Approximating Partially Bounded Degree Deletion on Directed Graphs

Toshihiro Fujito$^{(\boxtimes)}$ 🆔, Kei Kimura, and Yuki Mizuno

Toyohashi University of Technology, Toyohashi 441-8580, Japan
{fujito,kimura}@cs.tut.ac.jp, mizuno@algo.cs.tut.ac.jp

Abstract. The BOUNDED DEGREE DELETION problem (BDD) is that of computing a minimum vertex set in a graph $G = (V, E)$ with degree bound $b : V \to \mathbb{Z}_+$, such that, when it is removed from G, the degree of any remaining vertex v is no larger than $b(v)$. It is a classic problem in graph theory and various results have been obtained including an approximation ratio of $2 + \ln b_{\max}$ [30], where $b_{\max}$ is the maximum degree bound.

This paper considers BDD on *directed* graphs containing *unbounded* vertices, which we call PARTIALLY BOUNDED DEGREE DELETION (PBDD). Despite such a natural generalization of standard BDD, it appears that PBDD has never been studied and no algorithmic results are known, approximation or parameterized. It will be shown that (1) in case all the possible degrees are bounded, in-degrees by b^- and out-degrees by b^+, BDD on directed graphs can be approximated within $2 + \max_{v \in V} \ln(b^-(v) + b^+(v))$, and (2) although it becomes NP-hard to approximate PBDD better than $b_{\max}$ (even on *undirected* graphs) once unbounded vertices are allowed, it can be within $\max\{2, b_{\max} + 1\}$ when only in-degrees (and none of out-degrees) are partially bounded by b.

Keywords: Approximation algorithms · Bounded Degree Deletion
Partial cover

1 Introduction

The BOUNDED DEGREE DELETION problem is a well-known basic problem in graph theory. It has an application in various areas such as computational biology [15] and property testing [29], whereas its "dual problem" of finding maximum s-plexes, introduced in 1978 [32], has applications in social network analysis [1,28]. With degree bound of $b \in \mathbb{Z}_+$, b-BOUNDED DEGREE DELETION (or b-BDD for short) is the problem of computing a minimum cost vertex set X in a given weighted graph $G = (V, E)$ such that the degree of any remaining vertex v is bounded by b when all the vertices in X are removed from G.

Clearly, b-BDD is a generalization of the VERTEX COVER (VC) problem, and another generalization of VC has been recently introduced and actively studied. The k-PATH VERTEX COVER (k-PVC) problem [4–6, 22, 24], also known as VERTEX COVER P_k [34–37], P_k-HITTING SET [7], and k-PATH TRANSVERSAL [27], is the problem of computing a minimum vertex set C such that when all the vertices in C are removed from G, there remains no path on k vertices. A subset of vertices in a graph G is called a *dissociation set* if it induces a subgraph with maximum degree at most 1. The maximum cardinality of a dissociation set in G is called the *dissociation number* of G. The problem of computing the dissociation number was introduced by Yannakakis [40], who also proved it to be NP-hard in the class of bipartite graphs. See [31] for a survey on the dissociation number problem. Clearly, VC $\equiv$ 0-BDD $\equiv$ 2-PVC, 1-BDD $\equiv$ 3-PVC (but b-BDD $\not\equiv (b+2)$-PVC for $b \geq 2$), and a dissociation set is the complement of a 3-PVC (i.e., 1-BDD) solution.

We now summarize below algorithmic results known for b-BDD and related problems.

VC. It is known approximable within $2 - \Theta(1/\sqrt{\log n})$ [23], whereas VC has been shown hard to approximate within $10\sqrt{5} - 21 \approx 1.36$ unless P = NP [13] (or within $2 - \epsilon$ assuming the unique games conjecture [25]).

b-**BDD.** The first improvement over the simple $(b+2)$-approximation based on the hitting set formulation was attained in [17] by using the local ratio method and b-BDD was shown approximable within $\max\{2, b+1\}$. Okun and Barak considered more general b-BDD where $b : V \to \mathbb{Z}_+$ is an arbitrary function, and obtained an approximation bound of $2 + \ln b_{\max}$ by combination of the local ratio method and the greedy multicovering [30], where $b_{\max} = \max_{v \in V} b(v)$. Recently, a new approximation bound of $\max\{2, b_{\max}/2 + 1\}$ was obtained [19].

b-BDD has been extensively studied in parameterized complexity. It has been shown that, when parameterized by the size k of the deletion set, the problem is $W[2]$-hard for unbounded b and FPT for each fixed $b \geq 0$ [15], whereas, when parameterized by treewidth tw, it is FPT with parameters k and tw, and $W[2]$-hard with only parameter tw [3]. A linear vertex kernel of b-BDD has been developed by generalizing the Nemhauser-Trotter theorem for VC to b-BDD [10, 15, 39].

Besides, 2-BDD has been recently highlighted under the name of CO-PATH/CYCLE PACKING [9, 10, 16], mostly from the viewpoint of parameterized complexity, due to its important applications in bioinformatics.

