Approximating the minimum independent dominating set in perturbed graphs

Weitian Tong*,† Randy Goebel*,‡ Guohui Lin*,§

November 3, 2013

Abstract

We investigate the minimum independent dominating set in perturbed graphs $\mathfrak{g}(G, p)$ of input graph $G = (V, E)$, obtained by negating the existence of edges independently with a probability $p > 0$. The minimum independent dominating set (MIDS) problem does not admit a polynomial running time approximation algorithm with worst-case performance ratio of $n^{1-\epsilon}$ for any $\epsilon > 0$. We prove that the size of the minimum independent dominating set in $\mathfrak{g}(G, p)$, denoted as $i(\mathfrak{g}(G, p))$, is asymptotically almost surely in $\Theta(\log |V|)$. Furthermore, we show that the probability of $i(\mathfrak{g}(G, p)) \geq \sqrt{\frac{4|V|}{p}}$ is no more than $2^{-|V|}$, and present a simple greedy algorithm of proven worst-case performance ratio $\sqrt{\frac{4|V|}{p}}$ and with polynomial expected running time.

Keywords: Independent set, independent dominating set, dominating set, approximation algorithm, perturbed graph, smooth analysis

1 Introduction

An *independent set* in a graph $G = (V, E)$ is a subset of vertices that are pair-wise non-adjacent to each other. The independence number of G, denoted by $\alpha(G)$, is the size of a maximum independent set in G. One close notion to independent set is the *dominating set*, which refers to a subset of

*Department of Computing Science, University of Alberta. Edmonton, Alberta T6G 2E8, Canada.
†Email: weitian@ualberta.ca
‡Email: rgoebel@ualberta.ca
§Correspondence author. Email: guohui@ualberta.ca

vertices such that every vertex of the graph is either in the subset or is adjacent to some vertex in the subset. In fact, an independent set becomes a dominating set if and only if it is maximal. The size of a minimum independent dominating set of G is denoted by $i(G)$, while the domination number of G, or the size of a minimum dominating set of G, is denoted by $\gamma(G)$. It follows that $\gamma(G) \leq i(G) \leq \alpha(G)$.

Another related notion is the (vertex) *coloring* of G, in which two adjacent vertices must be colored differently. Note that any subset of vertices colored the same in a coloring of G is necessarily an independent set. The *chromatic number* $\chi(G)$ of G is the minimum number of colors in a coloring of G. Clearly, $\alpha(G) \cdot \chi(G) \geq |V|$.

The independence number $\alpha(G)$ and the domination number $\gamma(G)$ (and the chromatic number $\chi(G)$) have received numerous studies due to their central roles in graph theory and theoretical computer science. Their exact values are NP-hard to compute [4], and hard to approximate. Raz and Safra showed that the domination number cannot be approximated within $(1-\epsilon)\log|V|$ for any fixed $\epsilon > 0$, unless NP $\subset$ DTIME($|V|^{\log\log|V|}$) [9, 3]; Zuckerman showed that neither the independence number nor the chromatic number can be approximated within $|V|^{1-\epsilon}$ for any fixed $\epsilon > 0$, unless P = NP [14]; for $i(G)$, Halldórsson proved that it is also hard to approximate within $|V|^{1-\epsilon}$ for any fixed $\epsilon > 0$, unless NP $\subset$ DTIME($2^{o(|V|)}$) [5].

The above inapproximability results are for the worst case. For analyzing the average case performance of approximation algorithms, a probability distribution of the input graphs must be assumed and the most widely used distribution of graphs on n vertices is the random graph $G(n,p)$, which is a graph on n vertices, and each edge is chosen to be an edge of G independently and with a probability p, where $0 \leq p = p(n) \leq 1$. A graph property holds *asymptotically almost surely* (a.a.s.) in $G(n,p)$ if the probability that a graph drawn according to the distribution $G(n,p)$ has the property tends to 1 as n tends to infinity [1].

Let $\mathbb{L}n = \log_{1/(1-p)} n$. Bollobás [2] and Łuczak [7] showed that a.a.s. $\chi(G(n,p)) = (1+o(1))n/\mathbb{L}n$ for a constant p and $\chi(G(n,p)) = (1+o(1))np/(2\ln(np))$ for $c/n \leq p(n) \leq o(1)$ where c is a constant. It follows from these results that a.a.s. $\alpha(G(n,p)) = (1-o(1))\mathbb{L}n$ for a constant p and $\alpha(G(n,p)) = (1-o(1))2\ln(np)/p$ for $C/n \leq p \leq o(1)$. The greedy algorithm, which colors vertices of $G(n,p)$ one by one and picks each time the first available color for a current vertex, is known to produce a.a.s. in $G(n,p)$ with $p \geq n^{\epsilon-1}$ a coloring whose number of colors is larger than the $\chi(G(n,p))$ by only a constant factor (see Ch. 11 of the monograph of Bollobás [1]). Hence the largest color class produced by the greedy algorithm is a.a.s. smaller than $\alpha(G(n,p))$ only by a constant factor.

