Mininum Vertex Cover in Generalized Random Graphs with Power Law Degree Distribution

André L. Vignatti[a], Murilo V. G. da Silva[b]

[a]*DINF – Federal University of Paraná – Curitiba, Brazil*
[b]*DAINF – Federal University of Technology - Paraná – Curitiba, Brazil*

Abstract

In this paper we study the approximability of the minimum vertex cover problem in power law graphs. In particular, we investigate the behavior of a standard 2-approximation algorithm together with a simple pre-processing step when the input is a random sample from a generalized random graph model with power law degree distribution. More precisely, if the probability of a vertex of degree i to be present in the graph is $ci^{-\beta}$, where $\beta > 2$ and c is a normalizing constant, the expected approximation ratio is $1 + \zeta(\beta)^{-1}\mathrm{Li}_\beta(e^{-\rho(\beta)})$, where $\zeta(\beta)$ is the Riemann Zeta function of β, $\mathrm{Li}(\beta)$ is the polylogarithmic special function of β and $\rho(\beta) = \frac{\mathrm{Li}_{\beta-2}(\frac{1}{e})}{\zeta(\beta-1)}$.

Keywords: Chung-Lu random graph model, Britton random graph model, generalized random graph model, approximation algorithms, vertex cover problem, power-law graphs

1. Introduction

The empirical study of large real world networks in the late 1990's and early 2000's [1, 2, 3, 4, 5, 6, 7, 8, 9] showed that the vertex degree distribution of a variety of technological, biological and social networks can be approximated by a power law, i.e., the number of nodes of a given degree i is proportional to $i^{-\beta}$, where $\beta > 0$. There is some evidence that some optimization problems might be easier in networks with such degree distribution [9, 10, 11, 12, 13] than in general networks. More recently, some analytical results investigating optimization problems on model of power law graphs appeared in [13, 14, 15, 16, 17]. The problem that we deal in this paper fits in this context. We investigate the approximability of the minimum vertex cover problem in the generalized random graph model with power law degree distribution.

Random graphs with arbitrary degree distributions have been studied before [18, 19, 20, 21, 22], but in the last 15 years models for random graphs with a power law degree distribution, refered here as *power law graphs*, have received a special interest [23, 24, 25, 26, 27, 28] due to the availability of empirical data. These models can be roughly divided into three groups: *generalized random graph* model, *configuration* model and *preferential attachment* model.

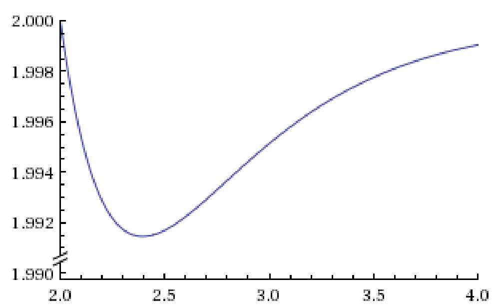
The generalized random graph model, proposed by Britton *et al* [27] in 2006, is a generalization of the Erdös-Rényi model [29, 30] where edges are added independently, but moderated by vertex weights that are used to generate arbitrary distribution of vertex degrees, including the power law distribution (note also that the case where every vertex have the same weight, this model is equivalent to the $G_{n,p}$ model). The well known Chung-Lu model [31] also fits in this category. The results in this paper are proved in the generalized random graph model, in particular due to the advantage of edge probabilities being independent. The configuration model uses a different approach where edge connections are probabilistic, but the list of vertex degrees (and therefore then number of edges) is

Email addresses: `vignatti@inf.ufpr.br` (André L. Vignatti), `murilo@dainf.ct.utfpr.br` (Murilo V. G. da Silva)

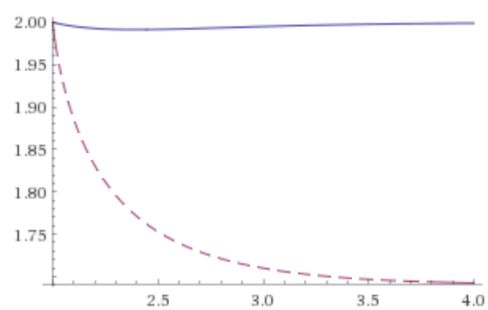
fixed. This scheme originated in the late 1970's [18, 19], but currently the most well know such model in the specific context of power law graphs is the ACL(α, β) model, proposed by Aiello *et al* [23]. A third well know approach is the preferential attachment model. The idea is to describe the growth of a graph driven by a random process in which new vertices connect to the current graph with probability proportional to the degree of the vertices already in place. This model was first described by Barabasi and Albert [2] and then more formally by Bollobás *et al* [24]. In [32], Hofstad gives a detailed and formal treatment of these three general categories for random power law graphs.

A vertex cover in a graph $G = (V, E)$ is a set of vertices $S \subseteq V$ such that every edge $e \in E$ has at least one endpoint in S. Finding a minimum vertex cover is a well known $\mathcal{NP}$-hard problem [33]. The problem is also conjectured not to admit approximation algorithms with constant factor smaller than 2, unless $\mathcal{P} = \mathcal{NP}$ [34]. Recently, Gast and Hauptmann [14] used ACL(α, β) model to show that there is an approximation algorithm with expected factor of approximation strictly smaller than 2 for random power law graphs (see Figure 1a for a plot of such approximation ratio for different values of β). In this paper we show a similar result, but in the generalized random graph model proposed by Britton *et al.* Furthermore, we also show that the same results hold for the Chung-Lu random graph model. We obtained a better expected approximation factor using a simpler algorithm, even though its worth mentioning that these approximation factors cannot be directly compared, since the underlying random graph models are not the same.