3-**PVC.** It was shown approximable within 2 [36, 37] (or within an expected approximation ratio of 23/11 by a randomized algorithm [24]) in general, and within 1.57 on cubic graphs [35].

1.1   Our Work and Contributions

We generalize BDD in two directions; in one to the problem of *directed* degree bounds, and in the other to the problem where some vertices are allowed to be

of unbounded degree. PARTIALLY b-BOUNDED DEGREE DELETION (b-PBDD) is, given a *directed* graph $G = (V \cup V_0, E)$ and a degree bound $b : V \to \mathbb{Z}_+$, to compute a minimum cost vertex subset $X \subseteq V \cup V_0$ such that the *in-degree* of any vertex $v \in V$ remaining after all the vertices in X are deleted from G is at most $b(v)$. Notice that the degree bound b is defined only on V and no bound is imposed on V_0. To the best of our knowledge, neither version, directed BDD nor partial BDD, has been previously studied, in either aspect of approximation complexity or parameterized one, except for the case of 1-BDD on directed graphs, which was shown approximated within 2 [17]. Certainly, 0-PBDD $\equiv$ 0-BDD $\equiv$ VC when $V_0 = \emptyset$, but b-PBDD $\not\equiv$ b-BDD even for $b = 1$.

Directed graphs provide more general computational models than undirected graphs, but problems tend to be harder to deal with on the former than the latter. Another type of generalization, in the setting of BDD, is to allow for "don't care" nodes. In fact the notion of "covering" or "domination" has been generalized to *partial* "covering/domination" and a significant amount of research work has been devoted to such extensions [2,8,11,14,20,21,26,33], where, instead of complete coverage or domination, only a prescribed fraction of covering or domination is required. Here we consider PBDD having unbounded vertices as defined above to be a natural extension of BDD to the partial version. The current work is partially motivated by the fact that the (logarithmically) bounded performance of the best algorithm for the standard BDD, however, becomes unbounded when applied to the partial version as will be explained in Sect. 4.1.

This paper presents that (1) in case all the possible degrees are bounded, in-degrees by b^- and out-degrees by b^+ (and $V_0 = \emptyset$), BDD on directed graphs can be approximated within $2 + \max_{v \in V} \ln(b^-(v) + b^+(v))$ by generalizing the algorithm of Okun and Barak [30], and (2) although it becomes NP-hard to approximate b-PBDD better than $b_{\max}$ (even on *undirected* graphs) once unbounded degrees are allowed, it can be within $\max\{2, b_{\max} + 1\}$ when only in-degrees (and none of out-degrees) are partially bounded by b.

1.2 Notations and Definitions

For a vertex set X in a digraph $G = (V, E)$, let $E(X) = \{(u, v) \in E \mid \{u, v\} \subseteq X\}$. Let $\delta^-(X)$ denote the set of arcs entering from outside of X to a vertex in X, i.e., $\delta^-(X) = \{(u, v) \in E \mid u \notin X, v \in X\}$ and $\delta(X)$ be the set of arcs incident to a vertex in X, i.e., $\delta(X) = \{(u, v) \in E \mid \{u, v\} \cap X \neq \emptyset\}$. Let $\delta^-(v)$ ($\delta(v)$, resp.) denote $\delta^-(\{v\})$ ($\delta(\{v\})$, resp.). The in-degree and out-degree of v is denoted by $d^-(v)$ ($= |\delta^-(v)|$) and $d^+(v)$, respectively. To restrict arcs under consideration within a certain arc set F, we use $\delta_F^-(X)$ and $d_F^-(v)$ to denote $\delta^-(X) \cap F$ and $|\delta^-(v) \cap F|$, respectively, and $d_{E(X)}^-(v)$ abbreviated to $d_X^-(v)$ for $X \subseteq V$. For the set of neighboring vertices of $u \in V$, let $N^+(u)$ and $N^-(u)$ denote $\{v \in V \mid (u, v) \in E\}$ and $\{v \in V \mid (v, u) \in E\}$, respectively.

We also use shorthand notations for functions b, d^-, and $\bar{w}$ (to be defined in Sect. 3) defined on V and $Z \subseteq V$ such as $b(Z) = \sum_{v \in Z} b(v)$, $d^-(Z) = \sum_{v \in Z} d^-(v)$, and $\bar{w}(Z) = \sum_{v \in Z} \bar{w}(v)$.

2   Approximating PBDD via Submodular Optimization

Assume that $b(v) \leq d^-(v), \forall v \in V$, for the rest of paper as one can always reset $b(v)$ to $d^-(v)$, without loss of generality, if $b(v) > d^-(v)$. A vertex $v \in V$ is called a *tight node* in what follows if $d^-(v) = b(v)$ (and it is *untight* if $d^-(v) > b(v)$).