For the domination number $\gamma(G(n,p))$, Wieland and Godbole showed that a.a.s. it is equal to either $\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))\rfloor + 1$ or $\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))\rfloor + 2$, for a constant p or a suitable function $p = p(n)$ [13]. It follows that a.a.s. $i(G(n,p)) \geq \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))\rfloor + 1$. Recently, Wang proved for $i(G(n,p))$ an a.a.s. upper bound of $\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))\rfloor + k + 1$, where $k = \max\{1, \mathbb{L}2\}$ [12].

Average case performance analysis of an approximation algorithm over random instances could be inconclusive, because the random instances usually have very special properties that distinguish them from real-world instances. For instance, for a constant p, the random graph $G(n,p)$ is expected to be dense. On the other hand, an approximation algorithm performs very well on most random instances can fail miserably on some "hard" instances. For instance, it has been shown by Kučera [6] that for any fixed $\epsilon > 0$ there exists a graph G on n vertices for which, even after a random permutation of vertices, the greedy algorithm produces a.a.s. a coloring using at least $n/\log_2 n$ colors, while $\chi(G) \leq n^{\epsilon}$. To overcome this, Spielman and Teng [10] introduced the smoothed analysis. This new analysis is a hybrid of the worst-case and the average-case analyses, and it inherits the advantages of both, by measuring the expected performance of the algorithm under slight random perturbations of the worst-case inputs. If the smoothed complexity of an algorithm is low, then it is unlikely that the algorithm will take long time to solve practical instances whose data are subject to slight noises and imprecision.

Formally, let A be an algorithm we want to analyze and Q be the quality measurement. Without loss of generality, we assume the larger Q the worse the algorithm A is, such as Q being the running time of algorithm A. Given x an input instance, $Q(A,x)$ denotes the performance of algorithm A on x. Let U denote the set of instances and U_n denote the subset of size-n instances. The smoothed performance measure of algorithm A under Q essentially is the worst of the expected performances of algorithm A in a small neighborhood of an instance, over all instances, and is defined precisely as

$$\mathcal{M}_{\text{smooth}}(n,\sigma) = \sup_{x \in U_n} E_r[Q(A, x + \sigma \cdot |x| \cdot r)],$$

where r is the random noise and σ is the perturbation parameter measures the strength of noise. Intuitively, $\sigma \cdot |x| \cdot r$ is the perturbation on instance x, in which the the magnitude of the perturbation is related to the magnitude of input. We have the following observations. If the smoothed performance measure of algorithm A under Q is good with some relatively small σ and some reasonable random model for r, then it is unlikely algorithm A would perform very bad in real world applications under quality measure Q, because real world instances are often subject to a slight amount of noise, especially when they are obtained from measurements of real-world phenomena. A classic example is the Simplex method for linear programming. Simplex method is a very practical algorithm, but it has exponential running time in the worst-case. Spielman and Teng [10] had shown that Simplex method has polynomial smoothed running time, which explains the above phenomenon perfectly. Though the smoothed analysis concept was originally introduced for the complexity of algorithms, we extend its idea to depict the essential properties of problems.

In this paper, we study the approximability of the minimum independent dominating set (MIDS) problem under the smoothed analysis, and we present a simple deterministic greedy algorithm beating the strong inapproximability bound of $n^{1-\epsilon}$, with polynomial expected running time. The MIDS problem, and the closely related independent set and dominating set problems, have important applications in wireless networks, and have been studied extensively in the literature. Our probabilistic model is the smoothed extension of random graph $G(n,p)$ (also called semi-random graphs in [8]), proposed by Spielman and Teng [11]: given a graph $G = (V, E)$, we define its perturbed graph $\mathfrak{g}(G,p)$ by negating the existence of edges independently with a probability of $p > 0$. That is, $\mathfrak{g}(G,p)$ has the same vertex set V as G but it contains edge e with probability p_e, where $p_e = 1-p$ if $e \in E$ or otherwise $p_e = p$. For sufficiently large p, Manthey and Plociennik presented an algorithm approximating the independence number $\alpha(\mathfrak{g}(G,p))$ with a worst-case performance ratio $O(\sqrt{np})$ and with polynomial expected running time [8].

