On the Hardness and Inapproximability of Optimization Problems on Power Law Graphs

Yilin Shen, Dung T. Nguyen, and My T. Thai

Department of Computer Information Science and Engineering
University of Florida, Gainesville, FL, 32611
{yshen,dtnguyen,mythai}@cise.ufl.edu

Abstract. The discovery of power law distribution in degree sequence (i.e. the number of vertices with degree i is proportional to $i^{-\beta}$ for some constant β) of many large-scale real networks creates a belief that it may be easier to solve many optimization problems in such networks. Our works focus on the hardness and inapproximability of optimization problems on power law graphs (PLG). In this paper, we show that the Minimum Dominating Set, Minimum Vertex Cover and Maximum Independent Set are still *APX*-hard on power law graphs. We further show the inapproximability factors of these optimization problems and a more general problem (ρ-Minimum Dominating Set), which proved that a belief of $(1 + o(1))$-approximation algorithm for these problems on power law graphs is not always true. In order to show the above theoretical results, we propose a general cycle-based embedding technique to embed any d-bounded graphs into a power law graph. In addition, we present a brief description of the relationship between the exponential factor β and constant greedy approximation algorithms.

Keywords: Theory, Complexity, Inapproximability, Power Law Graphs.

1 Introduction

In real life, the remarkable discovery shows that many large-scale networks follow a power law distribution in their degree sequences, ranging from biological networks, the Internet, the WWW to social networks [19] [20]. That is, the number of vertices with degree i is proportional to $i^{-\beta}$ for some constant β in these graphs, which is called power law graphs (PLG). The observations show that the exponential factor β ranges between 1 and 4 for most real-world networks [8]. Intuitively, the following theoretical question is raised: What are the differences in terms of complexity and inapproxamability of several optimization problems between on general graphs and on PLG?

Many experimental results on random power law graphs give us a belief that the problems might be much easier to solve on PLG. Eubank *et al.* [12] claimed that a simple greedy algorithm leads to a $1 + o(1)$ approximation ratio on Minimum Dominating Set (MDS) problem (without any formal proof) although MDS has been proved *NP*-hard to be approximated within $(1 - \epsilon) \log n$ unless

W. Wu and O. Daescu (Eds.): COCOA 2010, Part I, LNCS 6508, pp. 197–211, 2010.

NP=ZPP. The approximating result on Minimum Vertex Cover (MVC) was also much better than the 1.366-inapproximability on general graphs [10]. In [22], Gopal claimed that there exists a polynomial time algorithm that guarantees a $1 + o(1)$ approximation of the MVC problem with probability at least $1 - o(1)$. However, there is no such formal proof for this claim either. Furthermore, several papers also have some theoretical guarantees for some problems on PLG. Gkantsidis *et al.* [14] proved the flow through each link is at most $O(n \log^2 n)$ on power law random graphs (PLRG) where the routing of $O(d_u d_v)$ units of flow between each pair of vertices u and v with degrees d_u and d_v. In [14], the authors take advantage of the property of power law distribution by using the structural random model [1],[2] and show the theoretical upper bound with high probability $1 - o(1)$ and the corresponding experimental results. Likewise, Janson *et al.* [16] gave an algorithm that approximated Maximum Clique within $1 - o(1)$ on PLG with high probability on the random poisson model $G(n, \alpha)$ (i.e. the number of vertices with degree at least i decreases roughly as n^{-i}). Although these results were based on experiments and random models, it raises an interest in investigating hardness and inapproximability of classical optimization problems on PLG.

Recently, Ferrante *et al.* [13] had an initial attempt to show that MVC, MDS and Maximum Independent Set (MIS) ($\beta > 0$), Maximum Clique (Clique) and Minimum Graph Coloring (Coloring) ($\beta > 1$) still remain *NP*-hard on PLG. Unfortunately, there is a minor error in the proof of their Lemma 5 which makes the proof of *NP*-hardness of MIS, MVC, MDS with $\beta < 1$ no longer hold. Indeed, it is not trivial to fix that error and thus we present in Appendix A another way to show the *NP*-hardness of these problems when $\beta < 1$.

Our Contributions: In this paper, we show the *APX*-hardness and the inapproximability of MIS, MDS, and MVC according to a general *Cycle-Based Embedding Technique* which embeds any d-bounded graph into a power law graph with the exponential factor β. The inapproximability results of the above problems on PLG are shown in Table 1 with some constant c_1, c_2 and c_3. Then, the further inapproximability results on Clique and Coloring are shown by taking advantage of the reduction in [13]. We also analyze the relationship between β and constant greedy approximation algorithms for MIS and MDS.

In addition, recent studies on social networks have led to a new problem of spreading the influence through a social network [18] [17] by initially influencing a minimum small number of people. By formulating this problem as *ρ-Minimum Dominating Set* (ρ-MDS), we show that ρ-MDS is Unique Game-hard to be approximated within $2 - (2 + o_d(1)) \log \log d / \log d$ factor on d-bounded graphs and further leading to the following inapproximability result on PLG (shown in Table 1).

Organization: In Section 2, we introduce some problem definitions, the model of PLG, and corresponding concepts. In Section 3, the general embedding technique are introduced by which we can use to show the hardness and inapproximability of MIS, MDS, MVC in Section 4 and Section 5 respectively. In addition, the

Table 1. Inapproximability Results on Power Law Graph with Exponential Factor β

Problem	Inapproximability Factor	Condition
MDS	$1 + 2\left(\log c_3 - O(\log\log c_3) - 1\right)/((c_3+1)\zeta(\beta))$	NP$\not\subseteq$DTIME$\left(n^{O(\log\log n)}\right)$
MIS	$1 - 2\left(c_1 - O\left(\log^2 c_1\right)\right)/(c_1(c_1+1)\zeta(\beta))$	Unique Game Conjecture
MVC	$1 + 2\left(1 - (2 + o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)/((c_2+1)\zeta(\beta))$	Unique Game Conjecture
ρ-MDS	$1 + \left(1 - (2 + o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)/((c_2+1)\zeta(\beta))$	Unique Game Conjecture
Clique	$O\left(n^{1/(\beta+1)-\epsilon}\right)$	NP$\neq$ZPP
Coloring	$O\left(n^{1/(\beta+1)-\epsilon}\right)$	NP$\neq$ZPP

inapproximability result of Clique and Coloring are also shown in Section 5. In Section 6, we analyze the relationship between β and constant approximation algorithms, which further proves that the integral gap is typically small for optimization problems on PLG than that on general bounded graphs. We fix the *NP*-hardness proof for $\beta < 1$ presented in [13] in Appendix A.