Consider the following algorithm for vertex cover on an input graph G: *Step 1:* Insert in the vertex cover C every vertex that is adjacent to a vertex of degree 1; *Step 2:* Remove from G every vertex that is either of degree 1 or adjacent to a vertex of degree 1; *Step 3:* Run any 2-approximation algorithm (e.g., [35], chapter 35) on the remaining graph and let C' be the output of this algorithm; *Step 4:* Output $C \cup C'$ as a cover for G. We show that the expected approximation ratio of this algorithm is $1 + \zeta(\beta)^{-1}\mathrm{Li}_\beta(e^{-\rho(\beta)})$, where $\beta > 2$ is a parameter related to the exponent of the power law, ζ is the Riemann Zeta function, Li is the polylogarithmic special function and $\rho(\beta) = \frac{\mathrm{Li}_{\beta-2}(\frac{1}{e})}{\zeta(\beta-1)}$. This bound can be better understood in Figure 1b, where we plot this function for $2 < \beta \leq 4$, as well as the expected approximation ratio of the algorithm of Gast and Hauptman for the sake of comparison.



(a) Expected approximation ratio of the algorithm of Gast and Hauptman [14] for power law graphs in the ACL (α, β) model. The plotted curve is the approximation ratio of $2 - \frac{\zeta(\beta)-1-\frac{1}{2^\beta}}{2^\beta\zeta(\beta-1)\zeta(\beta)}$ for values of β from 2 to 4.



(b) The blue solid line refers to the expected approximation ratio of the algorithm of Gast and Hauptmann [14] and the red dashed line refers to the expected approximation ratio of algorithm described in the current paper for values of β from 2 to 4.

Figure 1: Expected approximation ratio of algorithms for power law graphs.

2. Generalized random graph model

In this section we describe generalized random graphs (GRG), introduced by Britton [27]. For a detailed introduction on this model and a few asymptotic equivalent variants we refer to [32].

In this model we start with a vertex set $V = \{1, 2, ..., |V|\}$ and each vertex $i \in V$ is associated to a weight w_i. Let w be a vector with entries $w_1, ..., w_{|V|}$. In the edge set E, every edge ij is created independently with probability $\Pr(ij \in E) = \frac{w_i w_j}{\ell_n + w_i w_j}$, where $\ell_n = \sum_{k \in V} w_k$. Naturally, the degree distribution of the random graph depends on the vector w. In order to create a power law random graph with a given exponent $\beta > 2$ we build w using principles similar to Aiello *et al* [23]: Let $\alpha = \ln(\frac{|V|}{\zeta(\beta)})$, where $\zeta(\beta) = \sum_{j=1}^{\infty} \frac{1}{j^\beta}$ is the Riemann Zeta function. Let $\Delta = \lfloor e^\alpha \rfloor$. For each $j = 1, ..., \Delta$, let

$$y_j = \begin{cases} \lfloor \frac{e^\alpha}{j^\beta} \rfloor & \text{if } n > 1 \\ \lfloor e^\alpha \rfloor & \text{otherwise} \end{cases}$$

Now construct w so that y_j of its entries are equal to j, for $j = 1, ..., \Delta$. In this way there are y_j vertices with weight j (note that, similarly, in the ACL model there are y_j vertices of a given fixed degree j). From the definition of α, observe that $|V| = e^\alpha \zeta(\beta)$. As discussed in [23], we can ignore rounding and deal with $\frac{e^\alpha}{i^\beta}$ and $e^{\frac{\alpha}{\beta}}$ as real numbers. Furthermore, note that in the ACL(α, β) model in the definition of y_i some extra care has to be taken, since the sum vertex degrees has to be even. In our model, no such restriction is necessary since y_i refers to the vector of weights. Throughout the paper a graph $G = (V, E)$ is a random sample in the GRG model.

Lemma 2.1. *Let $i, j \in \{1, \ldots, e^{\frac{\alpha}{\beta}}\}$. Then $e^\alpha \zeta(\beta - 1) + ij \approx e^\alpha \zeta(\beta - 1)$*

Proof. We show that

$$\lim_{\alpha \to \infty} \frac{e^\alpha \zeta(\beta - 1) + ij}{e^\alpha \zeta(\beta - 1)} = 1$$

The limit above is the same as

$$\lim_{\alpha \to \infty} 1 + \frac{ij}{e^\alpha \zeta(\beta - 1)} = 1 \quad \therefore \quad \lim_{\alpha \to \infty} \frac{ij}{e^\alpha \zeta(\beta - 1)} = 0$$

$$\therefore \quad \lim_{\alpha \to \infty} \frac{ij}{e^\alpha} = 0$$

But

$$\lim_{\alpha \to \infty} \frac{ij}{e^\alpha} \leq \lim_{\alpha \to \infty} \frac{e^{\frac{\alpha}{\beta}} e^{\frac{\alpha}{\beta}}}{e^\alpha} = \lim_{\alpha \to \infty} e^{\alpha(\frac{2}{\beta} - 1)}$$

Since $\beta > 2$, the limit goes to zero. □

From the definition of $\Pr(ij \in E)$ and Lemma 2.1, we have

$$\Pr(ij) = \frac{ij}{e^\alpha \zeta(\beta - 1) + ij} \approx \frac{ij}{e^\alpha \zeta(\beta - 1)} \tag{1}$$

We note that in the Chung-Lu [31] random graph model, the probability of an edge ij is defined to be $\Pr(ij \in E) = \frac{w_i w_j}{\ell_n}$ (instead of $\frac{w_i w_j}{\ell_n + ij}$, as we have in our definition). Equation 1 shows that $\frac{w_i w_j}{\ell_n} \approx \frac{w_i w_j}{\ell_n + ij}$ and therefore all results in this paper also hold in the Chung-Lu model for the input distribution that we are working here. Conversely, since the edge probability in Chung-Lu model is defined in way that the expected degree of a vertex i is w_i, we also have this property in our model.

