - \epsilon^2/10}{\left(4 \ln \frac{1}{\epsilon}\right) \frac{1+\epsilon^2/10}{\epsilon^2/10} \log_2 \frac{5c_1}{c_2}} < \frac{\epsilon}{8 \ln(1/\epsilon)} \quad (5.1)$$

which is less than 1 for $\epsilon < 0.89$, so it is also similar to our $(n''/2, 1 - \delta)$ vertex expander for the permutation model with $0 < \beta < 1$, $n'' = n/2$, and arbitrary small constant $\delta > 0$.

5.3 Connected Components and the Coin Toss Model

Table 1.1 highlights similarities between the sizes of our expanding subgraphs in the coin toss model from Section 3.2 and the largest connected components of power-law graphs from Aiello et al. [ACL01], although there are still some gaps.

For example, we know that the giant component exists for $\beta \in (0; 3.48)$ [ACL01, CL06]. For $3.48 < \beta < 4$, the size of each connected component is at most $\Theta\left(n^{2/\beta} \log n\right)$, and for $\beta > 4$ it is $\Theta\left(n^{2/(\beta+2)} \log n\right)$.

Our proof provides a linear size edge expander only for β up to 1.6, and then we have a single jump to $\Theta\left(n^{1/\beta}\right)$, which is close to a square root of the size of the corresponding largest component. The conclusion here is that the largest connected components of power-law graphs have strong expansion properties, even though their average degree $d = \omega(1)$, so they are not sparse.

5.4 Coin Toss and Permutation Models

If we consider the graphs from permutation model when $p = 1$, we notice that we get the same $(n'/2, d/2 - \delta)$ edge expander as in the coin toss model for any $0 < \beta \leq 1.6$. It demonstrates that power-law graphs with small β possess a similar structure regardless of whether we define degrees or frequencies of degrees to follow a power law.

Chapter 6

Conclusions

In this work we established existence of expanders in power-law graphs under several models and studied their behavior for various ranges of the exponent β . We also made an overview of the structure of random graphs, comparing them side by side with our findings to contrast some correspondences and differences. In particular, we showed that power-law graphs with small β have similar to $G(n, p)$ expansion properties, just as one would expect. Also the largest components of power-law graphs and our edge expanders have comparable sizes, so these components are presumably well-connected.

One possible direction for future research is the improvement of the current results. This includes tightening gaps between sizes of connected components and expanding subgraphs, increasing the quality of expansion, and weakening the conditions for having large expanding sets of size up to $n/2$. It would be useful to connect spectral and combinatorial expansion of power-law graphs. By analogy with connectivity in percolation theory, interesting open problem is to decide the expansion of sparse graphs obtained by randomly removing each edge with probability proportional to the degrees of its endpoints.

Another promising direction is the development of enhanced SAT algorithms for power-law formulas, facilitated by the available structural information. Self-similarity of a power law might be exploited to design recursive SAT algorithms. Finally, the results about random walks with lookahead suggest studying the expansion of k -th power of power-law graphs, for some small constant k .

Bibliography

- [ABBO⁺05] M. Alekhnovich, A. Borodin, J. Buresh-Oppenheim, R. Impagliazzo, A. Magen, and T. Pitassi. Toward a model for backtracking and dynamic programming. In *Proc. of the 20th IEEE Conf. on Computational Complexity*, pages 308–322, 2005.
- [ABL09a] C. Ansotegui, M. L. Bonet, and J. Levy. On the structure of industrial SAT instances. In *Proc. of the 15th Int. Conf. on Principles and Practice of Constraint Programming*, pages 127–141, 2009.
- [ABL09b] C. Ansotegui, M. L. Bonet, and J. Levy. Towards industrial-like random SAT instances. In *Proc. of the 21st Int. Joint Conf. on Artificial Intelligence*, pages 387–392, 2009.
- [ABL⁺16] Z. Allen Zhu, A. Bhaskara, S. Lattanzi, V. Mirrokni, and L. Orecchia. Expanders via local edge flips. In *Proc. of the 27th ACM-SIAM Symposium on Discrete Algorithms*, pages 259–269, 2016.
- [ABSRW04] M. Alekhnovich, E. Ben-Sasson, A. Razborov, and A. Wigderson. Pseudorandom generators in propositional proof complexity. *SIAM Journal on Computing*, 34(1):67–88, 2004.
- [ACL01] W. Aiello, F. Chung, and L. Lu. A random graph model for power law graphs. *Experimental Mathematics*, 10(1):53–66, 2001.
- [AHI05] M. Alekhnovich, E. A. Hirsch, and D. Itsykson. Exponential lower bounds for the running time of DPLL algorithms on satisfiable formulas. *Journal of Automated Reasoning*, 35(1):51–72, 2005.
- [AL06] A. Amit and N. Linial. Random lifts of graphs: Edge expansion. *Combinatorics, Probability and Computing*, 15(3):317–332, 2006.
- [BA99] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [BB03] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, 288:60–69, 2003.
- [Bol88] B. Bollobás. The isoperimetric number of random regular graphs. *European Journal of Combinatorics*, 9(3):241–244, 1988.