2 Preliminaries

This section provides several parts. First, we recall the definition of the new optimization problem ρ-Minimum Dominating Set. Next, the power law model and some corresponding concepts are proposed. Finally, we introduce some special graphs which will be used in the analysis throughout the paper.

2.1 Problem Definitions

The ρ-Minimum Dominating Set is defined as general version of MDS problem. In the context of influence spreading, the ρ-MDS problem says that given a graph modeling a social network, where each vertex v has a fix threshold $\rho|N(v)|$ such that the vertex v will adopt a new product if $\rho|N(v)|$ of its neighbors adopt it. Thus our goal is to find a small set DS of vertices such that targeting the product to DS would lead to adoption of the product by a large number of vertices in the graph in t propagations. To be simplified, we define ρ-MDS problem in the case that $t = 1$.

Definition 1 (ρ-Minimum Dominating Set). *Given an undirected graph $G = (V, E)$, find a subset $DS \subseteq V$ with the minimum size such that for each vertex $v_i \in V \setminus DS$, $|DS \cap N(v_i)| \geq \rho|N(v_i)|$, where $0 < \rho \leq 1/2$.*

2.2 Power Law Model and Concepts

A great number of models [5] [6] [1] [2] [21] on power law graphs are emerging in the past recent years. In this paper, we do the analysis based on the general (α, β) model, that is, the graphs only constrained with the distribution on the number of vertices with different degrees.

Definition 2 ((α, β) Power Law Graph Model). *A graph $G_{(\alpha,\beta)} = (V, E)$ is called a (α, β) power law graph where multi-edges and self-loops are allowed if the maximum degree is $\Delta = \lfloor e^{\alpha/\beta} \rfloor$ and the number of vertices of degree i is:*

$$y_i = \begin{cases} \lfloor e^{\alpha}/i^{\beta} \rfloor, & \text{if } i > 1 \text{ or } \sum_{i=1}^{\Delta} \lfloor e^{\alpha}/i^{\beta} \rfloor \text{ is even} \\ \lfloor e^{\alpha} \rfloor + 1, & \text{otherwise} \end{cases} \quad (1)$$

Definition 3 (d-Bounded Graph). *Given a graph $G = (V, E)$, G is a d-bounded graph if the degree of any vertex is upper bounded by an integer d.*

Definition 4 (Degree Set). *Given a power law graph $G_{(\alpha,\beta)}$, let $D_i(G_{(\alpha,\beta)})$ be the set of vertices with degree i on graph $G_{(\alpha,\beta)}$.*

2.3 Special Graphs

Definition 5 (Cubic Cycle CC_n). *A cubic cycle CC_n is composed of two cycles. Each cycle has n vertices and two i^{th} vertices in each cycle are adjacent with each other. That is, Cubic Cycle CC_n has $2n$ vertices and each vertex has degree 3. An example CC_8 is shown in Figure 1.*

Then a Cubic Cycle CC_n can be extended into a $\boldsymbol{d}$-Regular Cycle RC_n^d with the given vector $\boldsymbol{d}$. The definition is as follows.

Definition 6 ($\boldsymbol{d}$-Regular Cycle $RC_n^{\boldsymbol{d}}$). *Give a vector $\boldsymbol{d} = (d_1, \ldots, d_n)$, a $\boldsymbol{d}$-Regular Cycle RC_n^d is composed of a two cycles. Each cycle has n vertices and two i^{th} vertices in each cycle are adjacent with each other by $d - 2$ multi-edges. That is, $\boldsymbol{d}$-Regular Cycle RC_n^d has $2n$ vertices and the two i^{th} vertex has degree d_i. An example RC_8^d is shown in Figure 3.*

Definition 7 ($\boldsymbol{d}$-Cycle $C_n^{\boldsymbol{d}}$). *Give a vector $\boldsymbol{d} = (d_1, \ldots, d_n)$, a $\boldsymbol{d}$-Cycle C_n^d is a cycle with a even number of vertices n such that each vertex has degree d_i with $(d_i - 2)/2$ self-loops. An example C_8^d is shown in Figure 4.*

Definition 8 ($\boldsymbol{\kappa}$-Branch-$\boldsymbol{d}$-Cycle $\boldsymbol{\kappa}$-$BC_n^{\boldsymbol{d}}$). *Given a $\boldsymbol{d}$-Cycle and a vector $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)$, the $\boldsymbol{\kappa}$-Branch-$\boldsymbol{d}$-Cycle is composed of $|\boldsymbol{\kappa}|/2$ branches appending C_n^d, where $|\kappa|$ is a even number. An example is shown in Figure 5.*

Fact 1. *$\boldsymbol{\kappa}$-Branch-$\boldsymbol{d}$-Cycle has $|\boldsymbol{\kappa}|$ even number of vertices with odd degrees.*

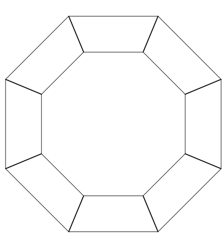
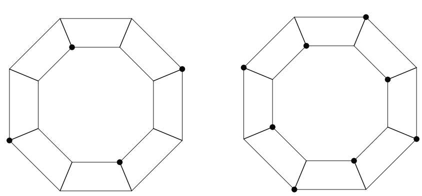