Re-define $\mathbb{L}n = \log_{1/p} n$. We first prove on $\gamma(\mathfrak{g}(G,p))$, and thus on $i(\mathfrak{g}(G,p))$ as well, an a.a.s. lower bound of $\mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))$ if $p > \frac{1}{n}$. We then prove on $\alpha(\mathfrak{g}(G,p))$, and thus on $i(\mathfrak{g}(G,p))$ as well, an a.a.s. upper bound of $2\ln n/p$ if $p < \frac{1}{2}$ or $2\ln n/(1-p)$ otherwise. Given that the a.a.s. values of $\alpha(G(n,p))$ and $i(G(n,p))$ in random graph $G(n,p)$, our upper bound comes with no big surprise; nevertheless, our upper bound is derived by a direct counting process which might be interesting by itself. Furthermore, we extend our counting techniques to prove on $i(\mathfrak{g}(G,p))$ a tail bound that, when $4\ln^2 n/n < p \leq \frac{1}{2}$, $\Pr[i(\mathfrak{g}(G,p)) \geq \sqrt{4n/p}] \leq 2^{-n}$. We then present a simple greedy algorithm to approximate $i(\mathfrak{g}(G,p))$, and prove that its worst case performance ratio is $\sqrt{4n/p}$ and its expected running time is polynomial.

2 A.a.s. bounds on the independent domination number

We need the following several facts.

Fact 1 $e^{\frac{x}{1+x}} \leq 1 + x \leq e^x$ *holds for all* $x \in [-1, 1]$.

Fact 2 $\left(\frac{n}{r}\right)^r \leq \binom{n}{r} \leq \left(\frac{ne}{r}\right)^r$ *holds for all* $r = 0, 1, 2, \ldots, n$.

Fact 3 (Jensen's Inequality) *For a real convex function* $f(x)$*, numbers* x_1*,* x_2*,* $\ldots$*,* x_n *in its domain, and positive weights* a_i*,* $f\left(\frac{\sum a_i x_i}{\sum a_i}\right) \leq \frac{\sum a_i f(x_i)}{\sum a_i}$*; the inequality is reversed if* $f(x)$ *is concave.*

Given any graph $G = (V, E)$, let $\mathfrak{g}(G, p)$ denote its perturbed graph, which has the same vertex set V as G and contains edge e with a probability of

$$p_e = \begin{cases} 1 - p, & \text{if } e \in E, \\ p, & \text{otherwise.} \end{cases}$$

2.1 An a.a.s. lower bound

Recall that $\gamma(\mathfrak{g}(G, p))$ and $i(\mathfrak{g}(G, p))$ are the domination number and the independent domination number of $\mathfrak{g}(G, p)$, respectively. Also, $\mathbb{L}n = \log_{1/p} n$.

Theorem 4 *For any graph* $G = (V, E)$ *and* $\frac{1}{n} < p \leq 1$*, a.a.s.*

$$\gamma(\mathfrak{g}(G, p)) \geq \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n)).$$

PROOF. Let $\mathcal{S}_r$ be the collection of all r-subsets of vertices in $\mathfrak{g}(G, p)$, and these $\binom{n}{r}$ sets of $\mathcal{S}_r$ are ordered in some way. Define I_j^r as a boolean variable to indicate whether or not the j-th r-subset of $\mathcal{S}_r$, V_j, is a dominating set; set $X_r = \sum_j I_j^r$.

Clearly, $\gamma(\mathfrak{g}(G, p)) < r$ implies that there are size-r dominating sets. Therefore,

$$\Pr[\gamma(\mathfrak{g}(G, p)) < r] \leq \Pr[X_r \geq 1] \leq E(X_r),$$

where $E(X_r)$ is the expected value of X_r. (We abuse the notation E a little, but its meaning should be clear at every occurrence.)