For a directed graph $G = (V \cup V_0, E)$ and $b : V \rightarrow \mathbb{Z}_+$, define the rank $r : 2^E \rightarrow \mathbb{Z}_+$ of $F \subseteq E$ such that

$$r(F) = \sum_{v \in V} \min\{d^-_F(v), b(v)\} + \sum_{v \in V_0} d^-_F(v).$$

Then (E, r) is a matroid, a direct sum of *partition matroids* and *free matroids*, and an arc set $F \subseteq E$ is independent iff $d^-_F(v) \leq b(v), \ \forall v \in V$. Thus, PBDD is the problem of computing $X \subseteq V$ of minimum cost such that the arc set induced by $V - X$ is independent in (E, r).

Definition 1. *For a matroid (E, r) let $r^d : 2^E \rightarrow \mathbb{Z}_+$ be such that*

$$r^d(S) = |S| - (r(E) - r(E \setminus S)).$$

Then, r^d is a matroid rank function and (E, r^d) is called the dual *of (E, r).*

Proposition 1. *Let (E, r) be the matroid defined by (G, b) as above.*

- $r(E) = b(V) + d^-(V_0)$ *(assuming that $b(v) \leq d^-(v), \ \forall v \in V$).*
- *In the dual matroid (E, r^d),*
  - $r^d(F) = |F| - (r(E) - r(E \setminus F))$

$$= \sum_{v \in V} \left(d^-_F(v) - \min\{b(v), d^-(v)\} + \min\{b(v), d^-_{E \setminus F}(v)\} \right)$$

$$+ \sum_{v \in V_0} \left(d^-_F(v) - d^-(v) + d^-_{E \setminus F}(v) \right)$$

$$= \sum_{v \in V} \left(\min\{b(v) + d^-_F(v), d^-(v)\} - \min\{b(v), d^-(v)\} \right)$$

$$= \sum_{v \in V} \min\{d^-_F(v), d^-(v) - b(v)\}.$$

  - $r^d(E) = |E| - r(E) = |E| - b(V) - d^-(V_0) = d^-(V) - b(V).$
  - $r^d(\delta(v)) = \sum_{u \in V} \min\{d^-_{\delta(v)}(u), d^-(u) - b(u)\}$

$$= \begin{cases} d^-(v) - b(v) + (\# \text{ of untight nodes in } N^+(v) \cap V) & \text{if } v \in V \\ (\# \text{ of untight nodes in } N^+(v) \cap V) & \text{if } v \in V_0 \end{cases}$$

Let $X \subseteq V \cup V_0$ be partitioned to $\tilde{X}, X_t$, and X_0 s.t. $X_0 = X \cap V_0$, $X_t = \{v \in X \setminus X_0 \mid v \text{ is tight}\}$, and $\tilde{X} = X \setminus (X_t \cup X_0)$. Likewise, for $Y = (V \cup V_0) \setminus X$

let $Y = \tilde{Y} \cup Y_t \cup Y_0$ s.t. $Y_0 = Y \cap V_0$, $Y_t = \{v \in Y \setminus Y_0 \mid v \text{ is tight}\}$, and $\tilde{Y} = Y \setminus (Y_t \cup Y_0)$. Then, since

$$\sum_{v \in X} r^d(\delta(v)) = \sum_{v \in \tilde{X}} r^d(\delta(v)) + \sum_{v \in X_t} r^d(\delta(v)) + \sum_{v \in X_0} r^d(\delta(v))$$

$$= \sum_{v \in \tilde{X}} (d^-(v) - b(v)) + \sum_{v \in X} (\# \text{ of untight nodes in } N^+(v) \cap V),$$

we have

Proposition 2.

$$\sum_{v \in X} r^d(\delta(v)) = d^-(\tilde{X}) - b(\tilde{X}) + \sum_{v \in X} \left| N^+(v) \cap (\tilde{X} \cup \tilde{Y}) \right|.$$

Note: Propositions 1 and 2 will be useful in proof of Lemma 1.

In general a subset $F \subseteq E$ is independent in a matroid iff F is spanning in its dual matroid. Thus, $X \subseteq V$ is a b-PBDD solution in $G = (V \cup V_0, E)$ iff $\delta(X)$ is spanning in (E, r^d). Therefore, b-PBDD on $G = (V, E)$ can be reduced to the problem of computing $X \subseteq V$ of minimum cost such that $\delta(X)$ is spanning in (E, r^d). More formally,

Proposition 3. *Define* $f : 2^V \to \mathbb{Z}_+$ *such that* $f(W) = r^d(\delta(W))$. *b-PBDD on* $G = (V, E)$ *can be formulated as the problem of computing* $X \subseteq V$ *of minimum cost such that* $f(X) = f(V)$.

It is known that f as defined above is nondecreasing and *submodular*, and the problem of computing minimum $X \subseteq V$ satisfying $f(X) = f(V)$ for such a function f is known as the *submodular set cover* problem.