Fig. 1. CC_8



(a) MDS (b) MVC, MIS

Fig. 2. Solutions on CC_8

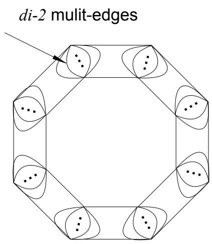


Fig. 3. RC_8^d

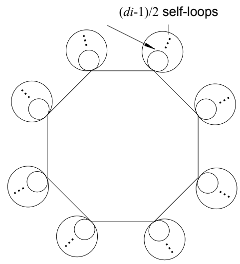


Fig. 4. C_8^d

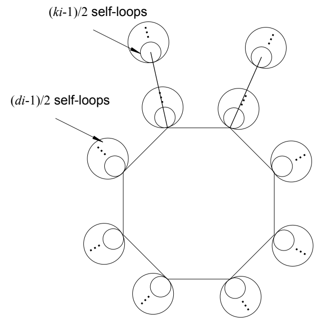


Fig. 5. $4\text{-}BC_6^d$

3 General Cycle-Based Embedding Technique

In this section, we present *General Cycle-Based Embedding Technique* on (α, β) power law graph model with $\beta > 1$. The idea on *Cycle-Based Embedding Technique* is to embed an arbitrary d-bounded graph into PLG with $\beta > 1$ with a $\boldsymbol{d_1}$-Regular Cycle, a $\boldsymbol{\kappa}$-Branch-$\boldsymbol{d_2}$-Cycle and a number of cliques K_2, where $\boldsymbol{d_1}$, $\boldsymbol{d_2}$ and $\boldsymbol{\kappa}$ are defined by α and β. Since the classical problems can be polynomially solved in both $\boldsymbol{d}$-Regular Cycles and $\boldsymbol{\kappa}$-Branch-$\boldsymbol{d}$-Cycle according to Corollary 1 and Lemma 2, Cycle-Based Embedding Technique helps to prove the complexity of such problem on PLG according to the complexity result of the same problem on bounded graphs.

Lemma 1. *MDS, MVC and MIS is polynomially solvable on Cubic Cycle.*

Proof. Here we just prove MDS problem is polynomially solvable on Cubic Cycle. The algorithm is simple. First we arbitrarily select a vertex, then select the vertex on the other cycle in two hops. The algorithm will terminate until all vertices are dominated. Now we will show that this gives the optimal solution. Let's take CC_8 as an example. As shown in Fig. 2(a), the size of MDS is 4. Notice that each node can dominate exact 3 vertices, that is, 4 vertices can dominate exactly 12 vertices. However, in CC_8, there are altogether 16 vertices, which have to be dominated by at least 4 vertices apart from the vertices in MDS. That is, the algorithm returns an optimal solution. Moreover, MVC and MIS can be proved similarly as shown in Fig. 2(b).

Corollary 1. *MDS, MVC and MIS is polynomially solvable on* $\boldsymbol{d}$*-Regular Cycle and* $\boldsymbol{d}$*-Cycle.*

Lemma 2. *MDS, MVC and MIS is polynomially solvable on* $\boldsymbol{\kappa}$*-Branch-*$\boldsymbol{d}$*-Cycle.*

Proof. Let us take the MDS as an example. First we select the vertices connecting both the branches and the cycle. Then by removing the branches, we will have a line graph regardless of self-loops, on which MDS is polynomially solvable. It is easy to see that the size of MDS will increase if any one vertex connecting both the branch and the cycle in MDS is replaced by some other vertices.

Theorem 1 (Cycle-Based Embedding Technique). *Any d-bounded graph G_d can be embedded into a power law graph $G_{(\alpha,\beta)}$ with $\beta > 1$ such that G_d is a maximal component and the above classical problems can be polynomially solvable on $G_{(\alpha,\beta)} \setminus G_d$.*

Proof. With the given β and $\tau(i) = \lfloor e^\alpha / i^\beta \rfloor - n_i$ where $n_i = 0$ when $i > d$, we construct the power law graph $G_{(\alpha,\beta)}$ as the following algorithm:

1. Choose a number α such that $e^\alpha = \max_{1 \le i \le d}\{n_i \cdot i^\beta\}$ and $e^{\alpha/\beta} \ge d$;
2. For the vertices with degree 1, add $\lfloor \tau(1)/2 \rfloor$ number of cliques K_2;
3. For $\tau(2)$ vertices with degree 2, add a cycle with the size $\tau(2)$;
4. For all vertices with degree larger than 2 and smaller than $\lfloor e^{\alpha/\beta} \rfloor$, construct a $\boldsymbol{d_1}$-Regular Cycle where $\boldsymbol{d_1}$ is a vector composed of $2\lfloor \tau(i)/2 \rfloor$ number of i elements for all i satisfying $\tau(i) > 0$;
5. For all leftover isolated vertices L such that $\tau(i) - 2\lfloor \tau(i)/2 \rfloor = 1$, construct a $\boldsymbol{d_2^1}$-Branch-$\boldsymbol{d_2^2}$-Cycle, where d_2^1 is a vector composed of the vertices in L with odd degrees and d_2^2 is a vector composed of the vertices in L with even degrees.

The last step holds since the number of vertices with odd degrees has to be even. Therefore, $e^\alpha = \max_{1 \le i \le d}\{n_i \cdot i^\beta\} \le n$, that is, the number of vertices in graph $G_{(\alpha,\beta)}$ $N = \zeta(\beta)n = \Theta(n)$ meaning that N/n is a constant. According to Corollary 1 and Lemma 2, since $G_{(\alpha,\beta)} \setminus G_d$ is composed of a $\boldsymbol{d_1}$-Regular Cycle and a k-Branch-$\boldsymbol{d_2}$-Cycle, it can be polynomially solvable.

4 Hardness of Optimization Problems on PLG

In this section, we prove that MIS, MDS, MVC are *APX*-hard on PLG.

Theorem 2 (Alimonti *et al.* [3]). *MDS is* APX-*hard on cubic graphs.*

Theorem 3. *MDS is* APX-*hard on PLG.*

Proof. According to Theorem 1, we use the *Cycle-Based Embedding Technique* to show $\mathcal{L}$-reduction from MDS on d-bounded graph G_d to MDS on power law graph $G_{(\alpha,\beta)}$. Let ϕ and φ be a feasible solution on G_d and $G_{(\alpha,\beta)}$ respectively.

We first consider MDS on different graphs. Notice that MDS on a K_2 is 1, $n/4$ on a $\boldsymbol{d}$-Regular Cycle according to Lemma 1 and $n/3$ on a cycle. Therefore, for a solution ϕ on G_d, we have a solution φ on $G_{(\alpha,\beta)}$ is $\varphi = \phi + n_1/2 + n_2/3 + n_3/4$, where n_1, n_2 and n_3 corresponds to $\tau(1)$, $\tau(2)$ and all leftover vertices in Theorem 1. Correspondingly, we have $OPT(\varphi) = OPT(\phi) + n_1/2 + n_2/3 + n_3/4$.