- [Che69] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Proc. of the Princeton conference in honor of Professor S. Bochner*, pages 195–199, 1969.
- [Chu97] F. Chung. *Spectral Graph Theory*. Number 92 in Regional Conf. Series in Mathematics. American Mathematical Society, 1997.
- [Chu07] F. Chung. Four proofs for the Cheeger inequality and graph partition algorithms. In *Proc. of the 4th Int. Congress of Chinese Mathematicians*, volume 2, 2007.
- [CL01] F. Chung and L. Lu. The diameter of sparse random graphs. *Advances in Applied Mathematics*, 26:257–279, 2001.
- [CL04] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2004.
- [CL06] F. Chung and L. Lu. *Complex Graphs and Networks*. Number 107 in Regional Conf. Series in Mathematics. American Mathematical Society, 2006.
- [Din07] I. Dinur. The PCP theorem by gap amplification. *Journal of the ACM*, 54(3), 2007.
- [ER59] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.
- [ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61, 1960.
- [FFF99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 29(4):251–262, 1999.
- [FGRV14] A. Frieze, N. Goyal, L. Rademacher, and S. Vempala. Expanders via random spanning trees. *SIAM Journal on Computing*, 43(2):497–513, 2014.
- [Fog08] A. Fog. Calculation methods for Wallenius’ noncentral hypergeometric distribution. *Communications in Statistics, Simulation and Computation*, 37(2):258–273, 2008.
- [Gil59] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [GMS03] C. Gkantsidis, M. Mihail, and A. Saberi. Conductance and congestion in power law graphs. *ACM SIGMETRICS Performance Evaluation Review*, 31(1):148–159, 2003.
- [GR54] B. Gutenberg and C. F. Richter. *Seismicity of the Earth and Associated Phenomena*. Princeton University Press, Princeton, New Jersey, 1954.

- [Hak62] S. Hakimi. On realizability of a set of integers as degrees of the vertices of a linear graph. *Journal of the Society for Industrial and Applied Mathematics*, 10(3):496–506, 1962.
- [HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.
- [Hop08] J. Hopcroft. CS 683 lecture notes: Advanced design & analysis of algorithms. <http://www.cs.cornell.edu/courses/cs683/2008sp/lecture%20notes/Lecture27.pdf>, 2008.
- [KB67] A. N. Kolmogorov and Y. M. Barzdin. On the realization of networks in three-dimensional space. *Problems of Cybernetics*, 19:261–268, 1967. [In Russian].
- [Kri18] M. Krivelevich. Finding and using expanders in locally sparse graphs. *SIAM Journal on Discrete Mathematics*, 32(1):611–623, 2018.
- [Lot26] A. J. Lotka. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323, 1926.
- [Mag06] A. Magen. CSC 2414H lecture notes: Metric embeddings. <http://www.cs.toronto.edu/~avner/teaching/S6-2414/TUT2.pdf>, 2006.
- [Mah09] M. Mahoney. CS369M lecture notes: Algorithms for modern massive data set analysis. <https://cs.stanford.edu/people/mmahoney/cs369m/Lectures/lecture11.pdf>, 2009.
- [Mih89] M. Mihail. Conductance and convergence of Markov chains — a combinatorial treatment of expanders. In *Proc. of the 30th Symposium on Foundations of Computer Science*, pages 526–531, 1989.
- [MS17] G. Moshkovitz and A. Shapira. Decomposing a graph into expanding subgraphs. *Random Structures & Algorithms*, 52(1):158–178, 2017.
- [MST06] M. Mihail, A. Saberi, and P. Tetali. Random walks with lookahead on power law random graphs. *Internet Mathematics*, 3(2):147–152, 2006.
- [Par97] V. Pareto. *Cours d’économie politique*, 1897. [In French].
- [Pin73] M. S. Pinsker. On the complexity of a concentrator. In *Proc. of the 7th Int. Teletraffic Conf.*, pages 318/1–318/4, 1973.
- [PRST16] T. Pitassi, B. Rossman, R. A. Servedio, and L. Tan. Poly-logarithmic Frege depth lower bounds via an expander switching lemma. In *Proc. of the 48th ACM Symposium on Theory of Computing*, pages 644–657, 2016.
- [Pud15] D. Puder. Expansion of random graphs: New proofs, new results. <https://arxiv.org/abs/1212.5216>, 2015.
- [Rao12] A. Rao. CSE531 lecture notes: Computational complexity theory. <https://homes.cs.washington.edu/~anuprao/pubs/CSE531Winter12/lecture11.pdf>, 2012.