Definition 2. *Let* f *be a nondecreasing submodular set function defined on the subsets of a finite ground set N, and* w_j *be a nonnegative cost associated with each element $j \in N$. The* SUBMODULAR SET COVER *problem (SSC) is to compute:*

$$\min_{S \subseteq N} \left\{ \sum_{j \in S} w_j \mid f(S) = f(N) \right\}.$$

The greedy algorithm, together with its performance analysis, is perhaps the most well-known heuristic for general SSC [38], but the primal-dual algorithm based on the following LP relaxation of SSC and its dual LP is also known to deliver better solutions for some of more specific SSC problems (See [18] for more details).

(P) $\min \sum_{j \in N} w_j x_j$ (D) $\max \sum_{S \subseteq N} f_S(N - S) y_S$

subject to: subject to:

$$\sum_{j \in N-S} f_S(j) x_j \geq f_S(N - S), \forall S \subseteq N \qquad \sum_{S : j \notin S} f_S(j) y_S \leq w_j, \forall j \in N$$

$$x_j \geq 0, \quad \forall j \in N \qquad\qquad\qquad y_S \geq 0, \quad \forall S \subseteq N$$

Here and in the algorithm called PD, the *contraction* of f onto $N - S$ is the function f_S defined on 2^{N-S} such that $f_S(X) = f(X \cup S) - f(S)$ for any $S \subseteq N$. If f is nondecreasing and submodular on N, so is f_S on $N - S$, and thus, another submodular set cover instance $(N - S, f_S)$ can be derived for any $S \subseteq N$. The performance of PD for general SSC can be estimated by the following theorem.

Theorem 1 ([18]). *The performance ratio of the primal-dual algorithm PD for an SSC instance (N, f) is bounded by*

$$\max \left\{ \frac{\sum_{j \in X} f_S(j)}{f_S(N - S)} \right\}$$

where $\max$ is taken over any $S \subseteq N$ and any minimal solution X in $(N - S, f_S)$.

It is more convenient, when applying Theorem 1 to an instance (G, b) of b-PBDD, to use it in the following form.

Corollary 1. *The performance ratio of PD, when applied to an instance $(G = (V, E), b)$ of b-PBDD, is bounded by*

$$\max \left\{ \frac{\sum_{v \in X} r^d(\delta(v))}{r^d(E)} \right\}$$

where (E, r) is the matroid defined by an instance (G, b) and $\max$ is taken over any graph G and any minimal solution X in G.

Proof. Consider the graph $\bar{G} = G - S$ obtained from G by removing all the vertices in S, and reformulate b-PBDD on $\bar{G} = (\bar{V}, \bar{E})$, where $\bar{V} = (V \cup V_0) - S, \bar{E} = E - \delta(S)$, as an SSC instance $(\bar{E}, \bar{f})$. To do so, let $\bar{r} : 2^{\bar{E}} \to \mathbb{Z}_+$ be the rank function of the matroid defined by $(\bar{E}, \bar{b})$, such that $\bar{r}(F) = \sum_{v \in \bar{V}} \min\{\bar{d}_F^-(v), \bar{b}(v)\}$ for $F \subseteq \bar{E}$, $\bar{r}^d$ be the dual of $\bar{r}$, and $\bar{f}(T) = \bar{r}^d(\bar{\delta}(T))$ for $T \subseteq \bar{V}$ (Note: Here, $\bar{\delta}(T) = \delta_{\bar{E}}(T), \bar{d}(v) = d_{\bar{E}}(v), \bar{b}(v) = \min\{b(v), \bar{d}^-(v)\}$ for all $T \subseteq \bar{V}$ and $v \in \bar{V}$). It can be shown then that $f_S(T) = \bar{f}(T)$ for any $S \subseteq V \cup V_0$ and $T \subseteq (V \cup V_0) - S$, and in particular, $\bar{f}(v) = f_S(v), \forall v \in \bar{V}$, and $\bar{f}(\bar{V}) = f_S((V \cup V_0) - S)$. Hence, we have

$$\max_{S \subseteq V \cup V_0} \left\{ \frac{\sum_{v \in X} f_S(v)}{f_S((V \cup V_0) - S)} \right\} = \max \left\{ \frac{\sum_{v \in X} \tilde{f}(v)}{\tilde{f}(\tilde{V})} \right\} \qquad (1)$$

where $\max$ in RHS is taken over any subgraph $\bar{G}$ of G induced by $\bar{V} \subseteq V \cup V_0$ and any minimal b-PBDD solution X in $\bar{G}$. It thus follows from Theorem 1 and Eq. (1) that the performance ratio of PD, when applied to b-PBDD, can be estimated by bounding

$$\frac{\sum_{v \in X} f(v)}{f(V \cup V_0)} = \frac{\sum_{v \in X} r^d(\delta(v))}{r^d(E)}$$