On one hand, for a d-bounded graph with vertices n, the optimal MDS is lower bounded by $n/(d+1)$. Thus, we know

$$\begin{aligned} &OPT(\varphi) = OPT(\phi) + n_1/2 + n_2/3 + n_3/4 \\ &\le OPT(\phi) + (N-n)/2 \le OPT(\phi) + (\zeta(\beta)-1)n/2 \\ &\le OPT(\phi) + (\zeta(\beta)-1)(d+1)OPT(\phi)/2 = [1 + (\zeta(\beta)-1)(d+1)/2]\, OPT(\phi) \end{aligned}$$

where N is the number of vertices in $G_{(\alpha,\beta)}$.

On the other hand, with $|OPT(\phi) - \phi| = |OPT(\varphi) - \varphi|$, we proved the $\mathcal{L}$-reduction with $c_1 = 1 + (\zeta(\beta)-1)(d+1)/2$ and $c_2 = 1$.

Theorem 4. *MVC is* APX-*hard on PLG.*

Proof. In this proof, we construct as *Cycle-Based Embedding Technique*, according to Theorem 1, to show $\mathcal{L}$-reduction from MVC on d-bounded graph G_d to MVC on power law graph $G_{(\alpha,\beta)}$. Let ϕ be a feasible solution on G_d and φ be a feasible solution on $G_{(\alpha,\beta)}$.

However, MVC on K_2, cycle, $\boldsymbol{d}$-Regular Cycle and $\boldsymbol{\kappa}$-Branch-$\boldsymbol{d}$-Cycle is $n/2$. Therefore, for a solution ϕ on G_d, we have a solution φ on $G_{(\alpha,\beta)}$ is $\varphi = \phi + (N-n)/2$. Correspondingly, we have $OPT(\varphi) = OPT(\phi) + (N-n)/2$.

On one hand, for a d-bounded graph with vertices n, the optimal MVC is lower bounded by $n/(d+1)$. Therefore, similarly as the proof in Theorem 3,

$$OPT(\varphi) \le [1 + (\zeta(\beta)-1)(d+1)/2]\, OPT(\phi)$$

On the other hand, with $|OPT(\phi) - \phi| = |OPT(\varphi) - \varphi|$, we proved the $\mathcal{L}$-reduction with $c_1 = 1 + (\zeta(\beta)-1)(d+1)/2$ and $c_2 = 1$.

Corollary 2. *MIS is* APX-*hard on PLG.*

5 Inapproximability of Optimization Problems on PLG

5.1 MDS, MIS, MVC

Theorem 5 (P. Austrin *et al.* [4]). *For every sufficiently large integer d, MIS on a graph d-bounded G is* UG-*hard to approximate within a factor* $O\left(d/\log^2 d\right)$.

Theorem 6 (P. Austrin *et al.* [4]). *For every sufficiently large integer d, MVC on a graph d-bounded G is* UG-*hard to approximate within a factor* $2 - (2 + o_d(1)) \log\log d / \log d$.

Theorem 7 (M. Chlebík *et al.* [9]). *For every sufficiently large integer d, there is no* $(\log d - O(\log\log d))$*-approximation for MDS on d-bounded graphs unless* $NP \subseteq DTIME\left(n^{O(\log\log n)}\right)$.

Theorem 8. *MIS is* UG-*hard to approximate to within a factor* $1 - \frac{2\left(c_1 - O\left(\log^2 c_1\right)\right)}{c_1(c_1+1)\zeta(\beta)}$ *on PLG.*

Proof. In this proof, we construct the power law graph based on *Cycle-Based Embedding Technique* in Theorem 1 and show the Gap-Preserving from MIS on d-bounded graph G_d to MIS on power law graph $G_{(\alpha,\beta)}$. Let ϕ be a feasible solution on G_d and φ be a feasible solution on $G_{(\alpha,\beta)}$. We show *Completeness* and *Soundness* with $m' = m + (N-n)/2$.

– If $OPT(\phi) = m \Rightarrow OPT(\varphi) = m'$
  Let $OPT(\phi) = m$ be the MIS on graph G_d, we have $OPT(\varphi)$ which is composed of several parts: (1) $OPT(\phi) = m$; (2) MIS on clique K_2, cycle and d-Regular Cycle are all exactly half number of all vertices. Therefore, MIS on $G_{(\alpha,\beta)} \setminus G_d$ is $(N-n)/2$, where N and n are respectively the number of vertices on $G_{(\alpha,\beta)}$ and G_d. We have $OPT(\varphi) = OPT(\phi) + (N-n)/2$. That is, $OPT(\varphi) = m'$ where $m' = m + (N-n)/2$.
– If $OPT(\phi) < O\left(\log^2 d/d\right) m \Rightarrow OPT(\varphi) < \left(1 - \frac{2\left(c_1 - O\left(\log^2 c_1\right)\right)}{c_1(c_1+1)\zeta(\beta)}\right) m'$

$$OPT(\varphi) = OPT(\phi) + \frac{N-n}{2} < O\left(\frac{\log^2 d}{d}\right) m + \frac{N-n}{2}$$
$$= \left(1 - \frac{\left(1 - O\left(\frac{\log^2 d}{d}\right)\right) m}{m + \frac{N-n}{2}}\right) m' < \left(1 - \frac{\left(1 - O\left(\frac{\log^2 d}{d}\right)\right)}{\frac{N}{2m}}\right) m'$$
$$< \left(1 - \frac{1 - O\left(\frac{\log^2 d}{d}\right)}{\frac{(d+1)N}{2n}}\right) m' < \left(1 - \frac{2n\left(1 - O\left(\frac{\log^2 d}{d}\right)\right)}{N(d+1)}\right) m'$$
$$= \left(1 - \frac{2n\left(1 - O\left(\frac{\log^2 d}{d}\right)\right)}{\zeta(\beta)(d+1)n}\right) m' \le \left(1 - \frac{2\left(c_1 - O\left(\log^2 c_1\right)\right)}{c_1(c_1+1)\zeta(\beta)}\right) m'$$

where c_1 is the minimum integer d satisfying Theorem 5.

Equation (1) holds since $1 \le OPT(\phi) < O\left(\frac{\log_2 d}{d}\right) m$. Since G_d is a d-bounded graph, $m \ge n/(d+1)$. The last step holds since it is easy to see that function $f(x) = (x - O\left(\log^2 x\right))/(x(x+1))$ is monotonously decreasing when $f(x) > 0$ for any $x > 0$.