- [Vad12] S. Vadhan. *Pseudorandomness*. Now Foundations and Trends, 1 edition, 2012.
- [Yul25] G. U. Yule. A mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London*, 213:21–87, 1925.
- [Zip49] G. K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Cambridge, MA, 1949.

Appendix A

Miscellaneous

A.1 Approximations

A.1.1 Euler-Maclaurin Formula

Euler-Maclaurin summation formula provides an approximation of the sum $\sum_{i=0}^n f(i)$ via the integral $\int_0^n f(x) dx$, and the error term is an integral with Bernoulli numbers.

$$\begin{aligned} \sum_{n=a}^b f(n) &= \int_a^b f(t) dt + \frac{1}{2}(f(b) + f(a)) + \\ &+ \sum_{i=2}^k \frac{b_i}{i!} (f^{(i-1)}(b) - f^{(i-1)}(a)) - \\ &- \int_a^b \frac{B_k(\{1-t\})}{k!} f^{(k)}(t) dt \end{aligned} \quad (\text{A.1})$$

where $\{x\}$ denotes the fractional part of x . One of its interesting applications is the Stirling's approximation formula.

If $f(x)$ and all its derivatives tend to 0 as $x \rightarrow \infty$, the formula can be simplified.

$$\sum_{n=a}^{\infty} f(n) = \int_a^{\infty} f(t) dt + \frac{1}{2}f(a) - \sum_{i=2}^k \frac{b_i}{i!} f^{(i-1)}(a) - \int_a^{\infty} \frac{B_k(\{1-t\})}{k!} f^{(k)}(t) dt \quad (\text{A.2})$$

For $k \geq 2$: $|B_k(\{x\})| \leq \frac{\pi^2}{3} \frac{k!}{(2\pi)^k} < 4 \frac{k!}{(2\pi)^k}$.

A.1.2 Lambert Function

Lambert function $W(x)$ is defined by the relation

$$z = W(ze^z) \tag{A.3}$$

When $x > 0$, $W(x) = \ln x - \ln W(x)$, and $\lim_{x \rightarrow \infty} \frac{W(x)}{\ln x} = 1$. Thus, we can approximate it as

$$W(x) \approx \ln x \tag{A.4}$$

A.2 Finding Expanding Subsets In Locally Sparse Graphs

Algorithm 1 Algorithmic proof of Theorem 2.14 ([Kri18])

```

procedure FINDEXPANDER( $G, \alpha$ )
   $V_1 \leftarrow V$ 
  for  $i \leftarrow 1, i_{max}$  do
     $d_i = \frac{|E_i|}{|V_i|} \geq c_2$ 
     $\delta =$  edge boundary of the sparsest cut  $(W_i, V_i \setminus W_i)$ 
    if  $|V_i| \leq \alpha n$  then ▷  $G_i$  is small and dense
      break
    else if there are isolated vertices then ▷ density increases
      remove isolated vertices
      continue
    else if  $\lambda_{G_i} > \frac{\delta^2}{2\Delta^2}$  then
       $G_i$  is edge-expander by (2.47)(2)
      and  $\Delta(G_i) \leq \Delta$  ▷  $G_i$  is a vertex-expander
      break
    else
      find  $W_i \subset V_i$  using the proof of (2.47)(1)
      s.t.  $\text{vol}(W_i) \leq \text{vol}(V_i)/2$ 
      and  $e_{G_i}(W_i, V_i \setminus W_i) \leq \delta|W_i|$ 
      if  $W_i$  touches at most  $d_i|W_i|$  edges then
        remove  $W_i$  ▷ density doesn't decrease
      else ▷  $W_i$  spans less than  $(d_i - \delta)|W_i|$  edges
        remove  $V_i \setminus W_i$  ▷ density might decrease
         $d_{i+1} \geq d_i - \delta$ 
      end if
    end if
  end for
end procedure

```