for any graph $G = (V \cup V_0, E)$ and any minimal solution X in G. $\square$

3   Fully Bounded Degree Deletion

It can be observed, modifying the undirected instance to be used in Sect. 4.1 to a directed one, that the greedy set cover approximation is embeddable even if all the in-degrees are bounded in directed graphs. On the other hand, BDD on directed graphs where both in-degree and out-degree are bounded at every vertex can be approximated in much the same way as in the case of undirected graphs. To explain this, suppose all the possible degree bounds are imposed on directed graph $G = (V, E)$, that is, the in-degree of v by $b^- : V \to \mathbb{Z}_+$ and the out-degree of v by $b^+ : V \to \mathbb{Z}_+$ for all the vertices $v \in V$ (and $V_0 = \emptyset$ here). Construct G_D from G by replacing each vertex v by two, v_1 and v_2, connecting all the incoming arcs of v to v_1 while outgoing arcs to v_2. When arc orientations are ignored, G_D becomes an undirected bipartite graph. An approximate solution for G can be computed by applying an existing algorithm $\mathcal{A}$ to G_D, and taking v into a solution iff either v_1 or v_2 (or both) in G_D is chosen by $\mathcal{A}$. This way of reducing directed BDD to undirected one yields a 2ρ-approximation when $\mathcal{A}$ is a ρ-approximation because the optimum with respect to G_D is bounded by twice the optimum with respect to G. So, the fully bounded version of BDD is approximable within $4 + 2 \ln \max_{v \in V} \{b^-(v), b^+(v)\}$ by running the Okun-Barak algorithm as $\mathcal{A}$.

The reduction based approach above can be further refined by rebuilding the Okun-Barak approach within the current framework of submodular optimization. Consider the partition matroids (E, r_-) and (E, r_+), defined both on E, based on b^- and b^+, respectively. Here, $X \subseteq V$ is a solution iff $\delta(X)$ is spanning in both (E, r_-^d) and (E, r_+^d), where r_-^d and r_+^d are the dual rank functions of r_- and r_+, respectively. Define $f : 2^V \to \mathbb{Z}_+$ such that $f(X) = r_-^d(\delta(X)) + r_+^d(\delta(X))$. Then, f is nondecreasing and submodular, and $X \subseteq V$ is a solution iff $f(X) = r_-^d(\delta(X)) + r_+^d(\delta(X)) = r_-^d(E) + r_+^d(E) = f(V)$. Therefore, the problem can be reduced to SSC (V, f, w).

Let us adopt the following strategy of two stage approximation from [30]; first apply the local ratio method and then the greedy method for SSC.

1st stage. Suppose $d^-(v) > b^-(v)$ for some $v \in V$. Let $S^- = \{v\} \cup N^-(v)$ and consider the subgraph $G[S^-]$ of G induced by S^-. Since any solution for G including an optimal one must contain v or otherwise, at least $(d^-(v) - b^-(v))$ from $N^-(v)$, we have a valid inequality

$$(d^-(v) - b^-(v))x_v + \sum_{u \in N^-(v)} x_u \geq (d^-(v) - b^-(v))$$

for any v with $d^-(v) > b^-(v)$, where $\boldsymbol{x} \in \{0,1\}^V$ denotes an incidence vector of a vertex subset. We may thus apply the local ratio reduction to the weighted graph (G, w) as follows. Define the vertex weight $\bar{w}$ within $G[S^-]$ such that $\bar{w}(v) = d^-(v) - b^-(v)$ and $\bar{w}(u) = 1, \forall u \in N^-(v)$. Let $\rho = \min\{w(u)/\bar{w}(u) \mid u \in S^-\}$ and $S_0^- = \{u \in S^- \mid w(u) = \rho\bar{w}(u)\}$ so that $S_0^- \neq \emptyset$ and $w(u) - \rho\bar{w}(u) > 0, \forall u \in S^- - S_0^-$.

Suppose a solution C is computed for $G - S_0^- = G[V - S_0^-]$ under the weight $w - \rho\bar{w}$ defined on $V - S_0^-$. Then, $C \cup S_0^-$ is a solution for G and our algorithm

returns it. The ratio of this solution to the optimum, local to $(G[S^-], \rho\bar{w})$, is bounded by

$$\frac{\bar{w}(S^- \cap (C \cup S_0^-))}{d^-(v) - b^-(v)} \leq \frac{\bar{w}(S^-)}{d^-(v) - b^-(v)}$$
$$= \frac{2d^-(v) - b^-(v)}{d^-(v) - b^-(v)}$$
$$= 2 + \frac{1}{d^-(v)/b^-(v) - 1}.$$

So, if C is a p-approximation for $(G - S_0^-, w - \rho\bar{w})$, the approximation ratio of $C \cup S_0^-$ for (G, w) can be estimated by the following bound

$$\max\left\{p, 2 + \frac{1}{d^-(v)/b^-(v) - 1}\right\}. \qquad (2)$$

Thus, we apply the local ratio approximation to $G[S^-]$ and reduce to the problem on $G - S_0^-$ when such a vertex is found whose in-degree is large enough relative to its degree bound. Likewise, we may apply the local ratio reduction to out-degrees, and for $v \in V$ with $d^+(v) > b^+(v)$ we have the approximation ratio of $C \cup S_0^+$ for (G, w) bounded by

$$\max\left\{p, 2 + \frac{1}{d^+(v)/b^+(v) - 1}\right\} \qquad (3)$$

when $C \subseteq V - S_0^+$ is a p-approximation for the reduced problem on $G - S_0^+$. We apply these local ratio reductions as long as there remains a vertex v with high degree/degree-bound ratio; i.e., any v with $d^-(v)/b^-(v)$ or $d^+(v)/b^+(v)$ exceeding the threshold β.