Theorem 9. *MVC is* UG*-hard to be approximated within* $1 + \frac{2\left(1-(2+o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)}{(c_2+1)\zeta(\beta)}$ *on PLG.*

Proof. The proof is similar to the inapproximability of MIS. We only show the *Soundness* here.

$$OPT(\varphi) = OPT(\phi) + \frac{N-n}{2} > \left(1 + \frac{1 - (2 + o_d(1))\frac{\log\log d}{\log d}}{1 + \frac{N-n}{2m}}\right) m'$$
$$> \left(1 + \frac{2n\left(1 - (2 + o_d(1))\frac{\log\log d}{\log d}\right)}{(d+1)\zeta(\beta)n}\right) m' > \left(1 + \frac{2\left(1 - (2 + o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)}{(c_2+1)\zeta(\beta)}\right) m'$$

where c_2 is the minimum integer d satisfying Theorem 6 and $m' = (N-n)/2$. The inequality holds since function $f(x) = (1-(2+o_x(1))\log\log x/\log x)/(x+1)$ is monotonously decreasing when $f(x) > 0$ for all x.

Theorem 10. *There is no* $1 + \frac{2(\log c_3 - O(\log\log c_3)-1)}{(c_3+1)\zeta(\beta)}$*-approximation for Minimum Dominating Set on PLG unless* $NP \subseteq DTIME\left(n^{O(\log\log n)}\right)$.

Proof. In this proof, we construct the power law graph based on *Cycle-Based Embedding Technique* in Theorem 1 and show the Gap-Preserving from MDS on d-bounded graph G_d to MDS on power law graph $G_{(\alpha,\beta)}$. Let ϕ and φ be feasible solutions on G_d and $G_{(\alpha,\beta)}$. We show *Completeness* and *Soundness.*

- If $OPT(\phi) = m \Rightarrow OPT(\varphi) = m'$
  Let $OPT(\phi) = m$ be the MDS on graph G_d, we have $OPT(\varphi)$ which is composed of several parts: (1) $OPT(\phi) = m$; (2) MDS on a K_2 is 1, $n/4$ on a $\boldsymbol{d}$-Regular Cycle according to Lemma 1 and $n/3$ on a cycle. That is, $OPT(\varphi) = m'$ where $m' = m + n_1/2 + n_2/3 + n_3/4$, where n_1, n_2 and n_3 corresponds to $\tau(1)$, $\tau(2)$ and all leftover vertices in Theorem 1.
- If $OPT(\phi) > (\log d - O(\log\log d))\, m \Rightarrow OPT(\varphi) > \left(1 + \frac{2(\log c_3 - O(\log\log c_3)-1)}{(c_3+1)\zeta(\beta)}\right) m'$

$$OPT(\varphi) = OPT(\phi) + n_1/2 + n_2/3 + n_3/4$$
$$> \left(1 + \frac{((\log d - O(\log\log d)) - 1)}{1+(N-n)/(2m)}\right) m' > \left(1 + \frac{2\left(\log c_3 - O(\log\log c_3) - 1\right)}{(c_3+1)\zeta(\beta)}\right) m'$$

  where $c_3 = \max\{\gamma_1, \gamma_2\}$, where γ_1 is the minimum integer d satisfying Theorem 7 and γ_2 satisfying $\frac{df(x)}{dx} = 0$ with function $f(x) = (\log x - O(\log\log x) - 1)/(x+1)$. Why we choose such c_3 is that γ_2 is the maxima of $f(x)$.

5.2 $\boldsymbol{\rho}$-Dominating Set Problem

Theorem 11. ρ*-PDS is* UG*-hard to be approximated into* $2 - (2 + o_d(1))\frac{\log\log d}{\log d}$ *on d-bounded graphs.*

Proof. In this proof, we show the Gap-Preserving from MVC on (d/ρ)-bounded graph $G = (V, E)$ to ρ-PDS on d-bounded graph $G' = (V', E')$. *w.l.o.g.*, we assume that d and d/ρ are integers. We construct a graph $G' = (V', E')$ by adding new vertices and edges to G. For each edge $(u, v) \in E$, create k new vertices $uv_1, \ldots, uv_k$ where $1 \le k \le \lfloor 1/\rho \rfloor$ and $2k$ new edges (uv_i, u) and (uv_i, u) for all $i \in [1, k]$ as shown in Fig. 6. Clearly, $G' = (V', E')$ is a d-bounded graph.

Let ϕ and φ be solutions to MVC on G and G' respectively. We claim that $OPT(\phi) = OPT(\varphi)$.

On one hand, if $\{v_1, v_2, \ldots, v_j\} \in V$ is minimum vertex cover on G. Then $\{v_1, v_2, \ldots, v_j\}$ is a ρ-PDS on G' because every old vertex in V has ρ of all neighbors in MVC and every new vertex in $V' \setminus V$ has at least one of two neighbors in MVC. Thus $OPT(\phi) \ge OPT(\varphi)$. One the other hand, we can prove that $OPT(\varphi)$ does not contain new vertices, that is, $V' \setminus V$. Consider a

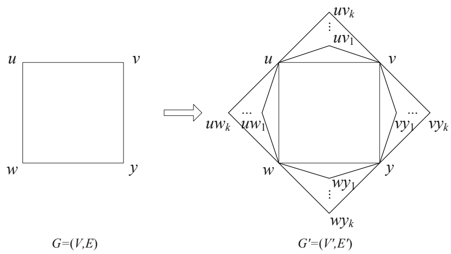


Fig. 6. Reduction from MVC to ρ-MDS

vertex $u \in V$, if $u \in OPT(\varphi)$, the new vertices uv_i for all $v \in N(u)$ and all $i \in [1,k]$ are not needed to be selected. If $u \notin OPT(\varphi)$, it has to be dominated by rho proportion of its all neighbors. That is, for each edge (u,v) incident to u, either v or all uv_i has to be selected since every uv_i has to be selected or dominated. If all uv_i are selected in $OPT(\varphi)$ for some edge (u,v), v is still not dominated by enough vertices if there are some more edges incident to v and the number of vertices uv_i k is great than 1, that is, $\lfloor 1/\rho \rfloor \geq 1$. In this case, therefore, v will be selected to dominate uv. Thus, $OPT(\varphi)$ does not contain new vertices. Since the verices in V selected is a solution to ρ-MDS, that is, for each vertex u in graph G, u will be selected or at least the number of neighbors of u will be selected. Therefore, the vertices in $OPT(\varphi)$ consist a Vertex Cover in G. Thus $OPT(\phi) \leq OPT(\varphi)$. Then we present the *Completeness* and *Soundness*.

