# On Randomized Greedy Matchings

by

Zevi Miller
Department of Mathematics and Statistics
Miami University
Oxford, Ohio 45056

Dan Pritikin*
Department of Mathematics and Statistics
Miami University
Oxford, Ohio 45056

*Part of this author's work was completed during a sabbatical provided by Miami University.

## Abstract

We analyze a randomized greedy matching algorithm (RGA) aimed at producing a matching with a large number of edges in a given weighted graph. RGA was first introduced and studied by Dyer and Frieze in [3] for unweighted graphs. In the weighted version, at each step a new edge is chosen from the remaining graph with probability proportional to its weight, and is added to the matching. The two vertices of the chosen edge are removed, and the step is repeated until there are no edges in the remaining graph. We analyze the expected size $\mu(G)$ of the number of edges in the output matching produced by RGA, when RGA is repeatedly applied to the same graph G. Let $r(G) = \mu(G)/m(G)$, where $m(G)$ is the maximum number of edges in a matching in G.

For a class $\mathcal{G}$ of graphs, let $\rho(\mathcal{G})$ be the infimum of the values r(G) over all graphs G in $\mathcal{G}$ (i.e. $\rho$ is the "worst" performance ratio of RGA restricted to the class $\mathcal{G}$). Our main results are bounds for $\mu$,r, and $\rho$. For example, the following results improve or generalize similar results obtained in [3] for the unweighted version of RGA;

$r(G) \geq \dfrac{1}{2 - \frac{|V|}{2|E|}}$ (if G has a perfect matching),

$\dfrac{\sqrt{26}-4}{2} \leq \rho(\text{SIMPLE PLANAR GRAPHS}) \leq .68436349$, and

$\rho(\text{SIMPLE } \Delta\text{-GRAPHS}) \geq \frac{1}{2} + \dfrac{\sqrt{(\Delta-1)^2+1} - (\Delta-1)}{2}$ (where the class $\Delta$-GRAPHS is the set of graphs of maximum degree at most $\Delta$).

## 1. INTRODUCTION

We consider loopless graphs, the edges being weighted. Here is careful notation concerning weighted graphs. Given a set V let $V^2$ denote $\{ \{x,y\} : x \neq y,\ x,y \in V \}$, the set of unordered pairs of elements of V. We write xy as shorthand for $\{x,y\}$. A <u>weight function</u> $\omega$ for a set V is an arbitrary function $\omega : V^2 \to [0,\infty)$, and corresponding to a weight function $\omega$ is the set $E = \{xy : \omega(xy) > 0\}$, the <u>edge set</u>. A <u>weighted graph</u> is then a triple $G = (V(G),E(G),\omega(G)) = (V,E,\omega)$ where $\omega$ is a weight function on $V^2$. We extend the domain of $\omega$ to the power set of $V^2$ by defining $\omega(F) = \sum_{e \in F} \omega(e)$ for any $F \subseteq V^2$, and we call $\omega(E)$ the <u>total weight</u> of G. When the range of $\omega$ (as applied to elements of $V^2$) is a subset of $\{0,1\}$ we say that $G = (V,E,\omega)$ is a <u>simple</u> graph (much as in the usual definition of a simple graph). We often specify a weighted graph $G = (V,E,\omega)$ by giving just the pair V and $\omega$, understanding E to be (as above) the set of unordered pairs having nonzero weight. Note that (V,E) will just be a "usual" simple graph, and we freely use standard graph theoretic terminology (e.g. the term *matching*) with the understanding that it refers to this simple graph underlying $(V,E,\omega)$.

Given weighted G we consider the following Randomized Greedy Algorithm (abbreviated as <u>RGA</u>) which produces a maximal matching in G:

```
begin
   M←∅
   while E(G) ≠ ∅ do
      begin
         choose e = uv ∈ E(G), with probability ω(e)/ω(E(G))
         G←G\{u,v}
         M←M∪{e}
      end
   output M and |M|
end
```

In other words we augment our matching at each stage by selecting the next matching edge with probability proportional to its weight from among the remaining edges. Here $|M|$ simply denotes the number of edges in M, ignoring the weights of the edges of M.

Dyer and Frieze [3] studied the average case behavior of $|M|$ in this algorithm when performed repeatedly on the same simple graph G, that is; when the next matching edge is selected with probability $\frac{1}{|E(G)|}$. They analyzed the expected ratio between the random output $|M|$ produced and the size of a maximum matching. In their concluding remarks, they mention the generalization to weighted graphs, i.e. this paper pursues one of their suggestions for further study. By introducing convexity and weighted graph considerations into their analysis, we obtain improved lower bounds for the average case behavior of $|M|$ produced by RGA. The weighted graph model is used here primarily as a tool for studying simple graphs.

The RGA and other greedy algorithms for matchings have been studied in [1, 2, 3, 4,5].

## 2. NOTATION, EXAMPLES, AND AN OVERVIEW OF RESULTS

For any $v \in V(G)$ we define the <u>degree</u> of v to be $\sum_{x \in V \setminus v} \omega(vx)$, and we denote it by $d_G(v)$. When G is specified by context we just write d(v) for the degree of v. Let $\bar{d} = \bar{d}(G) = \frac{1}{|V|} \sum_{v \in V} d(v) = \frac{2\omega(E)}{|V|}$, the <u>average degree</u> or <u>density</u> of G. A <u>Δ-graph</u> is a weighted graph with maximum degree at most Δ.

Let m(G) be the maximum size of a matching in G and let $\mu(G)$ be the expected size of the matching in G produced by the RGA. Let $r(G) = \mu(G)/m(G)$ if $\omega(E)>0$ and $r(G) = 1$ if $\omega(E)=0$. Given a class $\mathcal{G}$ of weighted graphs let $\rho(\mathcal{G}) = \inf_{G \in \mathcal{G}} r(G)$. That is, $\rho(\mathcal{G})$ is a measure of the worst-case performance of RGA as applied to graphs in $\mathcal{G}$.

Here are two examples illustrating RGA and the meanings of parameters $\omega$, m, $\mu$, r and $\rho$.

**Example 1**: Consider RGA as applied to the weighted graph G in Figure 1. Since the five edges are of distinct weights, we freely refer to an edge by indicating just its weight. Clearly m(G) = 2 and $\omega(E(G)) = 15$. To compute $\mu(G)$ we simply consider the possible options which RGA might follow, much as in a decision tree. As edges accumulate in the matching M we record the probabilities of each option at each stage, and when M becomes maximal we multiply the probability of the resulting maximal matching M by $|M|$. For example should edge 1 be the first edge introduced into M we calculate $\frac{1}{15}(\frac{2}{6}2 + \frac{4}{6}2)$, since upon removal of the ends of edge 1 the weight of the set of edges remaining is 6, so that with probability $\frac{2}{6}$ we will next select edge 2 and output a matching of size 2, and with probability $\frac{4}{6}$ we will next select edge 4 and output a matching of size 2. The full calculation of $\mu$ is then

$\mu(G) = \frac{1}{15}(\frac{2}{6}2 + \frac{4}{6}2) + \frac{2}{15}(\frac{1}{1}2) + \frac{3}{15}(\frac{4}{4}2) + \frac{4}{15}(\frac{1}{4}2 + \frac{3}{4}2) + \frac{5}{15}1 = \frac{5}{3}$.

Alternatively, we could have noticed that for any maximal matching M the size |M| is either 1 or 2, with |M| being 1 with probability $\frac{5}{15}$ and therefore being 2 with probability $\frac{10}{15}$, so that $\mu(G) = \frac{5}{15}1 + \frac{10}{15}2 = \frac{5}{3}$. Observe that for the maximal matching M = {1,2}, the probability that it is produced by RGA in the order "first 1 then 2" (in this case $\frac{1}{15}\,\frac{2}{6}$) is not equal to the probability that it is produced in the order "first 2 then 1" (in this case $\frac{2}{15}\,\frac{1}{1}$). Thus one expects that typically order matters in this context. Lastly, we have $r(G) = \frac{5/3}{2} = \frac{5}{6}$, which says that on average RGA produces a matching $\frac{5}{6}$ as large as a maximum matching in G.

**Example 2**: Consider RGA as applied to the simple graph H in Figure 1, each edge having weight 1. The eight edges labeled "e" are structurally equivalent within the graph, so we refer to edges as being of type e or not. While this graph is relatively complicated for an early example, we pursue its r value now since we will refer to it later and since we can use it now to illustrate the parameter $\rho$.

Clearly m(H) = 5 and $\omega(E(H)) = 15$. Observe that |M| = 3, 4 or 5 for all maximal matchings M. Clearly |M| = 3 if and only if $M = \{x_1,x_2,x_3\}$. Conveniently, no matter which of edges $x_1,x_2,x_3$ might be selected as the first edge of M, the remaining edges have total weight 10. Also, should two of those three edges be selected as the first two edges of M, the remaining edges have total weight 5. Therefore prob(|M|=3) = prob($M=\{x_1,x_2,x_3\}$) = $\frac{3}{15}\,\frac{2}{10}\,\frac{1}{5} = \frac{1}{125}$. Next observe that |M| = 5 if and only if M consists of $x_1$ and four independent edges of type e. Letting $M_i$ denote the ith edge selected by RGA, we then have the following computations, keeping in mind two things: H is highly symmetric, and the distribution of matchings produced by RGA applied to a disconnected graph is identical to the distribution of the union of matchings produced by RGA applied separately to each component of the disconnected graph.

prob(|M|=5 and $M_1=x_1$) = $\frac{1}{15}\,\frac{4}{5}\,\frac{4}{5} = \frac{16}{375}$,

prob(|M|=5 and $M_2=x_1$) = $\frac{8}{15}\,\frac{1}{10}\,\frac{4}{5} = \frac{16}{375}$,

prob(|M|=5 and $M_3=x_1$) = $\frac{8}{15}\left[\frac{1}{10}\,\frac{1}{8}\,\frac{4}{5} + \frac{4}{10}\,\frac{1}{5}\right] = \frac{18}{375}$ (The two terms in brackets correspond to whether or not $M_1$ and $M_2$ are on the <u>same</u> side of $x_1$ in H.),

prob(|M|=5 and $M_4=x_1$) = $\frac{8}{15}\left[\frac{1}{10}\,\frac{4}{8}\,\frac{1}{3} + \frac{4}{10}\,\frac{2}{5}\,\frac{1}{3}\right] = \frac{14}{375}$ (again, same side or not?),

prob(|M|=5 and $M_5=x_1$) = $\frac{8}{15}\left[\frac{1}{10}\,\frac{4}{8}\,\frac{1}{3} + \frac{4}{10}\,\frac{2}{5}\,\frac{1}{3}\right] = \frac{14}{375}$ (again, same side or not?).

Therefore prob(|M|=5) = $\frac{16}{375} + \frac{16}{375} + \frac{18}{375} + \frac{14}{375} + \frac{14}{375} = \frac{26}{125}$,

so prob(|M|=4) = $1 - \frac{1}{125} - \frac{26}{125} = \frac{98}{125}$.

Therefore $\mu(H) = \frac{1}{125}3 + \frac{98}{125}4 + \frac{26}{125}5 = 4.2$, and $r(H) = \frac{4.2}{5} = .84$.

Now H is a simple 3-regular graph (and therefore also a Δ-graph for Δ=3), and H is also planar (despite its illustration). Therefore this single example gives us upper bounds for the cases $\mathcal{G}$ = {SIMPLE 3-REGULAR GRAPHS}, {SIMPLE 3-GRAPHS} and {SIMPLE PLANAR GRAPHS}, i.e. ρ(SIMPLE 3-REGULAR GRAPHS) ≤ .84 and ρ(SIMPLE 3-GRAPHS) ≤ .84 and ρ(SIMPLE PLANAR GRAPHS) ≤ .84. But of course these facts are not very interesting unless we have reason to believe that H has particularly low r value within the corresponding class $\mathcal{G}$. Thus the more general results of this paper are the lower bounds given for $\rho(\mathcal{G})$ for interesting classes $\mathcal{G}$. For any class $\mathcal{G}$ our upper bound for $\rho(\mathcal{G})$ simply represents the r value of a particular graph for which RGA approximates m most poorly among graphs in $\mathcal{G}$ whose r values we happen to know.