2nd stage. We switch to the greedy algorithm for SSC (V, f, w) when vertices with high degree/degree-bound ratio are exhausted in the 1st stage. Here in the greedy mode, a vertex v with minimum $w(v)/f_C(\{v\})$ among the remaining vertices is repeatedly added to a solution set C as long as $f(C) < f(V)$.

We can show the following performance of this algorithm (details are omitted due to space limitations).

Theorem 2. *The (b^-, b^+)-BDD problem can be approximated within a factor of $2 + \max_{v \in V} \ln(b^-(v) + b^+(v))$ if $V_0 = \emptyset$.*

4   Partially Bounded Degree Deletion

4.1   Approximation Hardness

As was seen in the previous section, the Okun-Barak algorithm or its extension to directed graphs yields an $O(\log b_{\max})$-approximation. We observe here that such performance is possible only when all the degrees, both in-degrees and

out-degrees, are bounded, and if not, even at a single vertex, the performance becomes unbounded even if $b_{\max}$ is a fixed constant.

As already seen, the algorithm of Okun and Barak attains the best approximation bound of $2 + \ln b_{\max}$ for general b, by first applying the local ratio reduction to any v and its neighbors having high $d(v)$ to $b(v)$ ratio, and then by running the greedy approximation after $d(v)/b(v)$ becomes small enough for all the remaining vertices v. This approach is possible only when all the degrees are bounded since, if $d(v)$ is not bounded for some $v \in V$, there is now way of doing the local reduction at or around v with a reasonable ratio. Consider the following instance, for example: Let $G_b = (V_b, E_b)$ be a $(b-1)$-regular graph on n vertices, and V_b^c be a copy of V_b. Construct a graph $G = \{V_b \cup V_b^c \cup \{s\}, E\}$ from G_b by, besides having E_b entirely, connecting each vertex of V_b and its copy in V_b^c by an edge, and by having one more vertex s connected with every vertex in V_b by an edge. Since $d(v) = b+1$ if $v \in V_b$, $= 1$ if $v \in V_b^c$, and $= n$ if $v = s$, when the degree bound is set s.t. $b(v) = b$, $\forall v \in V_b$ and $b(v) = 1$, $\forall v \in V_b^c$ (and the degree of s is unbounded), the $d(v)/b(v)$ ratio can be made arbitrarily close to 1 at every $v \in V_b \cup V_b^c$, that there is nowhere to apply the local ratio reduction. So the algorithm simply runs the standard greedy approximation to G. Suppose that all the vertices in V_b are of heavy weight while the vertices in V_b^c are respectively assigned with weights of $1, 1/2, 1/3, \cdots, 1/n$ and s is assigned with $1 + \epsilon$. Then, the greedy algorithm outputs V_b^c as a solution of which weight is $\Theta(\log n)$ times that of the optimal solution $\{s\}$.

A more general approximation hardness of PBDD can be derived from that of Ek-VERTEX COVER (EkVC). This is the VERTEX COVER problem on k-uniform hypergraphs, and it is known to be NP-hard to approximate EkVC within a factor of $k - 1 - \epsilon$ for any $\epsilon > 0$ and $k \geq 3$ [12]. Let $H = (V, E_H)$ denote an instance of EkVC, i.e., a k-uniform hypergraph. Construct a bipartite instance $G = (V \cup E_H, E)$ of undirected PBDD from H s.t. $\{v, e_H\} \in E$, where $v \in V$ and $e_H \in E_H$, iff $v \in e_H$ in H. Set the weight of each vertex in E_H heavy enough that forces choice of vertices only from V and not from E_H. Notice that $d(e_H) = k$, $\forall e_H \in E_H$, and EkVC is reduced to undirected PBDD by setting the degree bound of $k - 1$ on each of them while leaving all the others (in V) unbounded. It follows from the approximation hardness of EkVC that

Theorem 3. *It is NP-hard to approximate PBDD, directed or undirected, within a factor of $b_{\max} - \epsilon$ for any $\epsilon > 0$ and $b_{\max} \geq 2$.*

4.2   Approximation Algorithm

Let us turn to an upper bound in approximation of b-PBDD, and next is a key lemma here:

Lemma 1. *For any minimal b-PBDD solution $X \subseteq V \cup V_0$ in $G = (V \cup V_0, E)$*

$$\sum_{v \in X} r^d(\delta(v)) \leq \max\{2, b_{\max} + 1\} r^d(E).$$

Proof. Omitted due to space limitations. □

It is immediate from Corollary 1 and Lemma 1 that

Theorem 4. *The b-PBDD problem can be approximated within* $\max\{2, b_{\max}+1\}$.