Definition 2.2 (**Set of vertices with same weight**). *Let W_k the set set of vertices with weight k, i.e., $W_k = \{l \in V | w_l = k\}$. Let p_{ij} be the probability of a random vertex of W_i be adjacent to a random vertex of W_j.*

Since vertices with a given weight are interchangeable we have

$$p_{ij} \approx \frac{ij}{e^{\alpha}\zeta(\beta-1)}$$

In the literature p_{ij} usually refers to the probability of an edge connecting vertex i and vertex j. The notation here is different since in the entire paper it is much clear to argue conditioned on the vertex weight instead of the vertex index.

We close this section giving some further graph theoretical notation and definitions.

Definition 2.3 (**Graphs theoretical definitions**). *The degree of $v \in V$ is denoted by $d(v)$. We denote V_k the set of vertices of degree k. The maximum degree of G is denoted by Δ, i.e., the largest k such that $|V_k| \neq \emptyset$. Let $V^- = V \setminus (V_0 \cup V_1)$. For $S \subseteq V$, denote $G[S]$ the graph induced by S and denote $N(S)$ the set o vertices containing at least one neighbor in S.*

3. Technical Lemmas

In this section our aim is to estimate the size of $N(V_1)$, which is the result of Lemma 3.13. We need such estimate since the first step of the approximation algorithm consists of including every vertex of $N(V_1)$ in the solution. The vertices that potentially belong to $N(V_1)$ are those in $V^- = V \setminus (V_0 \cup V_1)$. We estimate the probability of a vertex to belong to V_0 and V_1 in Lemmas 3.5 and 3.6 respectively. Instead of directly calculating the probability of a vertex of V^- being in $N(V_1)$, we found it easier to estimate the probability of the complementary event in Lemma 3.12.

Lemma 3.1. *Let $q_{ik} = 1 - p_{ik}$. Then,*

$$\prod_{k=1}^{\Delta} q_{ik}^{|W_k|} \approx \frac{1}{e^i}$$

Proof.

$$\begin{aligned}
\prod_{k=1}^{\Delta} q_{ik}^{|W_k|} &= \prod_{k=1}^{\Delta} \left(1 - \frac{ik}{e^{\alpha}\zeta(\beta-1)}\right)^{\frac{e^{\alpha}}{k^{\beta}}} \\
&= \prod_{k=1}^{\Delta} \left(1 - \frac{(ik)/(\zeta(\beta-1))}{e^{\alpha}}\right)^{e^{\alpha}\cdot 1/k^{\beta}} \\
&\approx \prod_{k=1}^{\Delta} \left(\frac{1}{e^{\frac{ik}{\zeta(\beta-1)}}}\right)^{\frac{1}{k^{\beta}}} \\
&= \prod_{k=1}^{\Delta} \left(\frac{1}{e^{\frac{i}{\zeta(\beta-1)}}}\right)^{\frac{1}{k^{\beta-1}}} \\
&= \left(\frac{1}{e^{\frac{i}{\zeta(\beta-1)}}}\right)^{\sum_{k=1}^{\Delta} \frac{1}{k^{\beta-1}}} \\
&\approx \left(\frac{1}{e^{\frac{i}{\zeta(\beta-1)}}}\right)^{\zeta(\beta-1)} \\
&= \frac{1}{e^i}.
\end{aligned}$$

□

Lemma 3.2. *Let $v \in V$. Then*

$$\Pr(v \in W_i) = \frac{\frac{e^\alpha}{i^\beta}}{e^\alpha \zeta(b)} = \frac{1}{i^\beta \zeta(b)}.$$

Proof. Directly from the definition of $|W_i|$ and $|V|$. □

Lemma 3.3. *Let $v \in W_i$. Then*

$$\Pr(v \in V_0 | v \in W_i) \approx \frac{1}{e^i}.$$

Proof. Let $v \in W_i$ and let X be the random variable for $d(v)$. Let X_j be the random variable counting the number of neighbors of v in W_j. By linearity of expectation, $X = \sum_{j=1}^{\Delta} X_j$. Hence,

$$\begin{aligned}
\Pr(v \in V_0 | v \in W_i) &= \Pr(X = 0) \\
&= \Pr(X_1 + X_2 + \ldots + X_\Delta = 0) \\
&= \Pr(X_1 = 0 \text{ and } X_2 = 0 \text{ and } \ldots \text{ and } X_\Delta = 0).
\end{aligned}$$

Since all X_j's are mutually independent,

$$\Pr(v \in V_0 | v \in W_i) = \prod_{j=1}^{\Delta} \Pr(X_j = 0).$$

Now we calculate $\Pr(X_j = 0)$. Note that X_j is a binomial random variable with parameters $n = |W_j| = \frac{e^\alpha}{j^\beta}$ and $p = p_{ij} = \frac{ij}{e^\alpha \zeta(\beta - 1)}$. Therefore

$$\Pr(X_j = 0) = \binom{n}{0} p_{ij}^0 (1 - p_{ij})^n = q_{ij}^{|W_j|}$$

where $q_{ij} = 1 - p_{ij}$. Hence

$$\Pr(v \in V_0 | v \in W_i) = \prod_{j=1}^{\Delta} q_{ij}^{|W_j|} \approx \frac{1}{e^i},$$