For the j-th r-subset V_j, let E_j be the subset of induced edges on V_j from the original graph $G = (V, E)$; let $V_j^c = V - V_j$, the complement subset of vertices. Also, for each vertex $u \in V_j^c$, define $E(u, V_j) = \{(u, v) \in E \mid v \in V_j\}$, and its size $n_{uj} = |E(u, V_j)|$. Using Fact 1, we can estimate $E(X_r)$ as follows:

$$\begin{aligned}
E(X_r) = \sum_{j=1}^{\binom{n}{r}} E(I_j^r) &= \sum_{j=1}^{\binom{n}{r}} \prod_{u \in V_j^c} \left(1 - \prod_{v \in V_j} \left(1 - p_{(u,v)}\right)\right) \\
&\leq \sum_{j=1}^{\binom{n}{r}} \prod_{u \in V_j^c} \exp\left(-\prod_{v \in V_j} \left(1 - p_{(u,v)}\right)\right) \\
&= \sum_{j=1}^{\binom{n}{r}} \exp\left(-\sum_{u \in V_j^c} \prod_{v \in V_j} \left(1 - p_{(u,v)}\right)\right) \\
&= \sum_{j=1}^{\binom{n}{r}} \exp\left(-\sum_{u \in V_j^c} p^{n_{uj}} (1-p)^{r - n_{uj}}\right) \\
&= \sum_{j=1}^{\binom{n}{r}} \exp\left(-\sum_{u \in V_j^c} \left(\frac{p}{1-p}\right)^{n_{uj}} (1-p)^r\right).
\end{aligned}$$

Since function $f(x) = (\frac{p}{1-p})^x$ is convex in the domain $[0, n]$, by Jensen's Inequality, the above becomes

$$E(X_r) \leq \sum_{j=1}^{\binom{n}{r}} \exp\left(-\left(\frac{p}{1-p}\right)^{\frac{1}{n-r} \sum_{u \in V_j^c} n_{uj}} (n-r)(1-p)^r\right).$$

Since function $g(x) = e^{-a^x b}$ with $a = (\frac{p}{1-p})^{\frac{1}{n-r}}$ and $b = (n-r)(1-p)^r$ is concave in the domain $[0, n^2]$, again by Jensen's Inequality, we further have

$$E(X_r) \leq \binom{n}{r} \exp\left(-\left(\frac{p}{1-p}\right)^{\frac{1}{(n-r)\binom{n}{r}} \sum_{j=1}^{\binom{n}{r}} \sum_{u \in V_j^c} n_{uj}} (n-r)(1-p)^r\right). \tag{2.1}$$

Recall that n_{uj} is number of edges in the original graph $G = (V, E)$ between u and vertices of V_j. Each edge $e \in E$ is thus counted towards the quantity $\left(\sum_{j=1}^{\binom{n}{r}} \sum_{u \in V_j^c} n_{uj}\right)$ exactly $2\binom{n-2}{r-1}$

times. That is,

$$\sum_{j=1}^{\binom{n}{r}} \sum_{u \in V_j^c} n_{uj} = 2\binom{n-2}{r-1}|E| = \frac{\binom{n}{r} r(n-r)|E|}{\binom{n}{2}}. \tag{2.2}$$

Using Eq. (2.2), Fact 2 and $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))$, Eq. (2.1) becomes

$$\begin{aligned} E(X_r) &\leq \binom{n}{r} \exp\left(-\left(\frac{p}{1-p}\right)^{\frac{r|E|}{\binom{n}{2}}} (n-r)(1-p)^r\right) \\ &\leq \binom{n}{r} \exp\left(-\left(\frac{p}{1-p}\right)^r (n-r)(1-p)^r\right) \\ &\leq \left(\frac{ne}{r}\right)^r \exp\Big(-p^r(n-r)\Big) \\ &\leq \exp\left(r\ln n + r - r\ln r - \frac{(\mathbb{L}n)(\ln n)}{n}(n-r)\right) \\ &= \exp((\mathbb{L}n)(\ln n) - \mathbb{L}((\mathbb{L}n)(\ln n))\ln n + r - r\ln r - (\mathbb{L}n)(\ln n) + r(\mathbb{L}n)(\ln n)/n) \\ &= \exp\left(-\mathbb{L}((\mathbb{L}n)(\ln n))\ln n - r\left(\ln r - (\mathbb{L}n)(\ln n)/n - 1\right)\right) \\ &\leq \exp\left(-\mathbb{L}((\mathbb{L}n)(\ln n))\ln n - r\left(\ln r - 2\right)\right). \end{aligned} \tag{2.3}$$

f The right hand side in Eq. (2.3) approaches 0 when $n \to +\infty$. Since $p > \frac{1}{n}$ guarantees $r \geq 1$, $\mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n))$ is an a.a.s. lower bound on $\gamma(\mathfrak{g}(G,p))$. This proves the theorem. □

Since $\Pr[i(\mathfrak{g}(G,p)) < r] \leq \Pr[\gamma(\mathfrak{g}(G,p)) < r]$, we have the following corollary:

Corollary 1 *For any graph $G = (V, E)$ and $\frac{1}{n} < p \leq 1$, a.a.s.*

$$i(\mathfrak{g}(G,p)) \geq \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\ln n)).$$

2.2 An a.a.s. upper bound

Recall that $\alpha(\mathfrak{g}(G,p))$ is the independence number of $\mathfrak{g}(G,p)$.