- If $OPT(\phi) = m \Rightarrow OPT(\varphi) = m$
- If $OPT(\phi) > \left(2-(2+o_d(1))\frac{\log\log(d/2)}{\log(d/2)}\right) m \Rightarrow OPT(\varphi) > \left(2-(2+o_d(1))\frac{\log\log d}{\log d}\right) m$

$$OPT(\varphi) > \left(2-(2+o_d(1))\frac{\log\log(d/\rho)}{\log(d/\rho)}\right) m > \left(2-(2+o_d(1))\frac{\log\log d}{\log d}\right) m$$

since the function $f(x) = 2 - \log x/x$ is monotonously increasing for any x.

Theorem 12. *ρ-PDS is* UG*-hard to be approximated into* $1+\frac{2\left(1-(2+o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)}{(c_2+1)\zeta(\beta)}$ *on PLG.*

Proof. In this proof, we will show the Gap-Preserving from ρ-MDS on bounded degree graph G_d to ρ-MDS on power law graph $G_{(\alpha,\beta)}$.

We use the same construction as in Theorem 8. Let ϕ be a solution on G'_d and φ be a solution on $G_{(\alpha,\beta)}$, we prove the *Completeness* and *Soundness*.

- If $OPT(\phi) = m \Rightarrow OPT(\varphi) = m'$

  Let $OPT(\phi) = m$ be the ρ-MDS on graph G_d, we have $OPT(\varphi)$ which is composed of several parts: (1) $OPT(\phi) = m$; (2) MDS on a K_2 is 1, $g(\rho)n$ on a ***d***-Regular Cycle according to Lemma 1 and $f(\rho)n$ on a cycle, where $f(\rho) = \begin{cases} \frac{1}{4}, & \rho \leq \frac{1}{3} \\ \frac{1}{3}, & \frac{1}{3} < \rho \leq \frac{1}{2} \end{cases}$ and $g(\rho) = \frac{1}{3}$ for all $\rho \leq \frac{1}{2}$.

Therefore, ρ-MDS on $G_{(\alpha,\beta)}$ to be m' where $m' = m+n_1/2+f(\rho)n_2+g(\rho)n_3$, where n_1, n_2 and n_3 corresponds to $\tau(1)$, $\tau(2)$ and all leftover vertices in Theorem 1.

– If $OPT(\phi) > \left(2-(2+o_d(1))\frac{\log\log d}{\log d}\right) m \Rightarrow OPT(\varphi) > \left(1 + \frac{1-(2+o_{c_2}(1))\frac{\log\log c_2}{\log c_2}}{(c_2+1)\zeta(\beta)}\right) m'$

$$OPT(\varphi) = OPT(\phi) + n_1/2 + f(\rho)n_2 + g(\rho)n_3$$
$$> \left(1 + \frac{2n\left(1-(2+o_d(1))\frac{\log\log d}{\log d}\right)}{(d+1)\zeta(\beta)n}\right) m' > \left(1 + \frac{2\left(1-(2+o_{c_2}(1))\frac{\log\log c_2}{\log c_2}\right)}{(c_2+1)\zeta(\beta)}\right) m'$$

Again, c_2 is the minimum integer d satisfying Theorem 6. The inequality holds since function $f(x) = (1-(2+o_x(1))\log\log x/\log x)/(x+1)$ is monotonously decreasing when $f(x) > 0$ for any x.

5.3 Maximum Clique, Minimum Coloring

Theorem 13 (Hastad [15]). *There is no $n^{1-\epsilon}$-approximation on Maximum Clique problem unless NP=ZPP.*

Lemma 3 (Ferrante *et al.* [13]). *Let $G = (V, E)$ be a simple graph with n vertices and $\beta \geq 1$. Let $\alpha \geq \max\{4\beta, \beta\log n + \log(n+1)\}$. Then, $G_2 = G \setminus G_1$ is a bipartite graph.*

Lemma 4. *Given a function $f(x)$ ($x \in \mathbb{Z}$, $f(x) \in \mathbb{Z}^+$) monotonously decreases, $\sum_x f(x) \leq \int_x f(x)$.*

Corollary 3. $e^\alpha \sum_{i=1}^{e^{\alpha/\beta}} \left(\frac{1}{d}\right)^\beta < (e^\alpha - e^{\alpha/\beta})/(\beta-1)$.

Theorem 14. *Maximum Clique cannot be approximated within $O\left(n^{1/(\beta+1)-\epsilon}\right)$ on large PLG with $\beta > 1$ and $n > 54$ for any $\epsilon > 0$ unless* NP=ZPP.

Proof. In [13], the authors proved the hardness of Maximum Clique problem on power law network. Here we use the same construction. According to Lemma 3, $G_2 = G \setminus G_1$ is a bipartite graph when $\alpha \geq \max\{4\beta, \beta\log n + \log(n+1)\}$ for any $\beta \geq 1$. Let ϕ be a solution on general graph G and φ be a solution on power law graph G_2. We show the *Completeness* and *Soundness*.

– If $OPT(\phi) = m \Rightarrow OPT(\varphi) = m$
  If $OPT(\phi) \leq 2$ on graph G, we can solve Clique problem in polynomial time by iterating the edges and their end vertices one by one, where G is not a general graph in this case. *w.l.o.g*, assuming $OPT(\phi) > 2$, then $OPT(\varphi) = OPT(\phi) > 2$ since the maximum clique on bipartite graph is 2.
– If $OPT(\phi) \leq m/n^{1-\epsilon} \Rightarrow OPT(\varphi) < O\left(1/(N^{1/(\beta+1)-\epsilon'})\right) m$
  In this case, we consider the case that $4\beta < \beta\log n + \log(n+1)$, that is, $n > 54$. According to Lemma 3, let $\alpha = \beta\log n+\log(n+1)$. From Corollary 3, we have

$$N = e^{\alpha} \sum_{d=1}^{e^{\alpha/\beta}} \left(\frac{1}{d}\right)^{\beta} < \frac{e^{\alpha} - e^{\alpha/\beta}}{\beta - 1} = \frac{n^{\beta}(n+1) - n(n+1)^{1/\beta}}{\beta - 1} < \frac{2n^{\beta+1} - n}{\beta - 1}$$

Therefore, $OPT(\varphi) = OPT(\phi) \leq m/n^{1-\epsilon} < O\left(m/\left(N^{1/(\beta+1)-\epsilon'}\right)\right)$.