Dyer and Frieze [3] proved that

i) For $\mathcal{G}$ a nontrivial class of simple graphs closed under vertex deletions, $\rho(\mathcal{G}) \geq \dfrac{1}{2 - \frac{1}{2}\kappa(\mathcal{G})}$ , where $\kappa(\mathcal{G}) = \inf_{G \in \mathcal{G}} \{ \frac{|V|}{|E|} : |E|>0 \}$.

ii) ρ(SIMPLE GRAPHS) $= \frac{1}{2}$.

iii) $\frac{6}{11} \leq$ ρ(SIMPLE PLANAR GRAPHS) $\leq \frac{11}{15}$.

iv) ρ(SIMPLE Δ-GRAPHS) $\geq \frac{\Delta}{2\Delta - 1}$.

v) ρ(SIMPLE FORESTS) $= \frac{2}{3} + 2\sum_{k=0}^{\infty} \frac{(-2)^k}{(2k+5)!!} \approx .7690397$, where $n!! = n(n-2)(n-4)\cdots(3)(1)$ for n odd.

vi) For simple G, the size of the matching produced by RGA is a.a. near $\mu(G)$.

We refer to these results by their number i),ii),iii),iv),v),vi) periodically in this paper. In i), observe that $\frac{1}{2}\kappa(\mathcal{G}) = 1/\sup_{G\in\mathcal{G}} \bar{d}(G)$. The lower bounds in iii) and iv) follow from i) in view of the facts that simple planar graphs (resp. simple Δ-graphs) have densities bounded above by 6 (resp. Δ). Concerning ii), the fact that ρ(SIMPLE GRAPHS) ≥ 1/2 is easy enough to explain here, independently of i), because *any* maximal matching M in G (i.e. any matching M which can be produced by RGA) has size at least m(G)/2, as for example known in [5]. To see this, let M' be a maximum matching. Each edge of M is adjacent to at most two edges of M', while each edge of M' is adjacent to at least one edge of M (since M is maximal). It follows that $|M| \geq \frac{|M'|}{2}$. Thus, for the same reasons as for ii) we have ρ(WEIGHTED GRAPHS) = 1/2.

Dyer and Frieze [3] remarked that the bounds in iii) are almost certainly not tight. Their intuition was correct. Among the results of this paper we prove that if simple G has a perfect matching then

$r(G) \geq \dfrac{1}{2 - \frac{|V|}{2|E|}} = \dfrac{\bar{d}}{2\bar{d} - 1}$ [as it turns out thereby rendering i) and iv) as corollaries], that

$\dfrac{\sqrt{26}-4}{2} \leq \rho(\text{SIMPLE PLANAR GRAPHS}) \leq .68436349$ [thereby improving iii) at both ends], and that

$\rho(\text{SIMPLE } \Delta\text{-GRAPHS}) \geq \frac{1}{2} + \frac{\sqrt{(\Delta-1)^2+1} - (\Delta-1)}{2}$ [thereby improving iv)].

These lower bounds follow from bounds we obtain in section 5 for $\mu(G)$ when G is a "pincushion" graph.

## 3. MONOTONICITY RESULTS

While [3] concerned only simple graphs, some of its results and proofs carry over without much difficulty. This section is devoted to proving some of those results (the ones we use) when recast in the more general setting of weighted graphs. The proofs borrow a great deal from [3], and are included for the sake of completeness. Given $S \subseteq V$ let $G\backslash S$ denote the weighted graph with vertex set $V\backslash S$ and weight function being the restriction of $\omega$ to $(V\backslash S)^2$. As is customary we write $G\backslash x$ as shorthand for $G\backslash\{x\}$. Given $F \subseteq E$ we let <u>G[F]</u> be the subgraph induced by the set of vertices at the ends of edges of F, i.e. $G[F] = G\backslash S$ where $S = V\backslash\{\text{ends of edges of } F\}$. The following is a generalization of Lemma 1 and Corollary 1 from [3].

**Lemma 1**: Let G be a weighted graph , M a maximum matching of G.

a) For all $v \in V$, $\mu(G) \geq \mu(G\backslash v) \geq \mu(G) - 1$.

b) If a vertex v of G is not an end of an edge of M, then $r(G) \geq r(G\backslash v)$. In particular, $r(G) \geq r(G[M])$.

<u>Proof</u>: Statement a) clearly holds when $|V| = 1$ and also when $\omega(E) = 0$, so we consider G with $|V|>1$, $\omega(E)>0$, and inductively assume that a) holds for all weighted graphs with fewer vertices than G. Given $v \in V$ let $H = G\backslash v$ and let $A = \{xy \in E : v \notin \{x,y\}\} = E(H)$. Observe that in applying RGA to G, if edge xy is the first matching edge selected then the continuation of RGA is just an application of RGA to $G\backslash\{x,y\}$. Consequently $1+\mu(G\backslash\{x,y\})$ is the expected size of the matching produced by RGA, conditioned on xy being the first edge selected. Therefore $\mu(G)$ is a weighted average of the quantities $1+\mu(G\backslash\{x,y\})$. Thus

$$\begin{aligned}
\mu(G) &= \frac{1}{\omega(E)} \sum_{xy\in E} \omega(xy)\,[1+\mu(G\backslash\{x,y\})] \\
&= \frac{1}{\omega(E)} \sum_{xy\in A} \omega(xy)\,[1+\mu(G\backslash\{x,y\})] + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,[1+\mu(H\backslash y)] \\
&\geq \frac{1}{\omega(E)} \sum_{xy\in A} \omega(xy)\,[1+\mu(H\backslash\{x,y\})] + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,[1+\mu(H\backslash y)] \qquad \text{(by induction)} \\
&= \left(1 - \frac{d(v)}{\omega(E)}\right) \mu(H) + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,[1+\mu(H\backslash y)]
\end{aligned}$$

$$= \mu(H) + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,[1+\mu(H\backslash y) - \mu(H)]$$

$$\geq \mu(H). \qquad\qquad \text{(by induction)}$$

Also,

$$\mu(G) - 1 = -1 + \frac{1}{\omega(E)} \sum_{xy\in A} \omega(xy)\,[1+\mu(G\backslash\{x,y\})] + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,[1+\mu(H\backslash y)]$$

$$= \frac{1}{\omega(E)} \sum_{xy\in A} \omega(xy)\,\mu(G\backslash\{x,y\}) + \frac{1}{\omega(E)} \sum_{y\in V\backslash v} \omega(vy)\,\mu(H\backslash y)$$

$$\leq \frac{1}{\omega(E)} \sum_{xy\in A} \omega(xy)\,[1+\mu(H\backslash\{x,y\})] + \frac{d(v)}{\omega(E)}\,\mu(H) \qquad \text{(two uses of induction )}$$

$$= (1 - \frac{d(v)}{\omega(E)})\,\mu(H) + \frac{d(v)}{\omega(E)}\,\mu(H)$$

$$= \mu(H).$$

This completes the proof of a).

As for b), since $m(G\backslash v) = m(G)$ it follows from a) that $r(G) \geq r(G\backslash v)$. Iterating the deletion process for each $v \in V$ which is not an end of an edge of M, we have $r(G) \geq r(G[M])$. ∎

Given an edge $u_1v_1$ in G we define a graph $G' = (V(G),E',\omega')$ derived from G. An informal description of G' is as follows. Let $V(G') = V(G)$, and let $E' = E(G')$ be obtained from E(G) by deleting each edge $v_1y$, $y\notin \{u_1,v_1\}$, adding its weight to the weight of $u_1y$, and leaving all other edges of G (i.e. $u_1v_1$, together with edges incident on neither $u_1$ or $v_1$) the same in G'. Thus we "transfer" the weight of each edge $v_1y$, $y\notin\{u_1,v_1\}$, to the edge $u_1y$. The result is that vertex $v_1$ in G' is an end point joined only to $u_1$, where the weights in G' of unordered vertex pairs $u_1y$ are given by $\omega'(u_1y) = \omega(v_1y) + \omega(u_1y)$ for each $y\notin \{u_1,v_1\}$, and all other unordered pairs in G' have the same weights they had in G. Formally, we can define the weights of all unordered pairs of vertices in G' by letting $\omega'(xy) = \omega(xy)$ if $x,y\notin \{u_1,v_1\}$, $\omega'(v_1y) = 0$ for each $y \in V\backslash\{u_1,v_1\}$, $\omega'(u_1v_1) = \omega(u_1v_1)$, and $\omega'(u_1y) = \omega(\{u_1y,v_1y\})$ for each $y \in V\backslash\{u_1,v_1\}$. The construction of the graph G' from a graph G is illustrated in Figure 2.

Suppose $G = (V,E,\omega)$ has a matching M consisting of the k edges $u_1v_1,u_2v_2,...,u_kv_k$. Next we define a graph $G^* = (V(G),E^*,\omega^*)$ derived from G. Again we begin with an informal description. Let $V(G^*) = V(G)$, and let $E^* = E(G^*)$ be obtained from E(G) by transferring the weight of each of the edges $u_rv_s$, $u_sv_r$, $v_sv_s$ to the edge $u_ru_s$ for all pairs $r\neq s$, and transferring the weight of each edge $xv_r$ to the edge $xu_r$ for all vertices x lying on no edge of M and for all $1\leq r\leq k$. Thus G* may be viewed as the graph obtained by iterating the "priming" operation above over all the matching edges $u_1v_1,u_2v_2,...,u_kv_k$. Formally, the definition of weights on unordered pairs of vertices of G* is given as follows. Let $\omega^*(xy) = \omega(xy)$ if $x,y\notin\{u_1,u_2,...,u_k,v_1,v_2,...,v_k\}$. For each $r<s$ let $\omega^*(u_ru_s) = \omega(\{u_ru_s,u_rv_s,v_ru_s,v_rv_s\})$ and $\omega^*(u_rv_s) = \omega^*(v_ru_s) = \omega^*(v_rv_s) = 0$. For each r let $\omega^*(u_rv_r) = \omega(u_rv_r)$. Finally for each vertex x lying on no edge of M and each $1\leq r\leq k$, let $\omega^*(v_rx) = 0$ and $\omega^*(u_rx) = \omega(\{u_rx,v_rx\})$. The construction of G* from G is illustrated in Figure 2.

Note that G, G' and G* will all have the same number of vertices and the same total weight. Now G' (resp. G*) depends upon the choice of $u_1v_1$ (resp. M) and also upon the choice(s) involved in designating which end of each edge $u_rv_r$ will be the "u" end, i.e. the end to which the edge weights get transferred. However, in this paper these choices will be either immaterial or clear from context and hence such choices are not incorporated into the notation for G' and G*. Having defined G' and G*, here is a weighted analogue of Lemma 2 of [3] which we will need later. Its claim a) is intuitively clear, since in G' the edges of G formerly incident at $u_1$ or $v_1$ are rearranged in star-like fashion joined at $u_1$ so that in G' if RGA selects any one of those edges then all of those edges are immediately deleted, in contrast to what would happen in G.

**Lemma 2**: For a weighted graph G,

a) Given an edge $u_1v_1$ in G, $\mu(G') \le \mu(G)$

b) Given a maximum matching $M = \{u_1v_1, u_2v_2, ..., u_mv_m\}$, $r(G') \le r(G)$ and $r(G^*) \le r(G)$, where here G' is taken relative to edge $u_1v_1$.