Theorem 5 *For any graph $G = (V, E)$, a.a.s.*

$$\alpha(\mathfrak{g}(G,p)) \leq \begin{cases} \frac{2\ln n}{p}, & \text{if } p \in (\frac{2\ln n}{n}, \frac{1}{2}], \\ \frac{2\ln n}{1-p}, & \text{if } p \in [\frac{1}{2}, 1 - \frac{2\ln n}{n}). \end{cases}$$

PROOF. Let $\mathcal{S}_r$ be the collection of all r-subsets of vertices in $\mathfrak{g}(G,p)$, and these $\binom{n}{r}$ sets of $\mathcal{S}_r$ are ordered in some way. Define I_j^r as a boolean variable to indicate whether or not the j-th r-subset of $\mathcal{S}_r$ is an independent set; set $X_r = \sum_j I_j^r$. Since $\alpha(\mathfrak{g}(G,p)) > r$ implies that there is at least one independent r-subset, i.e. $X_r > 0$, the probability of the event $\alpha(\mathfrak{g}(G,p)) > r$ is less than or equal to the probability of the event $X_r > 0$, i.e.

$$\Pr[\alpha(\mathfrak{g}(G,p)) > r] \leq \Pr[X_r > 0].$$

On the other hand, let A_j^r denote the event $I_j^r = 0$, i.e. the j-th r-subset is not independent. It follows that $X_r = 0$ is equivalent to the joint event $\cap_j A_j^r$, i.e.

$$\Pr[X_r = 0] = \Pr[\cap_j A_j^r] \geq \prod_j \Pr[A_j^r] = \prod_j (1 - \Pr[I_j^r = 1]).$$

Therefore, we have

$$\Pr[\alpha(\mathfrak{g}(G,p)) > r] \leq 1 - \prod_j (1 - \Pr[I_j^r = 1]). \tag{2.4}$$

Let E_j^r denote the subset of edges of $\mathfrak{g}(G,p)$, each of which connects two vertices in the j-th r-subset of $\mathcal{S}_r$. Note that $|E_j^r| \in [0, \binom{r}{2}]$. Among all the edges of E_j^r, assume there are n_j^r of them coming from the original edge set E of G. It follows that

$$\Pr[I_j^r = 1] = \prod_{e \in E_j^r} (1 - p_e) = \left(\frac{p}{1-p}\right)^{n_j^r} (1-p)^{\binom{r}{2}}.$$

Using this and Fact 1 in Eq. (2.4) gives us

$$\begin{aligned}
\Pr[\alpha(\mathfrak{g}(G,p)) > r] &\leq 1 - \prod_j (1 - \Pr[I_j^r = 1]) \\
&\leq 1 - \prod_{j=1}^{\binom{n}{r}} \exp\left(-\frac{\Pr[I_j^r = 1]}{1 - \Pr[I_j^r = 1]}\right) \\
&= 1 - \exp\left(-\sum_{j=1}^{\binom{n}{r}} \frac{\Pr[I_j^r = 1]}{1 - \Pr[I_j^r = 1]}\right) \\
&= 1 - \exp\left(-\sum_{j=1}^{\binom{n}{r}} \frac{\left(\frac{p}{1-p}\right)^{n_j^r} (1-p)^{\binom{r}{2}}}{1 - \left(\frac{p}{1-p}\right)^{n_j^r} (1-p)^{\binom{r}{2}}}\right).
\end{aligned} \tag{2.5}$$

Consider the function $f(x) = \frac{a^x b}{1 - a^x b}$ in Eq. (2.5), where $a = \frac{p}{1-p} > 0$, $b = (1-p)^{\binom{r}{2}} \in (0,1)$, and $0 \leq x \leq \binom{r}{2}$. Since its derivative

$$f'(x) = \frac{a^x b \ln a}{(1 - a^x b)^2} \begin{cases} < 0, & \text{if } a < 1, \\ = 0, & \text{if } a = 1, \\ > 0, & \text{if } a > 1, \end{cases}$$

$f(x)$ is strictly decreasing if $a < 1$, or strictly increasing if $a > 1$. Therefore, the maximum value of function $f(x)$ is achieved at $x = 0$ if $a \le 1$, or at $x = \binom{r}{2}$ if $a \ge 1$.