where in the last approximation we use Lemma 3.1. □

Lemma 3.4. *Let $v \in W_i$. Then*

$$\Pr(v \in V_1 | v \in W_i) \approx \frac{i}{e^i}.$$

Proof. Let $v \in W_i$ and let X be the random variable for $d(v)$. Let X_j be the random variable counting the number of neighbors of v in W_j. Therefore $X = \sum_{j=1}^{\Delta} X_j$. Hence,

$$\begin{aligned}
\Pr(v \in V_1 | v \in W_i) &= \Pr(X = 1) \\
&= \Pr(X_1 + X_2 + \ldots + X_\Delta = 1) \\
&= \Pr(X_1 = 1 \text{ and } X_j = 0, \forall j \neq 1) \\
&\quad + \Pr(X_2 = 1 \text{ and } X_j = 0, \forall j \neq 2) \\
&\quad + \\
&\quad \vdots \\
&\quad + \Pr(X_\Delta = 1 \text{ and } X_j = 0, \forall j \neq \Delta).
\end{aligned}$$

Since all variables X_j's are mutually independent, we have

$$\begin{aligned}\Pr(v \in V_1 | v \in W_i) = \Pr(X_1 = 1) \prod_{j \neq 1} \Pr(X_j = 0) \\ + \Pr(X_2 = 1) \prod_{j \neq 2} \Pr(X_j = 0) \\ + \\ \vdots \\ + \Pr(X_\Delta = 1) \prod_{j \neq \Delta} \Pr(X_j = 0).\end{aligned}$$

Now we calculate $\Pr(X_j = 0)$ and $\Pr(X_j = 1)$. Note that X_j is a binomial random variable with parameters $n = |W_j| = \frac{e^\alpha}{j^\beta}$ and $p = p_{ij} = \frac{ij}{e^\alpha \zeta(\beta-1)}$. Therefore

$$\begin{aligned}\Pr(X_j = 0) &= \binom{n}{0} p_{ij}^0 (1 - p_{ij})^n = q_{ij}^{|W_j|} \\ \Pr(X_j = 1) &= \binom{n}{1} p_{ij}^1 (1 - p_{ij})^{n-1} = |W_j| p_{ij} q_{ij}^{|W_j|-1} = |W_j| \frac{p_{ij}}{q_{ij}} q_{ij}^{|W_j|},\end{aligned}$$

where $q_{ij} = 1 - p_{ij}$. Therefore,

$$\begin{aligned}\Pr(v \in V_1 | v \in W_i) &= |W_1| \frac{p_{i1}}{q_{i1}} q_{i1}^{|W_1|} \cdot q_{i2}^{|W_2|} \cdot q_{i3}^{|W_3|} \cdot \ldots \cdot q_{i\Delta}^{|W_\Delta|} \\ &+ q_{i1}^{|W_1|} \cdot |W_2| \frac{p_{i2}}{q_{i2}} q_{i2}^{|W_2|} \cdot q_{i3}^{|W_3|} \cdot \ldots \cdot q_{i\Delta}^{|W_\Delta|} \\ &+ q_{i1}^{|W_1|} \cdot q_{i2}^{|W_2|} \cdot |W_3| \frac{p_{i3}}{q_{i3}} q_{i3}^{|W_3|} \cdot \ldots \cdot q_{i\Delta}^{|W_\Delta|} \\ &\vdots \\ &+ q_{i1}^{|W_1|} \cdot q_{i2}^{|W_2|} \cdot q_{i3}^{|W_3|} \cdot \ldots \cdot |W_\Delta| \frac{p_{i\Delta}}{q_{i\Delta}} q_{i\Delta}^{|W_\Delta|} \\ &= \prod_{k=1}^{\Delta} \left(q_{ik}^{|W_k|} \right) \left(\sum_{j=1}^{\Delta} |W_j| \frac{p_{ij}}{q_{ij}} \right).\end{aligned}$$

Now we bound the product and the sum separately. Let us first look at the sum:

$$\begin{aligned}\sum_{j=1}^{\Delta} |W_j| \frac{p_{ij}}{q_{ij}} &= \sum_{j=1}^{\Delta} \frac{e^\alpha}{j^\beta} \frac{\frac{ij}{e^\alpha \zeta(\beta-1)}}{1 - \frac{ij}{e^\alpha \zeta(\beta-1)}} \\ &= e^\alpha \sum_{j=1}^{\Delta} \frac{1}{j^\beta} \frac{ij}{e^\alpha \zeta(\beta-1) - ij} \\ &\approx e^\alpha \sum_{j=1}^{\Delta} \frac{1}{j^\beta} \frac{ij}{e^\alpha \zeta(\beta-1)} \\ &= \frac{i}{\zeta(\beta-1)} \sum_{j=1}^{\Delta} \frac{1}{j^{\beta-1}} \\ &\approx \frac{i}{\zeta(\beta-1)} \zeta(\beta-1) \\ &= i,\end{aligned}$$

where the first approximation is analogous to Lemma 2.1 (we omit the proof which is almost identical to the proof of Lemma 2.1). Now, using Lemma 3.1, we deal with the product

$$\prod_{k=1}^{\Delta} q_{ik}^{|W_k|} \approx \frac{1}{e^i}.$$

Therefore,

$$\Pr(v \in V_1 | v \in W_i) = \prod_{k=1}^{\Delta} \left(q_{ik}^{|W_k|} \right) \left(\sum_{j=1}^{\Delta} |W_j| \frac{p_{ij}}{q_{ij}} \right) \approx \frac{i}{e^i}.$$

□

In the next Lemma we need the special function *polylogarithm*, defined as $\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}$

Lemma 3.5. *Let $v \in V$. Then*

$$\Pr(v \in V_0) \approx \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)}$$

Proof.

$$\begin{aligned}
\Pr(v \in V_0) &= \sum_{i=1}^{\Delta} \Pr(v \in V_0 | v \in W_i) \Pr(v \in W_i) \\
&\approx \sum_{i=1}^{\Delta} \frac{1}{e^i} \frac{1}{i^\beta \zeta(\beta)} \\
&= \frac{1}{\zeta(\beta)} \sum_{i=1}^{\Delta} \frac{1}{e^i i^\beta} \\
&= \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)},
\end{aligned}$$

where in the last approximation we use Lemmas 3.3 and 3.2. □

The proof of Lemma 3.6 below is omitted since is similar the proof of Lemma 3.5. The key step is to use Lemma 3.4 instead of Lemma 3.3.