Proof: Since G* is the result of iterating m operations of the type involved in forming G' from G, to prove b) it suffices to show that $r(G') \le r(G)$. But since $m(G') = m(G)$ the statement $r(G') \le r(G)$ would follow from $\mu(G') \le \mu(G)$, which we now prove by induction on the number of vertices. In the base case when $|V| = 2$, $G' = G$ and the result is trivial. Given weighted G with $|V(G)| \ge 3$, assume that the inequality $\mu(H') \le \mu(H)$ holds for all weighted graphs H with $|V(H)| < |V(G)|$. Let $A = \{xy \in E(G): \{x,y\} \cap \{u_1,v_1\} = \varnothing\} = E(G \backslash \{u_1,v_1\})$ and $V' = V \backslash \{u_1,v_1\}$. Then

$$
\begin{aligned}
\mu(G) &= 1 + \frac{1}{\omega(E)}\Big( \sum_{xy\in A} \omega(xy)\, \mu(G\backslash\{x,y\}) + \omega(u_1v_1)\mu(G\backslash\{u_1,v_1\}) + \\
&\qquad \sum_{x\in V'} [\omega(u_1x)\, \mu(G\backslash\{u_1,x\}) + \omega(v_1x)\, \mu(G\backslash\{v_1,x\})] \Big) \\
&\ge 1 + \frac{1}{\omega(E)}\Big( \sum_{xy\in A} \omega'(xy)\, \mu(G'\backslash\{x,y\}) + \omega'(u_1v_1)\mu(G'\backslash\{u_1,v_1\}) + \\
&\qquad \sum_{x\in V'} [\omega(u_1x)\, \mu(G\backslash\{u_1,v_1,x\}) + \omega(v_1x)\, \mu(G\backslash\{u_1,v_1,x\})] \Big) \\
&= 1 + \frac{1}{\omega(E')}\Big( \sum_{xy\in A} \omega'(xy)\, \mu(G'\backslash\{x,y\}) + \omega'(u_1v_1)\, \mu(G'\backslash\{u_1,v_1\}) + \sum_{x\in V'} \omega'(u_1x)\, \mu(G'\backslash\{u_1,x\}) \Big) \\
&= \mu(G'),
\end{aligned}
$$

where the inductive hypothesis gives us that $\mu(G\backslash\{x,y\}) \ge \mu(G'\backslash\{x,y\})$ for $xy \in A$ (since $(G\backslash\{x,y\})' = G'\backslash\{x,y\}$), and Lemma 1a gives us that $\mu(G\backslash\{u_1,x\}) \ge \mu(G\backslash\{u_1,v_1,x\}) = \mu(G'\backslash\{u_1,x\})$ and $\mu(G\backslash\{v_1,x\}) \ge \mu(G\backslash\{u_1,v_1,x\}) = \mu(G'\backslash\{u_1,x\})$ for $x \in V'$. ∎

## 4. PINCUSHION GRAPHS

A pincushion graph is a weighted graph which has a perfect matching M with the feature that each edge of M has weight 1, each edge of M having at least one of its ends being of degree exactly 1. A pincushion graph has a unique perfect matching, and we refer to an edge as being either a <u>pin edge</u> or a <u>core edge</u> depending upon whether the edge is or is not part of that matching. For each pin edge we designate one of its ends of degree 1 as being a <u>pin vertex</u> and the other end as being a <u>core vertex</u>, so the vertices of a pincushion graph partition into pin vertices vs. core vertices. In a pincushion graph $G = (V,E,\omega)$ on 2n vertices we let $C = C(G)$, the <u>core</u> of G, denote the weighted graph induced by the core vertices. Also let E(C) denote the set of core edges, and let $\varepsilon = \varepsilon(G) = \omega(E(C))$ denote the total weight of just the core edges. Thus $\omega(E) = n + \varepsilon$. Among simple graphs on a given number 2n many vertices and a given number of edges, a pincushion graph is an attractive candidate for having relatively low $\mu$ value, since RGA applied to a pincushion graph yields a matching of size (n – # of core edges selected by RGA), so that intuition suggests that selection of core edges tends to make the RGA output a matching of relatively small size. This in turn suggests that the more total weight in the core, the lower the $\mu$ value. While these generalities are useful as heuristics, they are not valid in every case! For example, the weighted pincushion graph in Figure 3b has higher $\mu$ value than does the pincushion graph in Figure 3a, despite its additional core edge.

Here is a more tangible reason for studying pincushion graphs. If simple graph H has a maximum matching M then for $G = H[M]$ we have from Lemma 1b and Lemma 2b that $r(H) \geq r(G) \geq r(G^*)$, where $G^*$ was previously defined (relative to M). One can check that $G^*$ is a pincushion graph in which each core edge has weight either 1,2,3 or 4. So, to find a lower bound for r(H) we need only find a lower bound for $r(G^*)$. In [3] the authors didn't allow transforming simple H to $G^*$ when the result was not a simple graph. Having shown that many results from [3] carry over to weighted graphs, we have no such worries.

In weighted G let $f(G) = \sum_{v\in V} (d(v))^2 - \sum_{uv\in V^2} (\omega(uv))^2$ , an invariant of G which we study for technical reasons. Here is a lemma concerning f and the effect of transferring some or all of the weight of one edge to the weight of an adjacent edge, t being the amount of weight transferred.

**Lemma 3**: Given a real number t and three vertices x,y,z in weighted G such that $0 \leq t \leq \omega(xz)$ and $d_{G\backslash x}(y) \leq d_{G\backslash x}(z)$, let H denote the weighted graph on the same vertex set V, with weight function $\omega_H$ defined by modifying $\omega$ as follows: $\omega_H(xy) = \omega(xy) + t$, $\omega_H(xz) = \omega(xz) - t$, and $\omega_H(uv) = \omega(uv)$ for all $uv\in V^2\backslash\{xy,xz\}$. Then $f(H) \leq f(G)$.

<u>Proof</u>: $f(G) - f(H)$

$= (d_G(y))^2 - (d_G(y) + t)^2 + (d_G(z))^2 - (d_G(z) - t)^2 - (\omega(xy))^2 - (\omega(xz))^2 + (\omega(xy) + t)^2 + (\omega(xz) - t)^2$

$= 2t\,( [d_G(z) - \omega(xz)] - [d_G(y) - \omega(xy)] )$

$= 2t\,(d_{G\backslash x}(z) - d_{G\backslash x}(y)) \geq 0.$ ∎

When a core edge is selected in some stage of RGA, the removal of its ends necessarily renders two pin vertices as now isolated. This suggests that RGA will tend to produce a smaller matching when a core edge is selected first than when a pin edge is selected first. But the removal of the ends of a core edge also deletes along with it any core edges adjacent to the removed edge, leaving less total weight of core edges available for selection later. To measure the effect of removing a pin edge or core edge, it is convenient to give upper bounds for the expected total weight of the core edges that remain when a random pin edge or random core edge is removed. When a core edge of a pincushion graph is removed in the course of RGA we regard the two resulting isolated pin vertices as being automatically deleted, so that the result is again a pincushion graph. Let $\bar{\varepsilon}(G)$ denote the expected total weight of the remaining core edges given that RGA selects a random first edge which is a core edge; that is, $\bar{\varepsilon}(G)$ is the weighted average of the values $\omega(E(C\backslash\{u,v\})) = \varepsilon(G\backslash\{u,v\})$ where each such value receives weight $\frac{\omega(uv)}{\omega(E(C))}$. Let $M(n,\varepsilon)$ denote a pincushion graph on 2n vertices, its set of core edges being a matching comprised of $h = \lfloor n/2 \rfloor$ many edges, each of weight exactly $\varepsilon/h$ (i.e. a "balanced" matching). The pincushion graph $M(n,\varepsilon)$ plays an extremal role.

**Lemma 4**: Let G be a pincushion graph on 2n vertices with $\varepsilon$ being the total weight of its core edges.

a) The average of the values $\varepsilon(G\backslash\{u,v\})$ over all pin edges uv is $\frac{(n-2)\varepsilon}{n}$. (Thus we have a formula for the expected total weight of the remaining core edges in $G\backslash\{u,v\}$ when RGA selects a random first edge uv which happens to be a pin edge.)

b) The weighted average $\bar{\varepsilon}(G)$ of the values $\varepsilon(G\backslash\{u,v\})$ over all core edges uv is maximized when $G = M(n,\varepsilon)$. In particular, $\bar{\varepsilon}(G) \le \frac{(n-2)\varepsilon}{n}$. (Thus we have an upper bound for the expected total weight of the remaining core edges in $G\backslash\{u,v\}$ when RGA selects a random first edge uv which happens to be a core edge.)

c) If G is simple then $\bar{\varepsilon}(G) \le \varepsilon + 1 - \frac{4\varepsilon}{n}$.

<u>Proof</u>: Given that uv is a pin edge, the expected total weight of the core edges in $G\backslash\{u,v\}$ is

$\frac{1}{n} \sum_{v \in V(C)} \varepsilon(G\backslash v) = \frac{1}{n}[n\varepsilon - \sum_{v \in V(C)} d_C(v)] = \frac{1}{n}[n\varepsilon - 2\varepsilon] = \frac{(n-2)\varepsilon}{n}$, proving a).

For b), first observe that the total weight of the core edges remaining in $M(n,\varepsilon)$ when the ends of a core edge uv are removed is independent of the choice of core edge, so $\bar{\varepsilon}(M(n,\varepsilon)) = (h-1)\frac{\varepsilon}{h}$. When n is even this is $\frac{(n-2)\varepsilon}{n}$, and when n is odd this is $\frac{(n-3)\varepsilon}{(n-1)}$, which is still less than or equal to $\frac{(n-2)\varepsilon}{n}$. Therefore it only remains to prove that $\bar{\varepsilon}(G)$ is maximized when $G = M(n,\varepsilon)$.

We begin by relating $\overline{\varepsilon}(G)$ to f(C). Observe that

$$
\begin{aligned}
\overline{\varepsilon}(G) &= \frac{1}{\varepsilon} \sum_{uv\in E(C)} \omega(uv)\, \varepsilon(G\backslash\{u,v\}) \\
&= \frac{1}{\varepsilon} (\varepsilon^2 - \sum_{uv\in E(C)} \omega(uv)\, [d_C(u) + d_C(v) - \omega(uv)]) \\
&= \frac{1}{\varepsilon} (\varepsilon^2 + \sum_{uv\in E(C)} (\omega(uv))^2 - \sum_{x\in V(C)} (d_C(x))^2 ) \\
&= \frac{1}{\varepsilon} (\varepsilon^2 - f(C)).
\end{aligned}
$$

Thus it suffices to prove that f(C) is minimized when C is isomorphic to the core of M(n,ε).

First observe that the core of M(n,ε) minimizes f(C) among all cores C whose edges form only a matching. For if such a C consists of edges $e_1,e_2,...,e_k$ (whose total weight is of course ε), Then $f(C) = \sum_{1\le i\le k} (\omega(e_i))^2$. By elementary optimization principles f(C) is minimized when the matching edges are of equal weight, with k as large as allowable. In this case k is at most h, so the core of M(n,ε) does indeed minimize f(C) among those cores whose edges form only a matching.

We now show that M(n,ε) minimizes f(C) over arbitrary cores C. We will use induction on n, finally reducing to the case just handled when C is a matching. The base cases n=1 and n=2 are clear, since C must then be a matching on 0 or 1 edges, so suppose inductively that the core of M(n,ε) does minimize f(C) over all pincushion graphs on fewer vertices than |V(G)|. We are done if the core C of G is a matching, so suppose adjacent edges xy and xz exist in C. Without loss of generality, $d_{C\backslash x}(y) \le d_{C\backslash x}(z)$. Applying Lemma 3 with t = ω(xz), letting $G_1$ be the pincushion graph formed from G by transferring all of the weight of xz to xy (that is, $G_1$ is the graph H in the statement of Lemma 3), we have $f(C(G_1)) \le f(C)$. If in $C(G_1)$ there are still two or more core edges incident at x then we can apply Lemma 3 again to transfer the weight of one of those edges to the other without increasing f, so we can iterate this process to obtain a pincushion graph $G_2$ on 2n vertices with a core $C(G_2)$ of total weight ε with only one core edge incident at core vertex x (without loss of generality that edge being xy), with $f(C(G_2)) \le f(C)$. Now if y is incident to more than one core edge of $G_2$, Lemma 3 can be applied at y (the roles of x and y now being reversed) to transfer the weights of all of those core edges to the weight of xy (noting that x must have least degree in $G_2\backslash y$). We thereby obtain a pincushion graph $G_3$ on 2n vertices with a core $C(G_3)$ of total weight ε, with $f(C(G_3)) \le f(C)$, such that x and y and the edge between them form a single component of $C(G_3)$. By the inductive hypothesis $f(C(G_3\backslash\{x,y\})) \ge f(H)$ where H is the weighted graph induced by the vertices of $C(G_3\backslash\{x,y\})$ resulting from reorganizing the total weight of $C(G_3\backslash\{x,y\})$ into a "balanced" matching. Since f is additive over the connected components of any graph to which it is applied, it follows that $f(C(G_3)) \ge f(C(G_4))$, where $G_4$ is a pincushion graph on the vertices of $G_3$ whose core $C(G_4)$ is a matching of total weight ε. We have already observed by optimization that $f(C(G_4)) \ge f(M(n,\varepsilon))$, and hence we get $f(C) \ge f(M(n,\varepsilon))$ as desired.