The bound of $\max\{2, b_{\max}+1\}$ given in Lemma 1 or Theorem 4 is tight even if $V_0 = \emptyset$. Suppose a graph G consists of the vertex set $X \cup Y \cup \{z\}$ and the edge set $E = X \times (Y \cup \{z\})$ s.t. $b(v) = 0$, $\forall v \in X \cup \{z\}$ and $b(v) = b_{\max}$, $\forall v \in Y$ for some integer $b_{\max}$. Clearly, X here is a minimal solution for b-PBDD.

Let x and y denote $|X|$ and $|Y|$, respectively. We have

$$r^d(E) = |E| - b(V) = x(y + 1) - b_{\max}y = (x - b_{\max})y + x$$

and

$$\sum_{v \in X} r^d(\delta(v)) = x(y + 1) = (x - b_{\max})y + x + b_{\max}y$$

since $r^d(\delta(v)) = y + 1$, $\forall v \in X$. Set $x = b_{\max} + 1$. Then,

$$\frac{\sum_{v \in X} r^d(\delta(v))}{r^d(E)} = \frac{x + y + b_{\max}y}{x + y} = 1 + \frac{b_{\max}y}{y + b_{\max} + 1}$$

and $\sum_{v \in X} r^d(\delta(v))/r^d(E)$ becomes arbitrarily close to $1 + b_{\max}$ as $y \to \infty$.

Suppose now that each vertex of G is weighted s.t. $w(v) = r^d(\delta(v))$. The algorithm PD may return X as an approximate solution, whose weight is $(b_{\max} + 1)(y + 1)$, whereas $\{v, z\}$ is an optimal solution for any $v \in X$ when y is large enough, whose weight is $d(v) + d(z) = y + b_{\max} + 2$. Therefore, the ratio of weight of X to the optimal weight is

$$\frac{(b_{\max} + 1)(y + 1)}{y + b_{\max} + 2} = 1 + b_{\max} - \frac{b_{\max}^2 + 2b_{\max} + 1}{y + b_{\max} + 2}$$

and it approaches arbitrarily close to $1 + b_{\max}$ as y becomes larger.

References

1. Balasundaram, B., Butenko, S., Hicks, I.V.: Clique relaxations in social network analysis: the maximum k-plex problem. Oper. Res. **59**(1), 133–142 (2011)
2. Bar-Yehuda, R.: Using homogeneous weights for approximating the partial cover problem. J. Algorithms **39**(2), 137–144 (2001)
3. Betzler, N., Bredereck, R., Niedermeier, R., Uhlmann, J.: On bounded-degree vertex deletion parameterized by treewidth. Discret. Appl. Math. **160**(1–2), 53–60 (2012)
4. Brešar, B., Krivoš-Belluš, R., Semanišin, G., Šparl, P.: On the weighted k-path vertex cover problem. Discret. Appl. Math. **177**, 14–18 (2014)
5. Brešar, B., Jakovac, M., Katrenič, J., Semanišin, G., Taranenko, A.: On the vertex k-path cover. Discret. Appl. Math. **161**(13–14), 1943–1949 (2013)