When $p \le \frac{1}{2}$, that is $a = \frac{p}{1-p} \le 1$, Eq. (2.5) becomes

$$\begin{aligned} \Pr[\alpha(\mathfrak{g}(G,p)) > r] &\le 1 - \exp\left(-\sum_{j=1}^{\binom{n}{r}} \frac{(1-p)^{\binom{r}{2}}}{1-(1-p)^{\binom{r}{2}}}\right) \\ &= 1 - \exp\left(-\binom{n}{r}\frac{(1-p)^{\binom{r}{2}}}{1-(1-p)^{\binom{r}{2}}}\right). \end{aligned} \tag{2.6}$$

To prove $\Pr[\alpha(\mathfrak{g}(G,p)) > r] \to 0$ as $n \to +\infty$, we only need to prove that $\binom{n}{r}\frac{(1-p)^{\binom{r}{2}}}{1-(1-p)^{\binom{r}{2}}} \to 0$ as $n \to +\infty$. Using Fact 2, we have

$$\binom{n}{r}\frac{(1-p)^{\binom{r}{2}}}{1-(1-p)^{\binom{r}{2}}} = \frac{\binom{n}{r}}{\left(\frac{1}{1-p}\right)^{\binom{r}{2}} - 1} \le \frac{\left(\frac{ne}{r}\right)^r}{\left(\frac{1}{1-p}\right)^{\binom{r}{2}} - 1}. \tag{2.7}$$

Setting $r = 2\ln n/p$. We see that $r \to +\infty$ as $n \to +\infty$. On the other hand, when r is large enough, we have

$$\left(\frac{1}{1-p}\right)^{\binom{r}{2}} - 1 = \left(\frac{1}{1-p}\right)^{\binom{r}{2}}(1 - o(1)). \tag{2.8}$$

Using Eq. (2.8) and Fact 1, when n is sufficiently large, Eq. (2.7) becomes

$$\begin{aligned}
\binom{n}{r}\frac{(1-p)^{\binom{r}{2}}}{1-(1-p)^{\binom{r}{2}}} &\le \frac{\left(\frac{ne}{r}\right)^r}{\left(\frac{1}{1-p}\right)^{\binom{r}{2}}}(1+o(1)) = \left(\frac{ne}{r\left(\frac{1}{1-p}\right)^{\frac{r-1}{2}}}\right)^r (1+o(1)) \\
&= \left(\frac{ne}{r\left(1+\frac{p}{1-p}\right)^{\frac{r-1}{2}}}\right)^r (1+o(1)) \\
&\le \left(\frac{ne}{r\exp\left(\frac{\frac{p}{1-p}}{1+\frac{p}{1-p}}\cdot\frac{r-1}{2}\right)}\right)^r (1+o(1)) \\
&= \left(\frac{ne}{r\exp\left(p\cdot\frac{r-1}{2}\right)}\right)^r (1+o(1)) \\
&= \left(\frac{ne^{1+\frac{p}{2}}}{re^{\frac{rp}{2}}}\right)^r (1+o(1)) &(2.9) \\
&= \left(\frac{e^{1+\frac{p}{2}}}{r}\right)^r (1+o(1)) \\
&\le \left(\frac{e^{\frac{5}{4}}}{r}\right)^r (1+o(1)). &(2.10)
\end{aligned}$$

The quantity $\left(\frac{e^{\frac{5}{4}}}{r}\right)^r$ in Eq. (2.10) is less than 0.5^r when n is sufficiently large, the latter approaches 0 when $n \to +\infty$. This proves that when $p \le \frac{1}{2}$, $\Pr[\alpha(\mathfrak{g}(G,p)) > r] \to 0$ as $n \to +\infty$. That is, when $p \le \frac{1}{2}$, a.a.s. $\alpha(\mathfrak{g}(G,p)) \le 2\ln n/p$.