For c), since $\sum_{x\in V(C)} d_C(x) = 2\varepsilon$, the sum of the squares of the core degrees is minimized when the degrees are all equal to their average $\frac{2\varepsilon}{n}$. Thus we have $\sum_{x\in V(C)} (d_C(x))^2 \geq n\,\frac{4\varepsilon^2}{n^2} = \frac{4\varepsilon^2}{n}$. Recall that G is simple, so that $\omega(uv) = 1$ for all edges uv of G. Hence from the proof of b) we have that $\bar{\varepsilon}(G) = \frac{1}{\varepsilon}(\varepsilon^2 + \sum_{uv\in E(C)} (\omega(uv))^2 - \sum_{x\in V(C)} (d_C(x))^2)$

$= \frac{1}{\varepsilon}(\varepsilon^2 + \varepsilon - \sum_{x\in V(C)} (d_C(x))^2) \leq \varepsilon + 1 - \frac{4\varepsilon}{n}$. ∎

The next lemma in combination with Lemma 4 is the key idea behind our lower bounds for $\mu(G)$ which follow in the next section. It is fueled by Jensen's inequality, which for our purposes states that if f(x) is concave up over an interval containing inputs $x_1,x_2,...,x_k$ and if $\alpha_1,\alpha_2,...,\alpha_k$ are nonnegative constants which sum to $\alpha>0$, then

$$\sum \alpha_i\, f(x_i) \geq \alpha\, f\left(\frac{\sum \alpha_i x_i}{\alpha}\right).$$

**Lemma 5**: Suppose that $f_{n-1}(\varepsilon)$ and $f_{n-2}(\varepsilon)$ are functions of a real variable $\varepsilon$, both functions being concave up on the interval $[0,\infty)$. Further suppose that G is a pincushion graph on 2n vertices such that for each core vertex u, $\mu(G\backslash u) \geq f_{n-1}(\varepsilon(G\backslash u))$ and for each core edge uv, $\mu(G\backslash\{u,v\}) \geq f_{n-2}(\varepsilon(G\backslash\{u,v\}))$. Then $\mu(G) \geq 1 + \frac{1}{n+\varepsilon}\,[\, n\, f_{n-1}(\frac{(n-2)\varepsilon}{n}) + \varepsilon\, f_{n-2}(\bar{\varepsilon}(G))\,]$.

<u>Proof</u>: By handling separately the cases where the first edge selected by RGA on G is a pin edge versus a core edge we have $\mu(G) = 1 + \frac{1}{n+\varepsilon}\,[\sum_{u\in V(C)} 1\, \mu(G\backslash u) + \sum_{uv\in E(C)} \omega(uv)\, \mu(G\backslash\{u,v\})\,]$

$\geq 1 + \frac{1}{n+\varepsilon}\,[\sum_{u\in V(C)} f_{n-1}(\varepsilon(G\backslash u)) + \sum_{uv\in E(C)} \omega(uv)\, f_{n-2}(\varepsilon(G\backslash\{u,v\}))\,]$. By Lemma 4a the average of the inputs $\varepsilon(G\backslash u)$ is $\frac{(n-2)\varepsilon}{n}$, so by Jensen's inequality $\sum_{u\in V(C)} f_{n-1}(\varepsilon(G\backslash u)) \geq n\, f_{n-1}(\frac{(n-2)\varepsilon}{n})$. The weighted average $\frac{1}{\varepsilon}\sum_{uv\in E(C)} \omega(uv)\, \varepsilon(G\backslash\{u,v\})$ is, by definition, $\bar{\varepsilon}(G)$, so by Jensen's inequality

$\sum_{uv\in E(C)} \omega(uv)\, f_{n-2}(\varepsilon(G\backslash\{u,v\})) \geq \varepsilon\, f_{n-2}(\bar{\varepsilon}(G))$. The result immediately follows. ∎

The point here is that while clearly the total core weight will decrease during the execution of RGA, it will on average decrease by a measurable amount per step. Jensen's inequality, as applied in Lemma 5, allows us to take this into account when lower bounding $\mu$.

## 5. MAIN RESULTS

At last we are ready to produce improved bounds for the performance of RGA. Many of the lower bounds given for r(G) are expressed in the form $r(G) \geq \frac{1}{2} + x$ in order to indicate the extent to which a bound is an improvement over the trivial bound $r(G) \geq \frac{1}{2}$.

**Theorem 1**: If G is a simple pincushion graph on 2n vertices and ε many core edges, then $\mu(G) \geq \frac{n}{2}(1 + \frac{n}{n+\varepsilon})$. It follows that $r(G) \geq \frac{1}{2} + \frac{n}{2(n+\varepsilon)} = \frac{1}{2} + \frac{1}{2\bar{d}}$ for such a graph G.

Proof: Let $f_n(\varepsilon) = \frac{n}{2}(1 + \frac{n}{n+\varepsilon})$ for each natural number n. For the case n=1, G must have ε=0, so the claim is true since $\mu(G) = 1 = f_n(0)$. For the case n=2, E(C) consists of a single edge of weight ε, and it is easy to verify that $\mu(G) = 1 + \frac{2}{2+\varepsilon} = f_n(\varepsilon)$, so again the claim is true, completing our basis for induction. Inductively suppose the claim is true for all pincushion graphs on fewer than 2n vertices for some n > 2, and let G be a simple pincushion graph on 2n vertices. By Lemma 4c, $\bar{\varepsilon}(G) \leq \varepsilon + 1 - \frac{4\varepsilon}{n}$. Since $f_{n-2}(\varepsilon)$ is nonincreasing in ε, we have that $f_{n-2}(\bar{\varepsilon}(G)) \geq f_{n-2}(\varepsilon + 1 - \frac{4\varepsilon}{n})$. Since each of $f_{n-1}(\varepsilon)$ and $f_{n-2}(\varepsilon)$ are concave up on $[0,\infty)$, we have from Lemma 5 that

$$\mu(G) \geq 1 + \frac{1}{n+\varepsilon}\left[ n\, f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon\, f_{n-2}(\bar{\varepsilon}(G)) \right] \geq 1 + \frac{1}{n+\varepsilon}\left[ n\, f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon\, f_{n-2}\left(\varepsilon + 1 - \frac{4\varepsilon}{n}\right) \right]$$

$$= 1 + \frac{1}{n+\varepsilon}\left[ \frac{n(n-1)}{2}\left(1 + \frac{n-1}{n-1+\frac{(n-2)\varepsilon}{n}}\right) + \frac{\varepsilon(n-2)}{2}\left(1 + \frac{n-2}{n+\varepsilon-1-\frac{4\varepsilon}{n}}\right) \right]$$

$$= \frac{n}{2} + \frac{1}{2(n+\varepsilon)}\left[ n + \frac{n(n-1)^2}{n-1+\frac{(n-2)\varepsilon}{n}} + \frac{\varepsilon(n-2)^2}{n+\varepsilon-1-\frac{4\varepsilon}{n}} \right].$$

So to complete the inductive proof that $\mu(G) \geq \frac{n}{2}(1 + \frac{n}{n+\varepsilon}) = \frac{n}{2} + \frac{1}{2(n+\varepsilon)} n^2$, it suffices to show that $n + \frac{n(n-1)^2}{n-1+\frac{(n-2)\varepsilon}{n}} + \frac{\varepsilon(n-2)^2}{n+\varepsilon-1-\frac{4\varepsilon}{n}} \geq n^2$. Simplification shows this inequality equivalent to $(n-2)(n-1+\frac{(n-2)\varepsilon}{n}) \geq (n-1)(n+\varepsilon-1-\frac{4\varepsilon}{n})$, which is in turn equivalent to $\varepsilon \geq n-1$. So if $\varepsilon \geq n-1$, then we are done. To handle the remaining case, i.e. when $\varepsilon < n-1$, we have the useful knowledge that C is disconnected. Suppose that one of the connected components K of C(G) has p vertices and q edges, there being ε-q edges in the n-p vertex subgraph C(G)\K of C(G). Let $H_1$ and $H_2$ be the pincushion graphs with cores K and C(G)\K respectively. Then by the additivity of μ over connected components and induction on the number of points in pincushion graphs we have $\mu(G) = \mu(H_1) + \mu(H_2) \geq f_p(q) + f_{n-p}(\varepsilon-q)$ $= \frac{p}{2} + \frac{p^2}{2(p+q)} + \frac{n-p}{2} + \frac{(n-p)^2}{2(n-p+\varepsilon-q)} = \frac{n}{2} + \frac{1}{2}\left(\frac{p^2}{p+q} + \frac{(n-p)^2}{n-p+\varepsilon-q}\right)$. So it suffices to show that $\frac{p^2}{p+q} + \frac{(n-p)^2}{n-p+\varepsilon-q} \geq \frac{n^2}{n+\varepsilon}$. This last inequality holds because in general for a,c≥0, b,d>0, it is easily checked that $\frac{a^2}{b} + \frac{c^2}{d} \geq \frac{(a+c)^2}{b+d}$. ∎

**Corollary 1**:

a) If simple G has a perfect matching M, no edge of M in a 3-cycle and no two edges of M in a 4-cycle, then $r(G) \geq \frac{1}{2} + \frac{|V|}{4|E|} = \frac{1}{2} + \frac{1}{2\,\bar{d}(G)}$.

b) $\rho$(SIMPLE $\Delta$-GRAPHS OF GIRTH AT LEAST 5) $\geq \frac{1}{2} + \frac{1}{2\Delta}$.

c) $\rho$(SIMPLE PLANAR GRAPHS OF GIRTH AT LEAST 5) $\geq \frac{5}{8}$.

Proof: Suppose G is as in a). Then by Lemma 2 we have that $r(G^*) \leq r(G)$. The conditions in a) are designed so that G* will be a simple pincushion graph with density equal to that of G. Thus by Theorem 1 we have $r(G) \geq r(G^*) = \frac{1}{2} + \frac{1}{2\,\bar{d}(G^*)} = \frac{1}{2} + \frac{1}{2\,\bar{d}(G)} = \frac{1}{2} + \frac{|V|}{4|E|}$, proving a).

Given a simple $\Delta$-graph H of girth at least 5, let M be a maximum matching and let G = H[M]. Since G will be a $\Delta$-graph, $2\,|E(G)| \leq \Delta\,|V(G)|$. Since G will satisfy the hypotheses of a), from Lemma 1b we have that $r(H) \geq r(G) \geq \frac{1}{2} + \frac{|V(G)|}{4|E(G)|} \geq \frac{1}{2} + \frac{1}{2\Delta}$.

Given a simple planar graph H of girth at least 5, let M be a maximum matching and let G = H[M], with G having 2n vertices. Since G* will be a planar graph whenever G is a planar graph, it follows by standard properties of simple planar graphs that $|E(C(G^*))| \leq 3n$, and hence $\bar{d}(G^*) \leq 4$. Since G will satisfy the hypotheses of a), from Lemma 1b we have that

$r(H) \geq r(G) \geq r(G^*) \geq \frac{1}{2} + \frac{1}{2\,\bar{d}(G^*)} \geq \frac{5}{8}$. ∎

We pause to compare the results of Theorem 1 and Corollary 1 with results from [3]. For H a simple forest, let M be a maximum matching, and let G = H[M], a simple forest having a perfect matching. By Corollary 1a, $r(H) \geq r(G) \geq \frac{1}{2} + \frac{|V(G)|}{4|E(G)|} \geq \frac{3}{4}$, so $\rho$(SIMPLE FORESTS) $\geq .75$. This bound comes very close to the impressive exact result v) for $\rho$(SIMPLE FORESTS) in [3], which works out to roughly .7690397. So, this application of Corollary 1a can be interpreted to say that the fact that RGA performs well on simple forests is due mainly to the fact that forests have less edges than vertices, and has little to do with the acyclic defining feature of forests.

When applied to the restricted class of simple $\Delta$-graphs of girth at most 5, the result iv) from [3] that $\rho$(SIMPLE $\Delta$-GRAPHS) $\geq \frac{\Delta}{2\Delta - 1} = \frac{1}{2} + \frac{1}{4\Delta - 2}$ isn't as good as the lower bound $\frac{1}{2} + \frac{1}{2\Delta}$ given in Corollary 1b). Similarly for Corollary 1c).