Lemma 3.6. *Let $v \in V$. Then*

$$\Pr(v \in V_1) \approx \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}$$

In the following Lemmas we need to take in consideration the weights of the vertices in V^- and V_1, so we partition these sets according to Definition 3.7.

Definition 3.7. *Let*

$$\begin{aligned}
W_i^- &= \{v \in V | v \in W_i \ \textit{and} \ v \in V^-\} \\
W_i^{(1)} &= \{v \in V | v \in W_i \ \textit{and} \ v \in V_1\}
\end{aligned}$$

Lemma 3.8. *Let $v \in W_i$. Then*

$$\Pr(v \in W_i^- | v \in W_i) \approx 1 - \frac{(i+1)}{e^i}.$$

Proof. Let $v \in W_i$ and let X be the random variable for $d(v)$. Let X_j be the random variable counting the number of neighbors of v in W_j. Therefore $X = \sum_{j=1}^{\Delta} X_j$. Hence

$$\begin{aligned}\Pr(v \in V^-) &= \Pr(X \geq 2)\\ &= 1 - \Pr(X = 0 \text{ or } X = 1)\\ &= 1 - (\Pr(X = 0) + \Pr(X = 1) - \Pr(X = 0 \text{ and } X = 1))\end{aligned}$$

Note that $\Pr(X = 0 \text{ and } X = 1) = 0$, since X cannot assume both values at the same time. Therefore,

$$\begin{aligned}\Pr(v \in V^-) &= 1 - \Pr(X = 0) - \Pr(X = 1)\\ &= 1 - \Pr(v \in V_0 | v \in W_i) - \Pr(v \in V_1 | v \in W_i)\\ &\approx 1 - \frac{1}{e^i} - \frac{i}{e^i},\end{aligned}$$

where the approximation is obtained from Lemmas 3.3 and 3.4. ☐

Lemma 3.9.

$$\Pr(|W_i^{(1)}| = k) \approx \binom{|W_i|}{k} \left(\frac{i}{e^i}\right)^k \left(1 - \frac{i}{e^i}\right)^{|W_i|-k}$$

Proof. Directly from the fact that $|W_i^{(1)}|$ is a binomial random variable with parameters $n = |W_i|$ and $p = \Pr(v \in W_i^{(1)} | v \in W_i) \approx \frac{i}{e^i}$, where the approximation is obtained from Lemma 3.4. ☐

In the next Lemma, denote the event "$v \in W_i^-$ does not have a neighbor in $W_j^{(1)}$" by $v \nrightarrow W_j^{(1)}$.

Lemma 3.10.

$$\Pr(v \nrightarrow W_j^{(1)}) \lessapprox \left(\frac{1}{e}\right)^{\overline{\frac{i}{e^j j^{\beta-2} \zeta(\beta-1)}}}.$$

Proof. Using the Law of Total Probability and Lemma 3.9,

$$\begin{aligned}\Pr(v \nrightarrow W_j^{(1)}) &= \sum_{k=1}^{\Delta} \Pr(v \nrightarrow W_j^{(1)} \Big| |W_j^{(1)}| = k) \cdot \Pr(|W_j^{(1)}| = k)\\ &\approx \sum_{k=1}^{\Delta} (1 - p_{ij})^k \cdot \binom{|W_j|}{k} \left(\frac{j}{e^j}\right)^k \left(1 - \frac{j}{e^j}\right)^{|W_j|-k}\\ &= \sum_{k=1}^{\Delta} \underbrace{\binom{|W_j|}{k} \left((1 - p_{ij})\frac{j}{e^j}\right)^k \left(1 - \frac{j}{e^j}\right)^{|W_j|-k}}_{(\bigstar)}.\end{aligned}$$

If $\Delta \leq |W_j|$, then $\sum_{k=1}^{\Delta} (\bigstar) \leq \sum_{k=1}^{|W_j|} (\bigstar)$. Otherwise (i.e., $\Delta > |W_j|$), the every term $\binom{|W_j|}{k}$ is equal to 0 when $k > |W_j|$ and hence, $\sum_{k=1}^{\Delta} (\bigstar) = \sum_{k=1}^{|W_j|} (\bigstar)$. Therefore,

$$\Pr(v \nrightarrow W_j^{(1)}) \lessapprox \sum_{k=1}^{|W_j|} \binom{|W_j|}{k} \left((1 - p_{ij})\frac{j}{e^j}\right)^k \left(1 - \frac{j}{e^j}\right)^{|W_j|-k}.$$