6. Brešar, B., Kardoš, F., Katrenič, J., Semanišin, G.: Minimum k-path vertex cover. Discret. Appl. Math. **159**(12), 1189–1195 (2011)
7. Camby, E., Cardinal, J., Chapelle, M., Fiorini, S., Joret, G.: A primal-dual 3-approximation algorithm for hitting 4-vertex paths. In: 9th International Colloquium on Graph Theory and Combinatorics, ICGT 2014, p. 61 (2014)
8. Case, B.M., Hedetniemi, S.T., Laskar, R.C., Lipman, D.J.: Partial domination in graphs. arXiv e-prints (2017)
9. Chauve, C., Tannier, E.: A methodological framework for the reconstruction of contiguous regions of ancestral genomes and its application to mammalian genomes. PLoS Comput. Biol. **4**(11), e1000234 (2008)
10. Chen, Z.-Z., Fellows, M., Fu, B., Jiang, H., Liu, Y., Wang, L., Zhu, B.: A linear kernel for co-path/cycle packing. In: Chen, B. (ed.) AAIM 2010. LNCS, vol. 6124, pp. 90–102. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14355-7_10
11. Das, A.: Partial domination in graphs. arXiv e-prints (2017)
12. Dinur, I., Guruswami, V., Khot, S., Regev, O.: A new multilayered PCP and the hardness of hypergraph vertex cover. SIAM J. Comput. **34**(5), 1129–1146 (2005)
13. Dinur, I., Safra, S.: On the hardness of approximating minimum vertex cover. Ann. Math. (2) **162**(1), 439–485 (2005)
14. Elomaa, T., Kujala, J.: Covering analysis of the greedy algorithm for partial cover. In: Elomaa, T., Mannila, H., Orponen, P. (eds.) Algorithms and Applications. LNCS, vol. 6060, pp. 102–113. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12476-1_7
15. Fellows, M.R., Guo, J., Moser, H., Niedermeier, R.: A generalization of Nemhauser and Trotter's local optimization theorem. J. Comput. Syst. Sci. **77**(6), 1141–1158 (2011)
16. Feng, Q., Wang, J., Li, S., Chen, J.: Randomized parameterized algorithms for P_2-packing and co-path packing problems. J. Comb. Optim. **29**(1), 125–140 (2015)
17. Fujito, T.: A unified approximation algorithm for node-deletion problems. Discret. Appl. Math. **86**(2–3), 213–231 (1998)
18. Fujito, T.: On approximation of the submodular set cover problem. Oper. Res. Lett. **25**(4), 169–174 (1999)
19. Fujito, T.: Approximating bounded degree deletion via matroid matching. In: Fotakis, D., Pagourtzis, A., Paschos, V.T. (eds.) CIAC 2017. LNCS, vol. 10236, pp. 234–246. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57586-5_20
20. Gandhi, R., Khuller, S., Srinivasan, A.: Approximation algorithms for partial covering problems. J. Algorithms **53**(1), 55–84 (2004)
21. Halperin, E., Srinivasan, A.: Improved approximation algorithms for the partial vertex cover problem. In: Jansen, K., Leonardi, S., Vazirani, V. (eds.) APPROX 2002. LNCS, vol. 2462, pp. 161–174. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45753-4_15
22. Jakovac, M., Taranenko, A.: On the k-path vertex cover of some graph products. Discret. Math. **313**(1), 94–100 (2013)
23. Karakostas, G.: A better approximation ratio for the vertex cover problem. ACM Trans. Algorithms **5**(4), art. no. 41 (2009)
24. Kardoš, F., Katrenič, J., Schiermeyer, I.: On computing the minimum 3-path vertex cover and dissociation number of graphs. Theoret. Comput. Sci. **412**(50), 7009–7017 (2011)
25. Khot, S., Regev, O.: Vertex cover might be hard to approximate to within $2 - \epsilon$. J. Comput. Syst. Sci. **74**(3), 335–349 (2008)

26. Kneis, J., Mölle, D., Rossmanith, P.: Partial vs. complete domination: t-dominating set. In: van Leeuwen, J., Italiano, G.F., van der Hoek, W., Meinel, C., Sack, H., Plášil, F. (eds.) SOFSEM 2007. LNCS, vol. 4362, pp. 367–376. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-69507-3_31
27. Lee, E.: Partitioning a graph into small pieces with applications to path transversal. In: Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, pp. 1546–1558 (2017)
28. Moser, H., Niedermeier, R., Sorge, M.J.: Exact combinatorial algorithms and experiments for finding maximum k-plexes. J. Comb. Optim. **24**(3), 347–373 (2012)
29. Newman, I., Sohler, C.: Every property of hyperfinite graphs is testable. SIAM J. Comput. **42**(3), 1095–1112 (2013)
30. Okun, M., Barak, A.: A new approach for approximating node deletion problems. Inform. Process. Lett. **88**(5), 231–236 (2003)
31. Orlovich, Y., Dolgui, A., Finke, G., Gordon, V., Werner, F.: The complexity of dissociation set problems in graphs. Discret. Appl. Math. **159**(13), 1352–1366 (2011)
32. Seidman, S.B., Foster, B.L.: A graph-theoretic generalization of the clique concept. J. Math. Soc. **6**(1), 139–154 (1978)
33. Slavík, P.: Improved performance of the greedy algorithm for partial cover. Inform. Process. Lett. **64**(5), 251–254 (1997)
34. Tu, J.: A fixed-parameter algorithm for the vertex cover P_3 problem. Inform. Process. Lett. **115**(2), 96–99 (2015)
35. Tu, J., Yang, F.: The vertex cover P_3 problem in cubic graphs. Inform. Process. Lett. **113**(13), 481–485 (2013)
36. Tu, J., Zhou, W.: A factor 2 approximation algorithm for the vertex cover P_3 problem. Inform. Process. Lett. **111**(14), 683–686 (2011)
37. Tu, J., Zhou, W.: A primal-dual approximation algorithm for the vertex cover P_3 problem. Theoret. Comput. Sci. **412**(50), 7044–7048 (2011)
38. Wolsey, L.A.: An analysis of the greedy algorithm for the submodular set covering problem. Combinatorica **2**(4), 385–393 (1982)
39. Xiao, M.: On a generalization of Nemhauser and Trotter's local optimization theorem. J. Comput. Syst. Sci. **84**, 97–106 (2017)
40. Yannakakis, M.: Node-deletion problems on bipartite graphs. SIAM J. Comput. **10**(2), 310–327 (1981)