Corollary 4. *The Minimum Coloring problem cannot be approximated within $O\left(n^{1/(\beta+1)-\epsilon}\right)$ on large PLG with $\beta > 1$ and $n > 54$ for any $\epsilon > 0$ unless* NP=ZPP.

6 Relationship between β and Approximation Hardness

As shown in previous sections, many hardness results depend on β. In this section, we analyze the hardness of some optimization problems based on the value of β by showing that trivial greedy algorithms can achieve constant guarantee factor on MIS and MDS.

Lemma 5. *When $\beta > 2$, the size of MDS of a power law graph is greater than Cn where n is the number of vertices, C is some constant depended only on β.*

Proof. Let $MDS = (v_1, v_2, \ldots, v_t)$ with degrees $d_1, d_2, \ldots, d_t$ be the MDS of power-law graph $G = (V, E)$. The total of degrees of vertices in dominating set must be at least the number of vertices outside the dominating set. Thus $\sum_{i=1}^{i=t} d_i \geq |V \backslash DS|$. With a given total degrees, a set of vertices has minimum size when it includes highest degree vertices. With $\beta > 2$ the function $\zeta(\beta - 1) = \sum_{i=1}^{\infty} \frac{1}{i^{\beta-1}}$ is converged, there exists a constant $t_0 = t_0(\beta)$ such that

$$\sum_{i=t_0}^{\lfloor e^{\alpha/\beta} \rfloor} i \left\lfloor \frac{e^{\alpha}}{i^{\beta}} \right\rfloor \leq \sum_{i=1}^{t_0} \left\lfloor \frac{e^{\alpha}}{i^{\beta}} \right\rfloor$$

where α is any large enough constant. Thus the size of MDS is at least

$$\sum_{i=t_0}^{\lfloor e^{\alpha/\beta} \rfloor} \left\lfloor \frac{e^{\alpha}}{i^{\beta}} \right\rfloor \approx \left(\zeta(\beta) - \sum_{i=1}^{t_0} \frac{1}{i^{\beta}}\right) e^{\alpha} \approx C|V|$$

where $C = (\zeta(\beta) - \sum_{i=1}^{t_0} \frac{1}{i^{\beta}})/(\zeta(\beta))$.

Consider the greedy algorithm which selects vertices from the highest degree vertices to lowest one. In the worst case, it selects all vertices with degree greater than 1 and a half of vertices with degree 1 to form a dominating set. The approximation factor of this simple algorithm is a constant.

Corollary 5. *Given a power law graph with $\beta > 2$, the greedy algorithm that selects vertices in decreasing order of degrees provides a dominating set of size at most $\sum_{i=2}^{\lfloor e^{\alpha/\beta} \rfloor} \lfloor e^{\alpha}/i^{\beta} \rfloor + \frac{1}{2} e^{\alpha} \approx (\zeta(\beta) - 1/2) e^{\alpha}$. Thus the approximation ratio is $(\zeta(\beta) - \frac{1}{2})/(\zeta(\beta) - \sum_{i=1}^{t_0} 1/i^{\beta})$.*

Let us consider a maximization problem MIS, we propose a greedy algorithm Power-law-Greedy-MIS as follows. Sort the vertices in non-increasing order then start checking from the lowest degree vertex, if the vertex is not adjacent to any selected vertex, it is selected. The set of selected vertices forms an independent set with the size at least a half the number of vertices with degree 1 which is $e^{\alpha}/2$. The size of MIS is at most a half of number of vertices, then we have

Lemma 6. *Power-law-Greedy-MIS has factor* $1/(2\zeta(\beta))$ *on PLG with* $\beta > 1$.

Acknowledgment

This work is partially supported by NSF Career Award ♯ 0953284 and DTRA, Young Investigator Award, Basic Research Program ♯ HDTRA1-09-1-0061.

References

1. Aiello, W., Chung, F., Lu, L.: A random graph model for massive graphs. In: STOC 2000, pp. 171–180. ACM, New York (2000)
2. Aiello, W., Chung, F., Lu, L.: A random graph model for power law graphs. Experimental Math. 10, 53–66 (2000)
3. Alimonti, P., Kann, V.: Hardness of approximating problems on cubic graphs. In: Bongiovanni, G., Bovet, D.P., Di Battista, G. (eds.) CIAC 1997. LNCS, vol. 1203, pp. 288–298. Springer, Heidelberg (1997)
4. Austrin, P., Khot, S., Safra, M.: Inapproximability of vertex cover and independent set in bounded degree graphs. In: CCC 2009, pp. 74–80 (2009)
5. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
6. Bianconi, G., Barabási, A.L.: Bose-einstein condensation in complex networks (2000)
7. Bondy, J., Murty, U.: Graph theory with applications. MacMillan, London (1976)
8. Bornholdt, S., Schuster, H.G. (eds.): Handbook of Graphs and Networks: From the Genome to the Internet. John Wiley & Sons, Inc., New York (2003)
9. Chlebík, M., Chlebíková, J.: Approximation hardness of dominating set problems in bounded degree graphs. Inf. Comput. 206(11), 1264–1275 (2008)
10. Dinur, I., Safra, S.: On the hardness of approximating minimum vertex cover. Annals of Mathematics 162, 2005 (2004)
11. Erdos, P., Gallai, T.: Graphs with prescribed degrees of vertices. Mat. Lapok 11, 264–274 (1960)
12. Eubank, S., Kumar, V.S.A., Marathe, M.V., Srinivasan, A., Wang, N.: Structural and algorithmic aspects of massive social networks. In: SODA 2004, pp. 718–727. Society for Industrial and Applied Mathematics, Philadelphia (2004)
13. Ferrante, A., Pandurangan, G., Park, K.: On the hardness of optimization in power-law graphs. Theoretical Computer Science 393(1-3), 220–230 (2008)
14. Gkantsidis, C., Mihail, M., Saberi, A.: Conductance and congestion in power law graphs. SIGMETRICS Perform. Eval. Rev. 31(1), 148–159 (2003)
15. Hastad, J.: Clique is hard to approximate within $n^{1-\epsilon}$. In: FOCS 1996, Washington, DC, USA, p. 627. IEEE Computer Society, Los Alamitos (1996)
16. Janson, S., Luczak, T., Norros, I.: Large cliques in a power-law random graph (2009)