Using the Binomial Theorem,

$$\begin{aligned}
\Pr(v \nrightarrow W_j^{(1)}) &\lessapprox \left((1-p_{ij})\frac{j}{e^j} + 1 - \frac{j}{e^j}\right)^{|W_j|} \\
&= \left(1 - p_{ij}\frac{j}{e^j}\right)^{|W_j|} \\
&= \left(1 - \frac{ij}{e^\alpha \zeta(\beta-1)}\frac{j}{e^j}\right)^{\frac{e^\alpha}{j^\beta}} \\
&= \left(1 - \frac{\frac{ij^2}{e^j\zeta(\beta-1)}}{e^\alpha}\right)^{e^\alpha \cdot \frac{1}{j^\beta}} \\
&\approx \left(\frac{1}{e^{\frac{ij^2}{e^j\zeta(\beta-1)}}}\right)^{\frac{1}{j^\beta}} \\
&= \left(\frac{1}{e}\right)^{\frac{i}{e^j j^{\beta-2}\zeta(\beta-1)}}
\end{aligned}$$

□

Lemma 3.11. $\Pr(v$ *does not have a neighbor in* $V_1) \lessapprox \left(\frac{1}{e^i}\right)^{\rho(\beta)}$, *where* $\rho(\beta) = \frac{Li_{\beta-2}(\frac{1}{e})}{\zeta(\beta-1)}$.

Proof. Let ε_j be the event "$v \in W_i^-$ does not have a neighbor in $W_j^{(1)}$". Therefore

$$\begin{aligned}
\Pr(v \text{ does not have a neighbor in } V_1) &= \Pr(\varepsilon_1 \cap \varepsilon_2 \cap \ldots \cap \varepsilon_\Delta) \\
&= \prod_{j=1}^{\Delta} \Pr(\varepsilon_j) \\
&\lessapprox \prod_{j=1}^{\Delta} \left(\frac{1}{e}\right)^{\frac{i}{e^j j^{\beta-2}\zeta(\beta-1)}},
\end{aligned}$$

where the second step is a consequence of the independence of the events ε_j and the upper bound is consequence of Lemma 3.10. Therefore,

$$\begin{aligned}
\Pr(v \text{ does not have a neighbor in } V_1) &\lessapprox \prod_{j=1}^{\Delta} \left(\frac{1}{e}\right)^{\frac{i}{e^j j^{\beta-2}\zeta(\beta-1)}} \\
&= \left(\frac{1}{e}\right)^{\sum_{j=1}^{\Delta} \frac{i}{e^j j^{\beta-2}\zeta(\beta-1)}} \\
&= \left(\frac{1}{e}\right)^{\frac{i}{\zeta(\beta-1)} \sum_{j=1}^{\Delta} \frac{1}{e^j j^{\beta-2}}} \\
&= \left(\frac{1}{e}\right)^{\frac{i}{\zeta(\beta-1)} \mathrm{Li}_{\beta-2}(\frac{1}{e})}
\end{aligned}$$

□

In the next Lemma, denote the event "$v \in V^-$ does not have a neighbor in V_1" by $v \nrightarrow V_1$.

Lemma 3.12.

$$\Pr(v \nrightarrow V_1) \lessapprox \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)},$$

where $\rho(\beta) = \frac{Li_{\beta-2}(\frac{1}{e})}{\zeta(\beta-1)}$.

Proof. Using the Law of Total Probability,

$$\begin{aligned}
\Pr(v \nrightarrow V_1) &= \sum_{i=1}^{\Delta} \Pr(v \nrightarrow V_1 \text{ and } v \in W_i^-) \\
&= \sum_{i=1}^{\Delta} \Pr(v \nrightarrow V_1 | v \in W_i^-)\Pr(v \in W_i^-) \\
&= \sum_{i=1}^{\Delta} \Pr(v \nrightarrow V_1 | v \in W_i^-)\Pr(v \in W_i^- | v \in W_i)\Pr(v \in W_i).
\end{aligned}$$

Using Lemmas 3.11, 3.8 and 3.2, we have

$$\begin{aligned}
\Pr(v \nrightarrow V_1) &= \sum_{i=1}^{\Delta} \left(\frac{1}{e^i}\right)^{\rho(\beta)} (1 - 1/e^i - i/e^i) \left(\frac{1}{i^\beta \zeta(\beta)}\right) \\
&\lessapprox \frac{1}{\zeta(\beta)} \sum_{i=1}^{\Delta} \frac{1}{e^{i\rho(\beta)} i^\beta} \\
&= \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}.
\end{aligned}$$

□

In the next Theorem we estimate the size of the set of vertices adjacent to vertices of degree 1.

Theorem 3.13.

$$\mathrm{E}[|N(V_1)|] \gtrapprox |V| \left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right)\left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right)$$

Proof. We denote the event "$v \in V$ has a neighbor in V_1" by $v \to V_1$. Let X_v be the indicator random variable where $X_v = 1$ when $v \to V_1$ **and** $v \in V^-$ (otherwise $X_v = 0$). Therefore, by Lemmas 3.12, 3.5 and 3.6,

$$\begin{aligned}
\mathrm{E}[X_v] &= \Pr(X_v = 1) \\
&= \Pr(v \to V_1 \cap v \in V^-) \\
&= \Pr(v \to V_1 | v \in V^-) \cdot \Pr(v \in V^-) \\
&= \Pr(v \to V_1 | v \in V^-) \cdot \Pr(v \in V \setminus \{V_0 \cup V_1\}) \\
&\gtrapprox \left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right)\left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right).
\end{aligned}$$

Let X be a random variable such that $X = \sum_{v \in V} X_v$ (i.e., $X = |N(V_1)|$). Therefore, by the linearity of expectation and using the value of $\mathrm{E}[X_v]$,

$$\begin{aligned}
\mathrm{E}[|N(V_1)|] &= \sum_{v \in V} \mathrm{E}[X_v] \\
&\gtrapprox \sum_{v \in V} \left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right)\left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right) \\
&= |V| \left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right)\left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right)
\end{aligned}$$

□

4. Approximation Algorithm

The first step of the algorithm consists of including every vertex of $N(V_1)$ in the solution. In Lemma 4.1 we show that the optimal solution can be split into two parts and the optimal solution in one of these parts is $N(V_1)$. The optimal solution of the other part is bounded in Lemma 4.3.