Consider the bound i) from [3], that $\rho(\mathcal{G}) \geq \dfrac{1}{2 - \frac{1}{2}\kappa(\mathcal{G})}$ where $\kappa(\mathcal{G}) = \inf_{G \in \mathcal{G}} \{\frac{|V|}{|E|} : |E| > 0\}$, with $\mathcal{G}$ closed under vertex deletion. For a nontrivial simple graph G in such a class $\mathcal{G}$, with G also as in Theorem 1, we have $\dfrac{1}{2 - \frac{1}{2}\kappa(\mathcal{G})} \leq \dfrac{1}{2 - \frac{1}{2}\frac{|V(G)|}{|E(G)|}} = \dfrac{1}{2 - \frac{n}{n+\varepsilon}} = \dfrac{n+\varepsilon}{n+2\varepsilon}$

$= \frac{1}{2} + \frac{n}{2(n+2\varepsilon)} < \frac{1}{2} + \frac{n}{2(n+\varepsilon)}$. Therefore the lower bound for r(G) given in Theorem 1 is an improvement, regardless of the choice of class $\mathcal{G}$ containing G. However, Theorem 1 doesn't apply to all simple graphs having a perfect matching, it being only applicable to pincushion graphs (although we expect that pincushion graphs are good candidates for graphs with low r values). The corollary to the next result will effectively match or improve upon the bound $\rho(\mathcal{G}) \geq \dfrac{1}{2 - \frac{1}{2}\kappa(\mathcal{G})}$ of [3], and will apply to all simple graphs having a perfect matching, and is also independent of the choice of $\mathcal{G}$.

**Theorem 2**: For G a pincushion graph on 2n vertices with total core weight ε, $\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n+2\varepsilon}\right)$. Consequently, $r(G) \geq \frac{1}{2} + \frac{n}{2(n+2\varepsilon)}$.

<u>Proof</u>: Let $f_n(\varepsilon) = \frac{n}{2}\left(1 + \frac{n}{n+2\varepsilon}\right)$ for each natural number n. For the case n=1, G must have ε=0, so the claim is true since $\mu(G) = 1 = f_n(0)$. For the case n=2, E(C) consists of a single edge of weight ε, and as in Theorem 1, $\mu(G) = 1 + \frac{2}{2+\varepsilon} \geq f_n(\varepsilon)$, so again the claim is true, completing our basis for induction. Inductively suppose the claim is true for all pincushion graphs on fewer than 2n vertices for some n >2, and let G be a pincushion graph on 2n vertices. By Lemma 4b we have \o(ε,¯)(G) ≤ $\frac{(n-2)\varepsilon}{n}$. Since $f_{n-2}(\varepsilon)$ is nonincreasing in ε, we have that $f_{n-2}(\bar{\varepsilon}(G)) \geq f_{n-2}\left(\frac{(n-2)\varepsilon}{n}\right)$. Since each of $f_{n-1}(\varepsilon)$ and $f_{n-2}(\varepsilon)$ are concave up on $[0,\infty)$, we have from Lemma 5 that

$$\mu(G) \geq 1 + \frac{1}{n+\varepsilon}\left[ n\, f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon\, f_{n-2}(\bar{\varepsilon}(G)) \right] \geq 1 + \frac{1}{n+\varepsilon}\left[ n\, f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon\, f_{n-2}\left(\frac{(n-2)\varepsilon}{n}\right) \right]$$

$$= 1 + \frac{1}{n+\varepsilon}\left[ \frac{n(n-1)}{2}\left(1 + \frac{n-1}{n-1 + \frac{2(n-2)\varepsilon}{n}}\right) + \frac{\varepsilon(n-2)}{2}\left(1 + \frac{n-2}{n-2 + \frac{2(n-2)\varepsilon}{n}}\right) \right]$$

$$= \frac{n}{2} + \frac{1}{2(n+\varepsilon)}\left[ n + \frac{n(n-1)^2}{n-1 + \frac{2(n-2)\varepsilon}{n}} + \frac{\varepsilon n(n-2)}{n+2\varepsilon} \right].$$

So to complete the inductive proof that $\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n+2\varepsilon}\right) = \frac{n}{2} + \frac{1}{2(n+2\varepsilon)} n^2$, it suffices to show that $\frac{1}{n+\varepsilon}\left[ n + \frac{n(n-1)^2}{n-1 + \frac{2(n-2)\varepsilon}{n}} + \frac{\varepsilon n(n-2)}{n+2\varepsilon} \right] \geq \frac{n^2}{n+2\varepsilon}$. Simplification shows this equivalent to $(n-1)(n+2\varepsilon) \geq n(n-1+2\varepsilon - \frac{4\varepsilon}{n})$, which is in turn equivalent to $2\varepsilon \geq 0$, a true claim. ∎

**Corollary 2**: For any simple G with a perfect matching, $r(G) \geq \dfrac{1}{2 - \frac{1}{2}\frac{|V|}{|E|}} = \dfrac{\bar{d}}{2\bar{d} - 1}$.

Proof: Apply Theorem 2 to the pincushion graph G* on 2n = |V(G)| many vertices and total core weight ε to get $r(G) \geq r(G^*) \geq \frac{1}{2} + \frac{n}{2(n+2\varepsilon)} = \frac{n+\varepsilon}{n+2\varepsilon} = \dfrac{1}{2 - \frac{n}{n+\varepsilon}} = \dfrac{1}{2 - \frac{1}{2}\frac{|V|}{|E|}}$. ∎

Corollary 2 can give us the same lower bounds for ρ(SIMPLE PLANAR GRAPHS) and ρ(SIMPLE Δ-GRAPHS) that were obtained in [3], but we can make modest improvements in each of those bounds by employing the following lemma. Observe that the lower bound it gives for suitable μ(G) is between the lower bounds given in Theorems 1 and 2.

**Lemma 6**: Let A be a constant, $1 \leq A < 2$, and let $\mathcal{G}$ be a class of pincushion graphs closed under deletion of any pair of vertices at ends of the same pin edge. If every graph G in $\mathcal{G}$ has a core of total weight bounded by $\varepsilon \leq \frac{2A-2}{2A-A^2}\frac{n^2-n}{n-2}$ for $n \geq 3$ (where 2n = |V(G)|), then $\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n+A\varepsilon}\right)$ for every $G \in \mathcal{G}$.

Proof: Let $f_n(\varepsilon) = \frac{n}{2}\left(1 + \frac{n}{n+A\varepsilon}\right)$ for each natural number n. For the case n=1, G must have ε=0, so the claim is true since $\mu(G) = 1 = f_n(0)$. For the case n=2, E(C) consists of a single edge of weight ε, and $\mu(G) = 1 + \frac{2}{2+\varepsilon} \geq f_n(\varepsilon)$, so again the claim is true, completing our basis for induction. Inductively suppose the claim is true for all pincushion graphs in $\mathcal{G}$ on fewer than 2n vertices for some n > 2, and let $G \in \mathcal{G}$ be a pincushion graph on 2n vertices. By closure of the class $\mathcal{G}$, we know that removing from G two vertices at the ends of a single pin edge or at the ends of a single core edge results in a graph in $\mathcal{G}$ (in the case of the core edge by applying closure twice). Hence by induction we have $\mu(G\backslash u) \geq f_{n-1}(\varepsilon(G\backslash u))$ for each core vertex u and $\mu(G\backslash\{u,v\}) \geq f_{n-2}(\varepsilon(G\backslash\{u,v\}))$ for each core edge uv. By Lemma 4b we have $\bar{\varepsilon}(G) \leq \frac{(n-2)\varepsilon}{n}$, and since $f_{n-2}(\varepsilon)$ is nonincreasing in ε, we have that $f_{n-2}(\bar{\varepsilon}(G)) \geq f_{n-2}(\frac{(n-2)\varepsilon}{n})$. Since each of $f_{n-1}(\varepsilon)$ and $f_{n-2}(\varepsilon)$ are concave up on $[0,\infty)$, we have from Lemma 5 that

$$\mu(G) \geq 1 + \frac{1}{n+\varepsilon}\left[ n f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon f_{n-2}(\bar{\varepsilon}(G)) \right] \geq 1 + \frac{1}{n+\varepsilon}\left[ n f_{n-1}\left(\frac{(n-2)\varepsilon}{n}\right) + \varepsilon f_{n-2}\left(\frac{(n-2)\varepsilon}{n}\right) \right]$$

$$= 1 + \frac{1}{n+\varepsilon}\left[ \frac{n(n-1)}{2}\left(1 + \frac{n-1}{n-1+\frac{A(n-2)\varepsilon}{n}}\right) + \frac{\varepsilon(n-2)}{2}\left(1 + \frac{n-2}{n-2+\frac{A(n-2)\varepsilon}{n}}\right) \right]$$

$$= \frac{n}{2} + \frac{1}{2(n+\varepsilon)}\left[ n + \frac{n^2(n-1)^2}{n^2-n+A\varepsilon n-2A\varepsilon} + \frac{n\varepsilon(n-2)}{n+A\varepsilon} \right].$$

So to complete the inductive proof that $\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n+A\varepsilon}\right) = \frac{n}{2} + \frac{1}{2(n+A\varepsilon)}n^2$, it suffices to show that $n + \frac{n^2(n-1)^2}{n^2-n+A\varepsilon n-2A\varepsilon} + \frac{n\varepsilon(n-2)}{n+A\varepsilon} \geq \frac{n^2(n+\varepsilon)}{n+A\varepsilon}$. Simplifying, we see that this inequality is equivalent to

$$n^4+A\varepsilon n^3-2n^3-2A\varepsilon n^2+n^2+A\varepsilon n \geq n^4-n^3+2\varepsilon n^2-A\varepsilon n^2-n^3+n^2-2\varepsilon n+A\varepsilon n$$
$$+A\varepsilon n^3-A\varepsilon n^2+2A\varepsilon^2 n-A^2\varepsilon^2 n-2A\varepsilon n^2+2A\varepsilon n-4A\varepsilon^2+2A^2\varepsilon^2.$$

Now dividing by $\varepsilon$, simplifying, and gathering together terms involving $\varepsilon$, the last inequality is seen to be equivalent to $(2A-2)(n^2-n) \geq \varepsilon(2A-A^2)(n-2)$ which is the same as $\frac{2A-2}{2A-A^2}\frac{n^2-n}{n-2} \geq \varepsilon$. ∎

**Theorem 3**:

a) Suppose $\Delta$-graph G for $\Delta \geq 2$ has a perfect matching each edge of which has weight 1, and that its corresponding pincushion graph G* has 2n vertices and core weight $\varepsilon$. Then
$$\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n + \left(\frac{\Delta-2+\sqrt{(\Delta-1)^2+1}}{\Delta-1}\right)\varepsilon}\right).$$

b) $\rho(\text{SIMPLE } \Delta\text{-GRAPHS}) \geq \frac{1}{2} + \frac{\sqrt{(\Delta-1)^2+1} - (\Delta-1)}{2}$ for each natural number $\Delta$.

Proof: Let $\mathcal{G}$ be the class of all pincushion graphs H whose core graph C(H) is a $(2\Delta-2)$-graph. Since deletion of a vertex in a graph cannot yield a graph with higher maximum degree, it follows that $\mathcal{G}$ is closed under deletion of a pair of vertices at the ends of a pin edge. Also, for any $H \in \mathcal{G}$ on 2n vertices we have $2\varepsilon(H) \leq (2\Delta-2)n$, so $\varepsilon(H) \leq (\Delta-1)n \leq (\Delta-1)\frac{n^2-n}{n-2} = \frac{2A-2}{2A-A^2}\frac{n^2-n}{n-2}$ for the choice $A = \frac{\Delta-2+\sqrt{(\Delta-1)^2+1}}{\Delta-1}$, this A chosen as a solution to $\Delta-1 = \frac{2A-2}{2A-A^2}$, i.e. a solution to the quadratic $(\Delta-1)A^2 + (4-2\Delta)A - 2 = 0$. Clearly $1 \leq \frac{2\Delta-3}{\Delta-1} < A < 2$. By Lemma 6, for every $H \in \mathcal{G}$ we have $\mu(H) \geq \frac{n}{2}\left(1 + \frac{n}{n+A\varepsilon}\right)$. For $\Delta$-graph G as in a), its corresponding pincushion graph H = G* has a core which is a $(2\Delta-2)$-graph, so a) follows from the fact that $\mu(G) \geq \mu(G^*)$.