When $p \ge \frac{1}{2}$, that is $a = \frac{p}{1-p} \ge 1$, $q = 1-p \le \frac{1}{2}$ and exactly the same argument as when $p \le \frac{1}{2}$ applies by replacing p with $1-q$, which shows that a.a.s. $\alpha(\mathfrak{g}(G,p)) \le 2\ln n/(1-p)$. This proves the theorem. □

Since $\alpha(\mathfrak{g}(G,p)) \ge i(\mathfrak{g}(G,p))$, $\Pr[i(\mathfrak{g}(G,p)) > r] \le \Pr[\alpha(\mathfrak{g}(G,p)) > r]$ and thus we have the following corollary:

Corollary 2 *For any graph $G = (V,E)$, a.a.s.*

$$i(\mathfrak{g}(G,p)) \le \begin{cases} \frac{2\ln n}{p}, & \text{if } p \in (\frac{2\ln n}{n}, \frac{1}{2}], \\ \frac{2\ln n}{1-p}, & \text{if } p \in [\frac{1}{2}, 1-\frac{2\ln n}{n}). \end{cases}$$

3 A tail bound on the independent domination number

Theorem 6 *For any graph $G = (V, E)$ and $p \in (\frac{4\ln^2 n}{n}, \frac{1}{2}]$,*

$$\Pr[i(\mathfrak{g}(G,p)) \geq \sqrt{\frac{4n}{p}}] \leq \Pr[\alpha(\mathfrak{g}(G,p)) \geq \sqrt{\frac{4n}{p}}] \leq 2^{-n}.$$

PROOF. The proof of this theorem flows exactly the same as the proof of Theorem 5. In fact, with $p \leq \frac{1}{2}$, we have both Eq. (2.6) and Eq. (2.7) hold. Different from the proof of Theorem 5 where $r = 2\ln n/p$, we have now $r = \sqrt{\frac{4n}{p}} \geq 2\ln n/p$ and therefore Eq. (2.8) holds as well. Again, using Eq. (2.8) and Fact 1, when n is sufficiently large, Eq. (2.9) still holds. It then follows from Fact 1 that Eq. (2.6) becomes

$$\begin{aligned}\Pr[i(\mathfrak{g}(G,p)) \geq r] &\leq \Pr[\alpha(\mathfrak{g}(G,p)) \geq r] \\ &\leq 1 - \exp\left(-\left(\frac{ne^{1+\frac{p}{2}}}{re^{\frac{rp}{2}}}\right)^r (1+o(1))\right). \qquad (3.1)\end{aligned}$$

Using $r = \sqrt{\frac{4n}{p}}$, we prove in the following that $\left(\frac{ne^{1+\frac{p}{2}}}{re^{\frac{rp}{2}}}\right)^r (1+o(1)) = o(1)$. And consequently by Fact 1 again and $r = \sqrt{\frac{4n}{p}} \geq \sqrt{8n}$, Eq. (3.1) becomes

$$\begin{aligned}\Pr[i(\mathfrak{g}(G,p)) \geq r] &\leq \left(\frac{ne^{1+\frac{p}{2}}}{re^{\frac{rp}{2}}}\right)^r (1+o(1)) \\ &\leq \frac{e}{2}\left(\frac{ne^{1+\frac{p}{2}}}{re^{\frac{rp}{2}}}\right)^r \\ &= \frac{e}{2}\exp\left(-r\left(\ln r + \frac{1}{2}rp - \ln n - 1 - \frac{p}{2}\right)\right) \\ &= \frac{e}{2}\exp\left(-r\left(\ln r + \frac{1}{4}rp - \ln n - 1 - \frac{p}{2}\right) - \frac{1}{4}r^2 p\right) \\ &= \frac{e}{2}\exp\left(-r\left(\ln r + \frac{1}{4}rp - \ln n - 1 - \frac{p}{2}\right) - n\right). \qquad (3.2)\end{aligned}$$

The quantity $\left(\ln r + \frac{1}{4}rp - \ln n - 1 - \frac{p}{2}\right)$ in Eq. (3.2) is non-negative when $n \geq 2$, since

$$\begin{aligned}\ln r + \frac{1}{4}rp - \ln n - 1 - \frac{p}{2} &\geq \frac{1}{2}\ln(8n) + \frac{1}{4}\sqrt{4np} - \ln n - 1 - \frac{1}{4} \\ &\geq \frac{1}{2}\ln(8n) + \frac{1}{4}\sqrt{4n \cdot \frac{4\ln^2 n}{n}} - \ln n - 1 - \frac{1}{4} \\ &= \frac{1}{2}\left(\ln(8n) - \frac{5}{2}\right) \\ &\geq 0.\end{aligned}$$