For any $S \subseteq V$, let $\mathrm{OPT}(S)$ be the cardinality of a minimum vertex cover in $G[S]$. Let $V^* = V_1 \cup N(V_1)$. In Lemma 4.1, we give an auxiliary result that holds for any graph $G = (V, E)$ (i.e., no probabilistic argument is used in the proof). In this particular Lemma, quantities such as $\mathrm{OPT}(S)$ and $|N(V_1)|$ are not treated as expected values of random variables. In the rest of the paper these quantities are treated again as expected values of random variables.

Lemma 4.1. *The following three conditions hold:*

(i) G contains a minimum vertex cover C such that $N(V_1) \subseteq C$ and no vertex of V_1 is in C;

(ii) $\mathrm{OPT}(V^*) = |N(V_1)|$;

(iii) $\mathrm{OPT}(V) = \mathrm{OPT}(V^*) + \mathrm{OPT}(V \setminus V^*)$.

Proof. Let $uv \in E$ such that $u \in V_1$. Note that any minimum vertex cover C of G contains exactly one vertex of $\{u, v\}$ (otherwise C is not a cover or C is not minimum). Note that if $u \in C$, then $(C \setminus \{u\}) \cup \{v\}$ is also a minimum vertex cover. Therefore, using this same exchange argument, there is a minimum vertex cover containing every vertex of $N(V_1)$. Hence (i) holds. Applying (i) to the graph induced by V^*, $N(V_1)$ is a minimum vertex cover of $G[V^*]$, and hence (ii) holds. Now let $C = C' \cup N(V_1)$ be a minimum vertex cover respecting condition (i). There is no vertex cover C'' in $G[V \setminus V^*]$ such that $|C''| < |C'|$, otherwise $C'' \cup N(V_1)$ contradicts the minimality of C. Therefore $\mathrm{OPT}(V \setminus V^*) = |C'|$. Combining this with (ii), condition (iii) holds. □

Lemma 4.2.

$$\frac{\mathrm{OPT}(V^*)}{\mathrm{OPT}(V)} \geq \left(1 - \frac{\mathrm{Li}(e^{-\rho(\beta)})}{\zeta(\beta)}\right).$$

Proof. By Lemma 4.1 (i), $\mathrm{OPT}(V) \leq |V^-|$. Combining this with Lemmas 3.5 and 3.6 we have $\mathrm{OPT}(V) \leq |V^-| = |V| \left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right)$. By Lemma 4.1 (ii) and Theorem 3.13,

$$\mathrm{OPT}(V^*) = |N(V_1)| \geq |V| \left(1 - \frac{\mathrm{Li}_\beta(1/e)}{\zeta(\beta)} - \frac{\mathrm{Li}_{\beta-1}(1/e)}{\zeta(\beta)}\right)\left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right).$$

Combining these bounds we have

$$\frac{\mathrm{OPT}(V^*)}{\mathrm{OPT}(V)} \geq \left(1 - \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right).$$

□

Corollary 4.3.

$$\frac{\mathrm{OPT}(V \setminus V^*)}{\mathrm{OPT}(V)} \leq \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}.$$

Proof. By Lemma 4.1 (iii), $\frac{\mathrm{OPT}(V\setminus V^*)+\mathrm{OPT}(V^*)}{\mathrm{OPT}(V)} = 1$. By Lemma 4.2, the corollary holds. □

Algorithm 4.4.

1. *Let C be a cover obtained by a 2-approximation algorithm in $G[V \setminus V^*]$.*
2. *Output $C \cup N(V_1)$.*

Theorem 4.5. *The expected approximation factor of Algorithm 4.4 is $\left(1 + \frac{\mathrm{Li}(e^{-\rho(\beta)})}{\zeta(\beta)}\right)$ when applied to a random power law graph.*

Proof. The size of a solution obtained by Algorithm 4.4 is $|C \cup N(V_1)|$. Since C and $N(V_1)$ are disjoint, $|C \cup N(V_1)| = |C| + |N(V_1)|$. By Lemma 4.1 (ii), $|N(V_1)| = \mathrm{OPT}(V^*)$. Since C is obtained by a 2-approximation algorithm in $V \setminus V^*$, $|C| \leq 2\mathrm{OPT}(V \setminus V^*)$. Therefore

$$|C \cup N(V_1)| \leq \mathrm{OPT}(V^*) + 2\mathrm{OPT}(V \setminus V^*) = \mathrm{OPT}(V) + \mathrm{OPT}(V \setminus V^*),$$

where the last step is obtained by Lemma 4.1 (iii). By Corollary 4.3,

$$\begin{aligned} |C \cup N(V_1)| &= \mathrm{OPT}(V) + \mathrm{OPT}(V \setminus V^*) \\ &\leq \mathrm{OPT}(V) + \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)} \mathrm{OPT}(V) \\ &= \left(1 + \frac{\mathrm{Li}_\beta(e^{-\rho(\beta)})}{\zeta(\beta)}\right) \mathrm{OPT}(V). \end{aligned}$$

□

5. Conclusion

In this paper we study the approximability of the minimum vertex cover problem in power law graphs. We use the the *generalized random graph* [27] model (as well the Chung-Lu model) to analyze the expected approximation ratio of the algorithm. We observe that this graph model is potentially easier to handle than the model used in [14] mainly due to the fact that edges are added independently in the random process.