For b), the case $\Delta=1$ is trivial, so assume $\Delta \geq 2$. Given a simple $\Delta$-graph H, let M be a maximum matching and let G = H[M], with $|V(G)| = 2n$. Since G will be a $\Delta$-graph, $2\,|E(G)| \leq \Delta\,|V(G)| = 2n\Delta$, so $\varepsilon = \varepsilon(G^*) = |E(G)| - n \leq (\Delta-1)n$, so $\frac{\varepsilon}{n} \leq \Delta-1$. Since G will satisfy the hypotheses of a), from Lemma 1b and the lower bound of a) we have that
$$r(H) \geq r(G) \geq \frac{1}{2} + \frac{1}{2}\,\frac{1}{1 + \left(\frac{\Delta-2+\sqrt{(\Delta-1)^2+1}}{\Delta-1}\right)\frac{\varepsilon}{n}}$$
$$\geq \frac{1}{2} + \frac{1}{2}\,\frac{1}{1 + \Delta-2 + \sqrt{(\Delta-1)^2+1}} = \frac{1}{2} + \frac{\sqrt{(\Delta-1)^2+1} - (\Delta-1)}{2}.$$ ∎

**Theorem 4**:

a) For G simple and planar with a perfect matching M, if its corresponding pincushion graph G* has 2n vertices and core weight $\varepsilon$, then $\mu(G) \geq \frac{n}{2}\left(1 + \frac{n}{n + \left(\frac{4+\sqrt{26}}{5}\right)\varepsilon}\right)$.

b) $\rho(\text{SIMPLE PLANAR GRAPHS}) \geq \frac{\sqrt{26} - 4}{2} \approx .5495$.

Proof: Let $\mathcal{G} = \{ G^*[M'] : M' \subseteq M \}$. Note that each $G^*[M']$ really is the pincushion graph corresponding to the planar graph $G[M']$. Clearly $\mathcal{G}$ is closed under deletion of a pair of vertices at the ends of a pin edge. Each $G[M']$ on 2n vertices is simple and planar, so $|E(G[M'])| \leq 6n$. Since $|E(G[M'])| = \omega(E(G^*[M']))$ it follows that $n + \varepsilon(G^*[M']) \leq 6n$, so $2\varepsilon(G^*[M']) \leq 10n$. We use the same choice of $A = \frac{4 + \sqrt{5^2 + 1}}{5}$ that we used in the proof of Theorem 3 that corresponded to the choice $\Delta=6$, i.e. $2\Delta\text{-}2 = 10$. Letting $G^*[M']$ play the role of H in that proof, it is still true as it was in that proof that $\varepsilon(G^*[M']) = \varepsilon(H) \leq (\Delta\text{-}1)n$. The rest of the proof of Theorem 3 is then equally applicable here, giving us lower bounds identical to those in Theorem 3 for the case $\Delta = 6$. ∎

We now prove an upper bound for $\rho$ for a class of graphs, by considering what appears to be a graph with relatively low $\mu$ value for that class.

**Theorem 5**: $\rho$(SIMPLE PLANAR GRAPHS) $\leq .68436349$.

Computer-assisted proof: We devised a computer program to compute $\mu$ for arbitrary graphs, and ran it on the simple pincushion graph G whose core is isomorphic to the icosahedron. This G is planar since the icosahedron is planar, where the additional pin edges can be easily incorporated into a planar embedding of the icosahedron to obtain a planar embedding of G. Given an edge $e_1$ of G, our program ran sequentially through all sequences of distinct edges $e_1,e_2,...,e_k$ for which $M = \{e_1,e_2,...,e_k\}$ is a maximal matching in G, along the way computing the probability that such a sequence is produced by RGA applied to G (introducing at most a roundoff error of $5 \times 10^{-22}$). These probabilities were totaled in seven groups to obtain the various probabilities $p_k$ (k=6,7,8,9,10,11,12) that RGA when applied to G produces a maximal matching of size k in which the first selected edge is $e_1$. We ran this once when $e_1$ was a particular pin edge and once when $e_1$ was a particular core edge, and used the vertex-transitivity and edge-transitivity of the icosahedron to aid us in combining the results to obtain a grand total for an estimate for $\mu(G)$. This grand total is the sum of precisely 3,437,380,080 terms (one for each permutation of the edges in every possible maximal matching), each term of which was off by at most $6 \times 10^{-21}$, where the error in adding each term to the growing subtotal was at most $5 \times 10^{-18}$. Therefore our estimate for $\mu(G)$ was off by no more than $2 \times 10^{-8}$, and upon dividing by $m(G) = 12$ to obtain an estimate of $r(G)$ we arrived at the upper bound stated. Anyone interested can contact the authors for further documentation. ∎

## 6. AN INTERESTING WEIGHTED PINCUSHION GRAPH

Dyer and Frieze [3] considered the simple pincushion graph $G_n$ whose core is the complete graph $K_n$, and they proved that $\mu(G_n) = \sum_{0\leq i\leq n} u_i$, where $u_n = \frac{1}{2} + \frac{1}{2n}$ if n is odd,

$u_n = \frac{1}{2} + \frac{1}{2(n+1)}$ if n is even. Thus they were able to prove that $\mu(G_n) = \frac{1}{2}(n + \log 2n + \gamma - 1) + o(1)$ [where $\gamma$ is Euler's constant], and from there could easily prove that $\rho(\text{SIMPLE GRAPHS}) = \frac{1}{2}$.

We have found the following weighted generalization of $G_n$ to be an intriguing case. Let $G_n(x)$ denote the pincushion graph on 2n vertices whose core is complete, each core edge having weight equal to x. Thus $G_n = G_n(1)$. Let $f_n(x)$ denote $\mu(G_n(x))$. In this section we calculate $f_n(x)$ explicitly, and use the resulting expression to obtain lower bounds on $\mu(G)$, for certain graphs G, which are better than those derivable from Theorem 2.

On applying the RGA to $G_n(x)$, the first edge chosen will be a pin edge with probability $\frac{n}{\binom{n}{2}x+n}$ and will be a core edge with probability $1 - \frac{n}{\binom{n}{2}x+n}$. Hence (much as in the first sentence of the proof of Lemma 5) we obtain the following recurrence for the sequence of functions $f_n(x)$, n $\geq 0$, with $f_0(x) = 0$ and $f_1(x) = 1$;

$$f_n(x) = 1 + \left(1 - \frac{n}{\binom{n}{2}x+n}\right) f_{n-2}(x) + \left(\frac{n}{\binom{n}{2}x+n}\right) f_{n-1}(x). \qquad \text{(R)}$$

**Theorem 6**:

a) The functions $f_r(x)$ have an expansion of the form

$$f_r(x) = \sum_{i=0}^{r-1} \frac{\alpha_i(r)}{\frac{ix}{2} + 1} \qquad \text{(A)}$$

for some coefficients $\alpha_i(r)$ depending only on i and r (and not on x).

b) The coefficients $\alpha_i(r)$ satisfy

$\alpha_0(r) = \lceil \frac{r}{2} \rceil$

$$\alpha_i(r) = (-1)^{r+i+1}\left[\binom{r-1}{i} - \binom{r-2}{i} + \binom{r-3}{i} - \ldots \pm \binom{i}{i}\right]2^{i-1} \text{ for } i \geq 1. \qquad \text{(B)}$$

<u>Proof</u>: One can check directly that the result holds for r = 0, r = 1, and r = 2. Assume inductively that a) holds for $2 \leq r \leq n$ for some choice of coefficient values $\alpha_i(r)$, $0 \leq i \leq r-1$. Applying the recurrence (R) and the inductive hypothesis we have

$$f_{n+1}(x) = 1 + \left(1 - \frac{1}{\frac{nx}{2}+1}\right) \sum_{i=0}^{n-2} \frac{\alpha_i(n-1)}{\frac{ix}{2}+1} + \left(\frac{1}{\frac{nx}{2}+1}\right) \sum_{i=0}^{n-1} \frac{\alpha_i(n)}{\frac{ix}{2}+1}$$

$$= 1 + \alpha_0(n-1) + \frac{\alpha_0(n) - \alpha_0(n-1)}{\frac{nx}{2}+1} + \sum_{i=1}^{n-2} \frac{\alpha_i(n-1)}{\frac{ix}{2}+1} - \left(\frac{1}{\frac{nx}{2}+1}\right) \sum_{i=1}^{n-2} \frac{\alpha_i(n-1)}{\frac{ix}{2}+1} +$$

$$+ \left(\frac{1}{\frac{nx}{2}+1}\right) \sum_{i=1}^{n-1} \frac{\alpha_i(n)}{\frac{ix}{2}+1}. \qquad \text{(C)}$$

Viewing the right hand side of (C) as a sum of six terms, we can repeatedly apply the identity

$$\frac{1}{\left(\frac{nx}{2}+1\right)\left(\frac{ix}{2}+1\right)} = \frac{\frac{n}{n-i}}{\frac{nx}{2}+1} + \frac{\frac{i}{i-n}}{\frac{ix}{2}+1}$$

to rewrite the sum of the last two terms of the six as

$$\left(\frac{1}{\frac{nx}{2}+1}\right)\sum_{i=1}^{n-2} \frac{n}{n-i}\,(\alpha_i(n) - \alpha_i(n-1)) + \sum_{i=1}^{n-2} \left(\frac{1}{\frac{ix}{2}+1}\right) \frac{i}{i-n}\,(\alpha_i(n) - \alpha_i(n-1)) + \frac{-(n-1)\alpha_{n-1}(n)}{\frac{(n-1)x}{2}+1} + \frac{n\alpha_{n-1}(n)}{\frac{nx}{2}+1}.$$

Putting this into the right hand side of equation (C) and collecting terms with like denominators we get

$$f_{n+1}(x) = 1 + \alpha_0(n-1) + \sum_{i=1}^{n-2} \left(\frac{1}{\frac{ix}{2}+1}\right) \Big[\alpha_i(n-1) + \frac{i}{i-n}\,(\alpha_i(n) - \alpha_i(n-1))\Big] + \frac{-(n-1)\alpha_{n-1}(n)}{\frac{(n-1)x}{2}+1} +$$

$$+ \left(\frac{1}{\frac{nx}{2}+1}\right)\Big(\sum_{i=0}^{n-2} \Big[\frac{n}{n-i}\,(\alpha_i(n) - \alpha_i(n-1))\Big] + n\alpha_{n-1}(n)\Big).$$

Thus $f_{n+1}(x)$ has an expansion of form (A), completing the proof of a). To prove b), note that the above expression for $f_{n+1}(x)$ gives us a recursively defined choice for the coefficients $\alpha_i(n+1)$, as follows (with initial conditions given as in (B), as already checked for r=0,1,2):

$$\alpha_0(n+1) = 1 + \alpha_0(n-1) \qquad \text{(R1)}$$

$$\alpha_i(n+1) = \alpha_i(n-1) + \frac{i}{i-n}\,(\alpha_i(n) - \alpha_i(n-1)),\ \ 1 \le i \le n-2 \qquad \text{(R2)}$$

$$\alpha_{n-1}(n+1) = -(n-1)\alpha_{n-1}(n) \qquad \text{(R3)}$$

$$\alpha_n(n+1) = \sum_{i=0}^{n-2} \Big[\frac{n}{n-i}\,(\alpha_i(n) - \alpha_i(n-1))\Big] + n\alpha_{n-1}(n). \qquad \text{(R4)}$$

Claim b) holds for r=0,1,2. Inductively suppose that b) holds for $2 \le r \le n$ where the coefficients $\alpha_i(r)$ are recursively defined by (R1), (R2), (R3), (R4). We show that b) holds for r=n+1.

By (R1) we have $\alpha_0(n+1) = 1 + \lceil \frac{n-1}{2} \rceil = \lceil \frac{n+1}{2} \rceil$, as required. We now consider $\alpha_i(n+1)$ for $1 \le i \le n$.

For i = n-1, note that $\alpha_{n-1}(n) = 2^{n-2}$ by the induction hypothesis. By (R3) we have $\alpha_{n-1}(n+1) = -(n-1)2^{n-2}$, as required by the right side of (B) for r = n+1 and i = n-1.

Suppose now that i = n. Using formulas (B) with r = n and r = n-1 we have by the induction hypothesis that $\alpha_i(n) - \alpha_i(n-1) = (-1)^{n+i+1}\binom{n-1}{i}2^{i-1}$. Hence applying (R4) and $\frac{n}{n-i}\binom{n-1}{i} = \binom{n}{i}$ we get

$$\alpha_n(n+1) = \sum_{i=1}^{n-2} (-1)^{n+i+1}\binom{n}{i}2^{i-1} + \alpha_0(n) - \alpha_0(n-1) + n2^{n-2}$$

$$= (-1)^{n+1}\frac{1}{2}\left(\sum_{i=1}^{n-2}\binom{n}{i}(-2)^i\right) + \alpha_0(n) - \alpha_0(n-1) + n2^{n-2}$$

$$= (-1)^{n+1}\frac{1}{2}\left( (-1)^n - [\, 1 + (-1)^{n-1}n2^{n-1} + (-1)^n 2^n \,] \right) + \alpha_0(n) - \alpha_0(n-1) + n2^{n-2}$$

$$= 2^{n-1},$$

where the next to last line is obtained from the binomial expansion of $(-2+1)^n = (-1)^n$, and the last line can be checked separately for n even and n odd using the already proved $\alpha_0(r) = \lceil \frac{r}{2} \rceil$ for all r.