17. Kempe, D., Kleinberg, J., Tardos, É.: Influential nodes in a diffusion model for social networks. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M. (eds.) ICALP 2005. LNCS, vol. 3580, pp. 1127–1138. Springer, Heidelberg (2005)
18. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: KDD, pp. 137–146. ACM Press, New York (2003)
19. Kleinberg, J.: The small-world phenomenon: An algorithmic perspective. In: 32nd STOC, pp. 163–170 (2000)
20. Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A., Upfal, E.: Stochastic models for the web graph. In: FOCS 2000, p. 57 (2000)
21. Norros, I., Reittu, H.: On a conditionally poissonian graph process. In: Advances in Applied Probability, pp. 38–59 (2006)
22. Pandurangan, G., `http://www.cs.purdue.edu/homes/gopal/powerlawtalk.pdf`

Appendix A: Embedding Construction with $\beta < 1$

Ferrante *et. al.* [13] proved the *NP*-hardness of MIS, MDS, and MVC where $\beta < 1$ based on Lemma 7 which is invalid. A counter-example is as follows. Let $D_1 =< 3, 2, 2, 1 >$ and $D_2 =< 7, 6, 5, 4, 3, 2, 2, 1 >$ then D_1 is eligible and $Y_1 =< 1, 2, 1 >< Y_2 =< 1, 2, 1, 1, 1, 1, 1 >$ but D_2 is NOT eligible with $f_{D_2}(4) < 0$. In this part, we present an alternative lemma to prove the hardness of these problems on power-law graphs with $\beta < 1$.

Definition 9 (d-Degree Sequence). *Given a graph $G = (V, E)$, the d-degree sequence of G is a sequence $D =< d_1, d_2, \dots, d_n >$ of vertex degrees in non-increasing order.*

Definition 10 (y-Degree Sequence). *Given a graph $G = (V, E)$, the y-degree sequence of G is a sequence $Y =< y_1, y_2, ..., y_m >$ where m is the maximum degree of G and $y_i = |\{u|u \in V$ and $degree(u) = i\}|$.*

Definition 11 (Eligible Sequences). *A sequence of integers $S =< s_1, \dots, s_n >$ is eligible if $s_1 \geq s_2 \geq \dots \geq s_n$ and, for all $k \in [n]$, $f_S(k) \geq 0$, where*

$$f_S(k) = k(k-1) + \sum_{i=k+1}^{n} min\{k, s_i\} - \sum_{i=1}^{k} s_i$$

Lemma 7 (Invalid Lemma, [13]). *Let Y_1 and Y_2 be two y-degree sequences with m_1 and m_2 elements respectively such that (1) $Y_1(i) \leq Y_2(i)$, $\forall 1 \leq i \leq m_1$, and (2) two corresponding d-degree sequences D_1 and D_2 are contiguous. If D_1 is eligible then D_2 is eligible.*

Erdős and Gallai [11] showed that a sequence of integers to be graphic - d-degree sequence of an graph, *iff* it is eligible and the total of all elements is even. Then Havel and Hakimi [7] gave an algorithm to construct a simple graph from a degree sequence.

Lemma 8 ([7]). *A sequence of integers $D =< d_1, \dots, d_n >$ is graphic if and only if it is non-increasing, and the sequence of values $D' =< d_2 - 1, d_3 - 1, \dots, d_{d_1+1} - 1, d_{d_1+2}, \dots, d_n >$ when sorted in non-increasing order is graphic.*

We now prove the following lemma, which can substitute Lemma 7 for the NP-hardness proof in [13].

Lemma 9. *Given an undirected graph $G = (V, E)$, $0 < \beta < 1$, there exists polynomial time algorithm to construct power-law graph $G' = (V', E')$ of exponential factor β such that G is a set of maximal components of G'.*

Proof. To construct G', we choose $\alpha = max\{\beta \ln(n-1) + \ln(n+2), 3\ln 2\}$ then $\lfloor e^{\alpha}/((n-1)^{\beta})\rfloor > n+2$, i.e. if there are a least 2 vertices of G' having degree d, there are at least 2 vertices of $G'\backslash G$ having degee d. According to the definition, the total degrees of all vertices in G' and G are even. Therefore, the lemma will follow if we prove that the degree sequence D of $G'\backslash G$ is eligible.

In D, the maximum degree is $\lfloor e^{\alpha/\beta}\rfloor$. There is only one vertex of degree i if $1 \le e^{\alpha}/i^{\beta} < 2$ and furthermore $e^{\alpha/\beta} \ge i > e^{(\alpha-\ln 2)/\beta} = (e^{\alpha}/2)^{1/\beta}$.

We check $f_D(k)$ in two cases:

1. **Case 1**: $k \le \left\lfloor e^{\alpha/\beta}/2 \right\rfloor$

$$
\begin{aligned}
f_D(k) &= k(k-1) + \sum_{i=k+1}^{n} min\{k, d_i\} - \sum_{i=1}^{k} d_i \\
&> k(k-1) + \sum_{i=k}^{T-k} k + \sum_{i=B}^{k-1} i + \sum_{i=1}^{B-1} 2 - \sum_{i=1}^{k} (T-k+1) \\
&= k(T-k) + (k-B)(k-1+B)/2 + B(B-1) - k(2T-k+1)/2 \\
&= (B^2 - B)/2 - k
\end{aligned}
$$

where where $T = \left\lfloor e^{\alpha/\beta} \right\rfloor$ and $B = \left\lfloor (e^{\alpha}/2)^{1/\beta} \right\rfloor + 1$. Note that with $\alpha > 3\ln 2$, $\alpha/\beta > \ln 2\,(2/\beta + 1)$. Hence $\left(\left\lfloor (e^{\alpha}/2)^{1/\beta} \right\rfloor + 1\right)\left(\left\lfloor (e^{\alpha}/2)^{1/\beta} \right\rfloor\right) > \left\lfloor e^{\alpha/\beta} \right\rfloor \ge 2k$, so $f_D(k) > 0$.

2. **Case 2**: $k > \left\lfloor e^{\alpha/\beta}/2 \right\rfloor$

$$
f_D(k+1) \ge f_D(k) + 2k - 2d_{k+1} \ge f_D(k) \ge \ldots \ge f_D(\left\lfloor e^{\alpha/\beta}/2 \right\rfloor) > 0
$$