A question that might arise is how much the approximation factor can be improved by repeating the first step of the algorithm recursively for the residual graph until no more vertices of degree one are left. Even though this idea seems natural, the analysis appears to be more complicated, since the recursive steps give rises to probabilities that are not independent. Also, one might hope to prove a concentration inequality for the probability that the algorithm returns a good solution.

As a future work we are also interested in investigating the approximability of other optimization problems on power law graphs.

[1] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the internet topology, SIGCOMM Comput. Commun. Rev. 29 (4) (1999) 251–262.

[2] A.-L. Barabsi, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.

[3] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, A. S. Tomkins, The web as a graph: Measurements, models, and methods, in: Proceedings of the 5th Annual International Conference on Computing and Combinatorics, COCOON'99, Springer-Verlag, Berlin, Heidelberg, 1999, pp. 1–17.

[4] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Graph structure in the web, Comput. Netw. 33 (1-6) (2000) 309–320.

[5] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, E. Upfal, Stochastic models for the web graph, in: Proceedings of the 41st Annual Symposium on Foundations of Computer Science, FOCS '00, IEEE Computer Society, Washington, DC, USA, 2000, pp. 57–.

[6] J. Kleinberg, S. Laurence, The structure of the web, Science 294 (5548) (2001) 1849–1850.

[7] N. Guelzim, S. Bottani, P. Bourgine, F. Képès, Topological and causal structure of the yeast transcriptional regulatory network, Nature genetics 31 (1) (2002) 60–63.

[8] G. Siganos, M. Faloutsos, P. Faloutsos, C. Faloutsos, Power laws and the as-level internet topology, IEEE/ACM Trans. Netw. 11 (4) (2003) 514–524.

[9] S. Eubank, V. S. Anil, K. Madhav, V. Marathe, A. Srinivasan, N. Wang, Structural and algorithmic aspects of massive social networks, in: in Proceedings of 15th ACM-SIAM Symposium on Discrete Algorithms (SODA 2004), 711-720, SIAM, 2004.

[10] K. Park, H. Lee, On the effectiveness of route-based packet filtering for distributed dos attack prevention in power-law internets, SIGCOMM Comput. Commun. Rev. 31 (4) (2001) 15–26.

[11] C. Gkantsidis, M. Mihail, A. Saberi, Conductance and congestion in power law graphs, in: Proceedings of the 2003 ACM SIGMETRICS, ACM, New York, NY, USA, 2003, pp. 148–159.

[12] M. O. da Silva, G. A. Gimenez-Lugo, M. V. G. da Silva, Vertex cover in complex networks, International Journal of Modern Physics C 24 (11).

[13] E. D. Demaine, F. Reidl, P. Rossmanith, F. S. Villaamil, S. Sikdar, B. D. Sullivan, Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs, Manuscript.

[14] M. Gast, M. Hauptmann, Approximability of the vertex cover problem in power-law graphs, Theor. Comput. Sci. 516 (2014) 60–70.

[15] M. Gast, M. Hauptmann, M. Karpinski, Inapproximability of dominating set on power law graphs, Theoretical Computer Science 562 436–452.

[16] M. Gast, M. Hauptmann, M. Karpinski, Improved approximation lower bounds for vertex cover on power law graphs and some generalizations, Computing Research Repository (CoRR) arXiv:1210.2698.

[17] Y. Shen, D. T. Nguyen, Y. Xuan, M. T. Thai, New techniques for approximating optimal substructure problems in power-law graphs, Theoretical Computer Science 447 107–119.

[18] E. A. Bender, The asymptotic number of labeled graphs with given degree sequences, Journal of Combinatorial Theory A. 24 296–307.

[19] N. Wormald, Some problems in the enumeration of labelled graphs, Doctoral Thesis, Dept. of Math., Univ. Newcastle, Australia.

[20] B. Bollobás, Random graphs, Academic Press, 1985.

[21] M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence, Random Structures and Algorithms 6 (2-3) (1995) 161–180.

[22] M. Molloy, B. Reed, The size of the giant component of a random graph with a given degree sequence, Comb. Probab. Comput. 7 (3) (1998) 295–305.

[23] W. Aiello, F. Chung, L. Lu, A random graph model for power law graphs, Experimental Mathematics 10 (2001) 53–66.

[24] B. Bollobás, O. Riordan, J. Spencer, G. Tusnády, The degree sequence of a scale-free random graph process, Random Structures and Algorithms 18 (3) (2001) 279–290.

[25] F. Chung, L. Lu, Connected components in random graphs with given expected degree sequences, Annals of Combinatorics 6 (2) (2002) 125–145.

[26] F. Chung, L. Lu, The average distance in a random graph with given expected degrees, PNAS 99 (25) (2002) 15879–15882.

[27] T. Britton, M. Deijfen, A. Martin-Lof, Generating simple random graphs with prescribed degree distribution, Journal of Statistical Physics 124 (6) (2006) 1377–1397.

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

[29] P. Erdös, A. Rényi, On random graphs, i, Publicationes Mathematicae (Debrecen) 6 (1959) 290–297.

[30] E. N. Gilbert, Random graphs, The Annals of Mathematical Statistics 30 (4) (1959) 1141–1144.

[31] F. Chung, L. Lu, Connected components in random graphs with given expected degree sequences, Annals of Combinatorics 6 (2) (2002) 125–145.

[32] R. Hofstad, Random graphs and complex networks, Manuscript.

[33] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman & Co., New York, NY, USA, 1979.

[34] S. Khot, O. Regev, Vertex cover might be hard to approximate to within 2-ϵ, J. Comput. Syst. Sci. 74 (3) (2008) 335–349.

[35] T. H. Cormen, C. Leiserson, R. Rivest, C. Stein, Introduction to Algorithms, Third Edition, 3rd Edition, The MIT Press, 2009.