This completes the proof of (B) for r = n+1 and i = n, and it now remains to prove (B) for r = n+1 and $1 \le i \le n-2$. Applying the inductive hypothesis and collecting terms with like coefficients we have $n\alpha_i(n-1) - i\alpha_i(n) = (-1)^{n+i}2^{i-1}\left( i\binom{n-1}{i} + (n-i)[\, \binom{n-2}{i} - \binom{n-3}{i} + \binom{n-4}{i} + \ldots \pm \binom{i}{i} \,] \right)$. Thus using (R2) we get

$$\alpha_i(n+1) = \frac{1}{n-i}\left(n\alpha_i(n-1) - i\alpha_i(n)\right)$$

$$= (-1)^{n+i}2^{i-1}\left( \frac{i}{n-i}\binom{n-1}{i} + [\, \binom{n-2}{i} - \binom{n-3}{i} + \binom{n-4}{i} + \ldots \pm \binom{i}{i} \,] \right)$$

$$= (-1)^{n+i}2^{i-1}\left( \binom{n}{i} - \binom{n-1}{i} + \binom{n-2}{i} - \binom{n-3}{i} + \binom{n-4}{i} + \ldots \pm \binom{i}{i} \right)$$

as required. The inductive step is completed, thereby proving the theorem. ∎

We will see that Theorem 6 yields lower bounds for the $\mu$ value of certain graphs which improve the lower bounds for $\mu$ implied by Theorem 2. This will follow from the following lemma and a construction to be given afterwards. We continue with the notation $f_n(x) = \mu(G_n(x))$ for the function calculated in Theorem 6, and we let $P_n(x) = \frac{n}{2}\left(\frac{(n-1)x + 4}{(n-1)x + 2}\right)$.

**Lemma 7**: We have $f_n(x) > P_n(x)$ for x = 1,2, or 3 and n sufficiently large.

<u>Proof</u>: Let $g_n(x) = f_n(x) - f_{n-1}(x)$ for $n \ge 2$. From the basic recurrence for the $f_n$'s we get $g_n(x) = 1 - \left(\frac{(n-1)x}{(n-1)x + 2}\right)g_{n-1}(x)$. An easy induction then gives the following formulas;

$$g_n(1) = \begin{cases} \frac{1}{2} + \frac{1}{2(n+1)} & \text{for n even} \\ \frac{1}{2} + \frac{1}{2n} & \text{for n odd} \end{cases}, \text{ and}$$

$$g_n(2) = \begin{cases} \frac{1}{2} & \text{for n even} \\ \frac{1}{2} + \frac{1}{2n} & \text{for n odd.} \end{cases}$$

For any positive integer $n \geq 3$, we let $S(n) = \sum_{\substack{3 \leq k \leq n \\ k \text{ odd}}} \frac{1}{k}$.

Suppose first $x=1$. Although $f_n(1)$ was computed in [3], we recompute it here for the sake of completeness along lines that will ease the transition to the subsequent calculation of $f_n(2)$. Assuming first that n is even, we have

$$\begin{aligned} f_n(1) &= f_2(1) + \sum_{3 \leq k \leq n} g_k(1) \\ &= \frac{5}{3} + \frac{n}{2} - 1 + \frac{1}{2}\Bigg( \sum_{\substack{4 \leq k \leq n \\ k \text{ even}}} \frac{1}{k+1} + \sum_{\substack{3 \leq k \leq n \\ k \text{ odd}}} \frac{1}{k} \Bigg) \\ &= \frac{2}{3} + \frac{n}{2} + \frac{1}{2}\left[2S(n) - \frac{1}{3} + \frac{1}{n+1}\right]. \end{aligned}$$

Thus the inequality $f_n(1) > P_n(1)$ for n even is equivalent to

$$\frac{2}{3} + \frac{n}{2} + \frac{1}{2}\left[2S(n) - \frac{1}{3} + \frac{1}{n+1}\right] > \frac{n}{2}\left(1 + \frac{2}{n+1}\right),$$

which obviously holds for n sufficiently large. In fact simplification shows the latter inequality equivalent to $\frac{3}{2(n+1)} + S(n) > \frac{1}{2}$, which is already true for $n \geq 4$. For n odd with $n \geq 5$, we have

$$\begin{aligned} f_n(1) &= f_2(1) + \sum_{3 \leq k \leq n} g_k(1) \\ &= \frac{5}{3} + \frac{n+1}{2} - \frac{3}{2} + \frac{1}{2}\Bigg( \sum_{\substack{4 \leq k \leq n \\ k \text{ even}}} \frac{1}{k+1} + \sum_{\substack{3 \leq k \leq n \\ k \text{ odd}}} \frac{1}{k} \Bigg) \\ &= \frac{1}{6} + \frac{n+1}{2} + \frac{1}{2}\left[2S(n) - \frac{1}{3}\right] \\ &= \frac{n+1}{2} + S(n). \end{aligned}$$

Thus to verify $f_n(1) > P_n(1)$ for n odd we are led to

$$\frac{n+1}{2} + S(n) > \frac{n}{2}\left(1 + \frac{2}{n+1}\right).$$

Again this is clearly true for n large enough, and simplification leads to $\frac{1}{n+1} + S(n) > \frac{1}{2}$ which holds for $n \geq 5$.

Consider now the case $x = 2$. Supposing first that n is even we get

$$\begin{aligned} f_n(2) &= f_2(2) + \sum_{3 \leq k \leq n} g_k(2) \\ &= \frac{3}{2} + \frac{n}{2} - 1 + \frac{1}{2}\Bigg( \sum_{\substack{3 \leq k \leq n \\ k \text{ odd}}} \frac{1}{k} \Bigg) \\ &= \frac{n+1}{2} + \frac{1}{2} S(n). \end{aligned}$$

The inequality $f_n(2) > P_n(2)$ is then verified for n sufficiently large by observing that

$$\frac{n+1}{2} + \frac{1}{2} S(n) > \frac{n}{2}\left(1 + \frac{1}{n}\right)$$

holds for n large enough. For n odd we similarly get

$f_n(2) = f_2(2) + \frac{n+1}{2} - \frac{3}{2} + \frac{1}{2} S(n) = \frac{n+1}{2} + \frac{1}{2} S(n)$, and again it is easily seen that $f_n(2) > P_n(2)$ holds for n sufficiently large.

Finally we pass to the case x = 3. Here we did not find a simple closed form for $g_n(3)$, and the proof relies on the recurrence R for the $f_n$'s and induction. The inequality $f_k(3) > P_k(3)$ being clearly true for small k, assume inductively that the same inequality holds for k < n and we will prove that it holds for k = n. Then we have

$$f_n(3) = 1 + \frac{1}{\frac{3(n-1)}{2} + 1} \left( \frac{3(n-1)}{2} f_{n-2}(3) + f_{n-1}(3) \right) \qquad \text{(by recurrence R)}$$

$$\geq 1 + \frac{1}{\frac{3(n-1)}{2} + 1} \left( \frac{3(n-1)}{2} \frac{(n-2)}{2} \frac{(3n-3)}{(3n-5)} + \frac{(n-1)}{2} \frac{3n}{(3n-2)} \right) \qquad \text{(by induction)}$$

$$= 1 + \frac{3}{2} \left( \frac{9n^4 - 36n^3 + 53n^2 - 38n + 12}{27n^3 - 72n^2 + 51n - 10} \right)$$

$$= \frac{n}{2} + \frac{1}{3} + \left( \frac{5n^2 - 18n - \frac{34}{3}}{27n^3 - 72n^2 + 51n - 10} \right)$$

$$> P_n(3) \quad \text{(for n large enough)}$$

where the last line follows from observing that $P_n(3) = \frac{n}{2} \left( \frac{3n+1}{3n-1} \right) = \frac{n}{2} + \frac{1}{3} + \frac{1}{9n-3}$ . ∎

We will now show that for at least one class of graphs G Theorem 6 yields lower bounds for $\mu(G)$ which are stronger than the bound given by Theorem 2.

For any $x \in \{1,2,3\}$ and any integer $n \geq 2$, define $\Gamma_n(x)$ to be the set of all simple graphs G on 2n vertices for which;

(a) G has a perfect matching M, say comprised of some n edges $e_1, e_2, ..., e_n$ .

(b) For any integer pair i,j with $1 \leq i < j \leq n$, G has exactly x of the 4 possible distinct edges having one end a vertex of $e_i$ and the other end a vertex of $e_j$.

Also let $\Gamma(x) = \bigcup_{n=2}^{\infty} \Gamma_n(x)$ , viewed as the class of all graphs belonging to $\Gamma_n(x)$ for some n. We illustrate graphs in the sets $\Gamma_3(2)$ and $\Gamma_3(3)$ in Figure 4. We have restricted our attention to the cases $x \in \{1,2,3\}$ in order to ensure that each graph G considered is simple (hence $x \leq 4$) and not the trivial case $G = K_{2n}$ (hence $x \neq 4$).

Observe that for any graph $G \in \Gamma_n(x)$ we have $|E(G)| = n + \binom{n}{2}x$. For such a G the graph G* is just the pincushion graph with complete core, each core edge having weight x; i.e. $G^* = G_n(x)$, the graph discussed above. Hence by Lemma 2 we have $\mu(G) \geq \mu(G^*) = f_n(x)$. On the other hand, the lower bound for $\mu(G)$ given by Theorem 2, noting that $\varepsilon = \binom{n}{2}x$, is

$$\mu(G) \geq \frac{n}{2}\left( 1 + \frac{n}{n+2\binom{n}{2}x} \right).$$

But $P_n(x) = \frac{n}{2}\left(\frac{(n-1)x + 4}{(n-1)x + 2}\right) = \frac{n}{2}\left( 1 + \frac{n}{n+\binom{n}{2}x} \right) > \frac{n}{2}\left( 1 + \frac{n}{n+2\binom{n}{2}x} \right).$

Combining this with Lemma 7 we obtain the following corollary.

**Corollary 3**: Let G be a graph in the class $\Gamma(x)$, x = 1, 2, or 3, with $2n = |V(G)|$ large enough. Then the lower bound $\mu(G) \geq f_n(x)$ implied by Lemma 2 is strictly stronger than the lower bound for $\mu(G)$ provided by Theorem 2.

From $P_n(x) = \frac{n}{2}\left( 1 + \frac{n}{n+\binom{n}{2}x} \right) = \frac{n}{2}\left( 1 + \frac{n}{n+\varepsilon} \right)$, where $\varepsilon = \varepsilon(G_n(x))$, we see that $\mu(G_n(x))$ satisfies the stronger lower bound $\mu(G) \geq \frac{n}{2}\left( 1 + \frac{n}{n+\varepsilon} \right)$ provided by Theorem 1 (compared to the corresponding bound of Theorem 2) despite the fact that $G_n(x)$ does not satisfy the hypothesis of Theorem 1. So Corollary 3 is related to the fact that Theorem 1 gives a stronger bound than does Theorem 2.

In view of the corollary, it would be interesting to know if the function $f_n(x)$ computed explicitly in Theorem 6 has a simpler, possibly closed, form.

## 7. CONCLUDING REMARKS

Naturally, our results have concerned classes of graphs for which our main techniques yield interesting bounds. Thus far we have largely ignored interesting classes such as $\mathcal{G}$ = {SIMPLE k-REGULAR GRAPHS}. The following easy result, which uses none of our previous techniques, addresses this issue.

**Theorem 7**:

a) In a simple graph on n vertices, with minimum degree $\delta$ and maximum degree $\Delta$, every maximal matching M satisfies $|M| \geq \frac{n}{2}\,\frac{\delta}{\Delta+\delta-1}$.

b) In a simple k-regular graph G, every maximal matching M satisfies $|M| \geq \frac{k}{2k-1}\,m(G)$, so $r(G) \geq \frac{k}{2k-1}$.