It follows that Eq. (3.2) becomes

$$\begin{aligned}\Pr[i(\mathfrak{g}(G,p)) \geq r] &\leq \frac{e}{2}\exp\left(-r\left(\ln r + \frac{1}{4}rp - \ln n - 1 - \frac{p}{2}\right) - n\right)\\ &\leq \frac{e}{2}\, e^{-n}\\ &< 2^{-n}.\end{aligned}$$

This proves the theorem. □

4 Approximating the independent domination number

We present next a simple algorithm, denoted as *Approx-IDS*, for computing an independent dominating set in $\mathfrak{g}(G,p)$. In the first phase, algorithm Approx-IDS repeatedly picks a maximum degree vertex and updates the graph by deleting the picked vertex and all its neighbors; it terminates when there is no more vertex and returns a subset I of V. If $|I| \leq \sqrt{\frac{4n}{p}}$, algorithm Approx-IDS terminates and outputs I; otherwise it moves into the second phase. In the second phase, algorithm Approx-IDS performs an exhaustive search over all subsets of V, and returns the minimum independent dominating set I^*.

Theorem 7 *For any graph* $G = (V,E)$ *and* $p \in (\frac{4\ln^2 n}{n}, \frac{1}{2}]$, *algorithm* Approx-IDS *is a* $\sqrt{\frac{4n}{p}}$-*approximation to the* MIDS *problem on the perturbed graph* $\mathfrak{g}(G,p)$*, and it has polynomial expected running time.*

PROOF. Note that $i(\mathfrak{g}(G,p)) \geq 1$. The subset I of V computed by algorithm Approx-IDS is a dominating set, since every vertex of V is either in I, or is a neighbor of some vertex in I. Also, no two vertices of I can be adjacent, since otherwise one would be removed in the iteration its neighbor was picked by the algorithm. Therefore, I is an independent dominating set of $\mathfrak{g}(G,p)$. It follows that if algorithm Approx-IDS terminates after the first phase, $|I| \leq \sqrt{\frac{4n}{p}} \cdot i(\mathfrak{g}(G,p))$. Also clearly the first phase takes $O(n^3)$ time.

In the second phase, a maximum of 2^n subsets of V are examined by the algorithm. Since checking each of them to be an independent dominating set or not takes no more than $O(n^2)$ time, the overall running time is $O(2^n n^2)$. Note that this phase returns I^* with $|I^*| = i(\mathfrak{g}(G,p))$. As $\alpha(\mathfrak{g}(G,p)) \geq |I| > \sqrt{\frac{4n}{p}}$, Theorem 6 tells that the probability of executing this second phase is no more than 2^{-n}. Therefore, the expected running time of the second phase is $O(n^2)$. This proves the theorem. $\square$

5 Conclusions

We have performed a smooth analysis for approximating the minimum independent dominating set problem. The probabilistic model we used is the perturbed graph $\mathfrak{g}(G, p)$ of the input graph $G = (V, E)$ [11]. We have proved a.a.s. bounds and a tail bound on the independent domination number of $\mathfrak{g}(G, p)$, and presented an algorithm with the worst-case performance ratio of $\sqrt{\frac{4|V|}{p}}$ and with polynomial expected running time.

Acknowledgement

This research was supported in part by NSERC.

References

[1] B. Bollobás. *Random Graphs.* Academic Press, New York, 1985.

[2] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8:49–55, 1988.

[3] U. Feige. A threshold of for approximating set cover. *Journal of the ACM*, 45:634–652, 1998.

[4] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness.* W. H. Freeman and Company, San Francisco, 1979.

[5] M. M. Halldórsson. Approximating the minimum maximal independence number. *Information Processing Letters*, 46:169–172, 1993.

[6] L. Kučera. The greedy coloring is a bad probabilistic algorithm. *Journal of Algorithms*, 12:674–684, 1991.

[7] T. Łuczak. The chromatic number of random graphs. *Combinatorica*, 11:45–54, 1991.

[8] B. Manthey and K. Plociennik. Approximating independent set in perturbed graphs. *Discrete Applied Mathematics*, 162:1761–1768, 2013.

[9] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)*, pages 475–484, 1997.

[10] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51:385–463, 2004.

[11] D. A. Spielman and S.-H. Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice. *Communications of the ACM*, 52:76–84, 2009.

[12] C. Wang. The independent domination number of random graph. *Utilitas Mathematica*, 82:161–166, 2010.

[13] B. Wieland and A. P. Godbole. On the domination number of a random graph. *The Electronic Journal of Combinatorics*, 8:#R37, 2001.

[14] D. Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3:103–128, 2007.