<u>Proof</u>: Suppose M is a maximal matching in G, spanning a set X of 2|M| many vertices, and let Y be the set of n-2|M| remaining vertices. By maximality of M, Y is independent. Let $\Sigma_X$ (resp. $\Sigma_Y$) denote the sum of the degrees in G of the vertices of X (resp. Y). Then $\Sigma_Y$ is the number of edges with one end in Y, so

$$|M| + \Sigma_Y \le |E(G)| = \frac{1}{2}(\Sigma_X + \Sigma_Y) = \Sigma_Y + \frac{1}{2}(\Sigma_X - \Sigma_Y) \le \Sigma_Y + \frac{1}{2}(\Delta 2|M| - \delta(n-2|M|)).$$

Upon solving the resulting inequality for |M| we obtain a). Should G be k-regular, upon substitution of k for $\delta$ and $\Delta$ we have $|M| \ge \frac{n}{2}\,\frac{k}{2k-1}$. Since $\frac{n}{2} \ge m(G)$ we obtain b). ∎

In particular, from Theorem 7b alone we obtain that ρ(SIMPLE k-REGULAR GRAPHS) ≥ $\frac{k}{2k-1}$, matching the result iv) of [3] as applies to the restricted set of simple k-regular graphs. Better yet, we learn that it is not just the *average* performance of RGA when applied to such graphs which has this lower bound, but that *every* application of RGA to a k-regular simple graph produces a matching of size within a factor of $\frac{k}{2k-1}$ from being optimal. Of course our Theorem 3b provides us with the better lower bound $r(G) \ge \frac{1}{2} + \frac{\sqrt{(k-1)^2+1} - (k-1)}{2}$ for G any simple k-regular graph.

Concerning the number ρ(SIMPLE 3-REGULAR GRAPHS), it can be upper bounded by the r-value for the graph H in Example 2 (shown in Figure 1), and lower bounded by Theorem 3b with k = 3 to yield the following corollary. A large gap is to be expected, since simple pincushion graphs can't possibly be 3-regular. We suspect that the exact value is close to the upper bound.

**Corollary 4**: $.618 \approx \frac{\sqrt{5} - 1}{2} \le$ ρ(SIMPLE 3-REGULAR GRAPHS) ≤ .84.

In our research we have been guided by some heuristics that have proven fruitful concerning how to design a pincushion graph G on 2n vertices with core weight ε yielding relatively low r value. The use of Jensen's inequality suggests that such a G be nearly regular so that the $\varepsilon(G\backslash u)$ values in the proof of Lemma 5 are concentrated near their mean. Given that G is nearly regular, the use and proof of Lemma 4 suggest that G be chosen so that $\sum_{uv \in E(C)} (\omega(uv))^2$ is nearly maximized (given whatever other constraints we might place on G), with not much variation in the weights of the edges of C. Based on this intuition, we make the following conjectures.

**Conjecture 1**: For G a pincushion graph on 2n vertices, with core of weight ε, $r(G) \ge \frac{1}{2} + \frac{n}{2(n+\varepsilon)}$, with equality achieved if and only if n is even.

In particular, for n even this conjecture implies that RGA performs most poorly on M(n,ε) among all pincushion graphs of the same order and size (noting that $r(M(n,\varepsilon)) = \frac{1}{2} + \frac{n}{2(n+\varepsilon)}$ ). That is, ρ(PINCUSHION GRAPHS ON 2n VERTICES WITH CORE WEIGHT ε) = $\frac{1}{2} + \frac{n}{2(n+\varepsilon)}$. This conjecture also implies that, if a weighted G has a perfect matching, those matching edges being of weight 1, then $r(G) \geq \frac{1}{2} + \frac{|V|}{4|E|}$. This in turn would imply that ρ(SIMPLE Δ-GRAPHS) $\geq \frac{1}{2} + \frac{1}{2\Delta}$ and that ρ(SIMPLE PLANAR GRAPHS) $\geq \frac{7}{12}$.

For n even, $n \geq 4$, let $AC_n$ denote the pincushion graph on 2n vertices whose core is an n-cycle, every other edge of which has weight 1, the other core edges being of weight 3. Let $AC_2 = G_2(4)$.

**Conjecture 2**: For G a simple pincushion graph on 2n vertices with n even, with core C(G) being a connected 4-graph, $r(G) \geq r(AC_n)$.

Let $I = \inf \{r(AC_n)\}$. Conjecture 2 implies that ρ(SIMPLE PINCUSHION GRAPHS ON 4n VERTICES HAVING A 4-GRAPH FOR A CORE) = I. Conjecture 2 implies that ρ(SIMPLE BRIDGELESS 3-REGULAR GRAPHS) ≥ I, since if simple G is bridgeless and 3-regular then G has a perfect matching M, and each component of the corresponding G* is a pincushion graph with a 4-graph for a core, so $r(G) \geq r(G^*) \geq I$. Similarly, Conjecture 2 implies that ρ(SIMPLE 3-GRAPHS WITH A MAXIMUM MATCHING OF EVEN SIZE) ≥ I.

Of course each of these conjectures could conceivably fall to a simply chosen counterexample. The interest in our posing them is in asking for a proof of either conjecture, since such a proof would provide an exact value for $\rho(\mathcal{G})$ for an interesting class $\mathcal{G}$ and would provide a nice lower bound for the ρ value of an interesting class of simple graphs.

Here is some discussion concerning ρ(SIMPLE PLANAR GRAPHS). The upper bound of 11/15 ≈ .733333 in iii) from [3] comes from a simple planar pincushion graph in which the core is $K_4$, i.e. in which the core is a tetrahedron, a 3-regular graph. Though not previously mentioned, we were first able to reduce this bound to 3716/5265 ≈ .705793 by considering instead the simple planar pincushion graph in which the core is an octahedron, a 4-regular graph. The graph G employed in Theorem 5, yielding the best known upper bound, is the simple planar pincushion graph in which the core is an icosahedron, a 5-regular graph. While each of these cores is a Platonic solid, really we were just following the intuition from the results of this paper that a regular or nearly regular core in a relatively dense pincushion graph tends to produce relatively low r value. The core of a simple planar pincushion graph necessarily has density $\bar{d} < 6$, and the tetrahedron, octahedron and icosahedron have densities 3,4,5 respectively. Thus we suspect that the bound in Theorem 5 can be improved slightly by computing the r value of a well chosen simple planar pincushion graph with a nearly 6-regular core. Such a graph would necessarily be much larger than the graph G of Theorem 5, so that even if such a graph were chosen so as to possess symmetries useful in speeding up the computation (note that vertex-transitivity is impossible in the non-regular core of such a graph) we suspect it would be presently unfeasible to wait for the halting of a computer program such as the one we ran, when the input is such a graph. On the other hand Theorem 1 informs us that these planar simple pincushion examples always yield r values bounded below by .625, and the current lower bound for ρ(SIMPLE PLANAR GRAPHS) falls far short of that. Thus, without a significantly new idea, it appears there will continue to be a gap between the lower and upper bounds for ρ(SIMPLE PLANAR GRAPHS).

Finally for G simple with a perfect matching M we have mentioned that the core edges of G* must have weights 1,2,3 or 4, but thus far we have been unsuccessful at improving any interesting general lower bounds based on this fact.

## REFERENCES


[1] J. Aronson, M. E. Dyer, A. M. Frieze and S. Suen, Randomized greedy matching. II, *Random Struct. Alg.* **6**, 55-73 (1995).

[2] J. Aronson, A. M. Frieze and B. Pittel, Analysis of a randomized greedy matching algorithm, in preparation.

[3] M. E. Dyer and A. M. Frieze, Randomized greedy matching, *Random Struct. Alg.* **2**, 29-46 (1991).

[4] M. E. Dyer, A. M. Frieze and B. Pittel, On the average performance of the greedy algorithm for finding a matching in a graph, *Ann. Appl. Probab.*, to appear.

[5] B. Korte and D. Hausmann, An analysis of the greedy algorithm for independence systems, in *Algorithmic Aspects of Combinatorics*, B. Alspach, P. Hall and D. J. Miller, Eds., *Annals of Discrete Mathematics* **2**, 65-74 (1978).

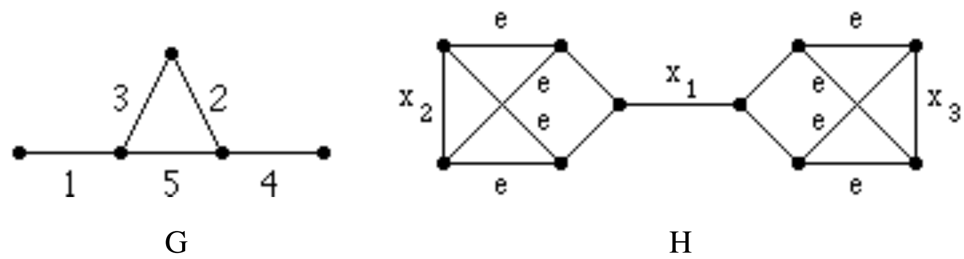



Figure 1: The graphs discussed in Example 1 and Example 2.

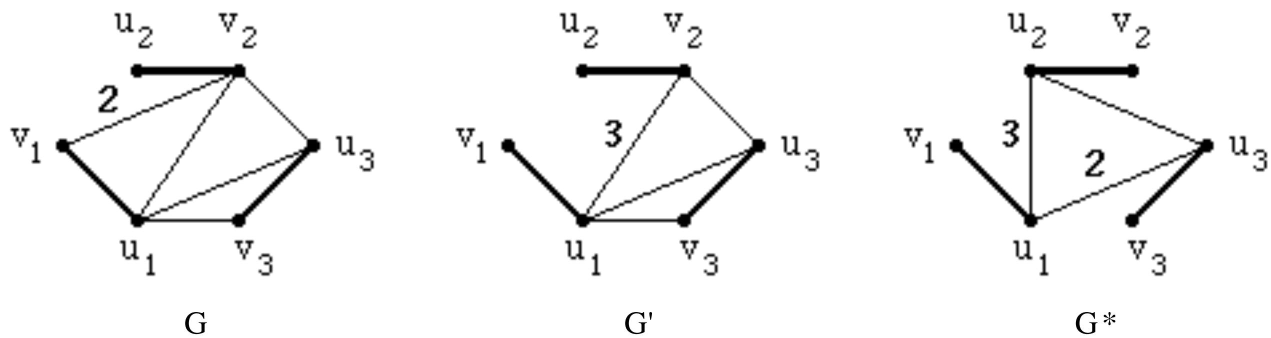
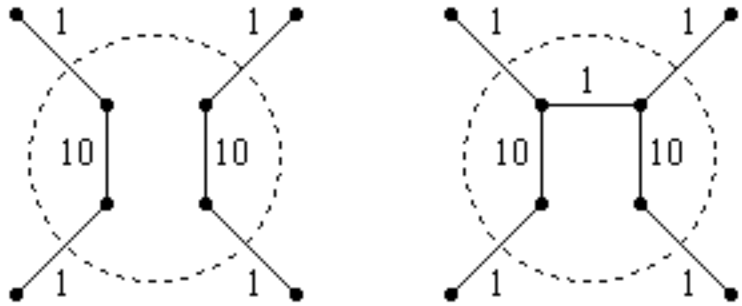
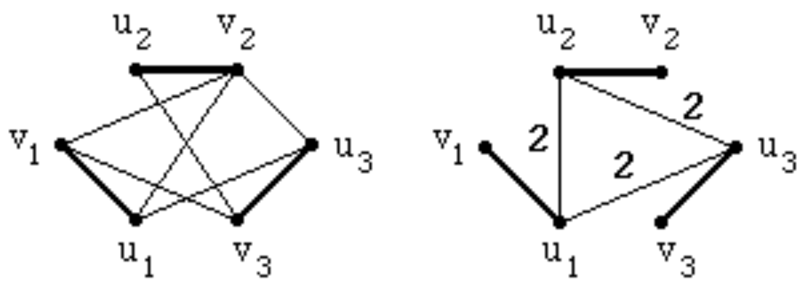



Figure 2: The graphs G' and G* obtained from an example G.
Here all edges have weight 1 except as indicated.




a) b)

Figure 3: Adding edges to the core doesn't necessarily reduce $\mu$.




$G \in \Gamma_3(2)$ $G^* \cong G_3(2)$

$H \in \Gamma_3(3)$ $H^* \cong G_3(3)$

Figure 4: Graphs in the sets $\Gamma_3(2)$ and $\Gamma_3(3)$.