Integer and fractional packing of families of graphs

Raphael Yuster *
Department of Mathematics
University of Haifa at Oranim
Tivon 36006, Israel

Abstract

Let $\mathcal{F}$ be a family of graphs. For a graph G, the *$\mathcal{F}$-packing number*, denoted $\nu_{\mathcal{F}}(G)$, is the maximum number of pairwise edge-disjoint elements of $\mathcal{F}$ in G. A function ψ from the set of elements of $\mathcal{F}$ in G to $[0,1]$ is a *fractional $\mathcal{F}$-packing* of G if $\sum_{e \in H \in \mathcal{F}} \psi(H) \leq 1$ for each $e \in E(G)$. The *fractional $\mathcal{F}$-packing number*, denoted $\nu^*_{\mathcal{F}}(G)$, is defined to be the maximum value of $\sum_{H \in \binom{G}{\mathcal{F}}} \psi(H)$ over all fractional $\mathcal{F}$-packings ψ. Our main result is that $\nu^*_{\mathcal{F}}(G) - \nu_{\mathcal{F}}(G) = o(|V(G)|^2)$. Furthermore, a set of $\nu_{\mathcal{F}}(G) - o(|V(G)|^2)$ edge-disjoint elements of $\mathcal{F}$ in G can be found in randomized polynomial time. For the special case $\mathcal{F} = \{H_0\}$ we obtain a simpler proof of a recent difficult result of Haxell and Rödl [9] that $\nu^*_{H_0}(G) - \nu_{H_0}(G) = o(|V(G)|^2)$. Their result can be implemented in deterministic polynomial time. We also prove that the error term $o(|V(G)|^2)$ is asymptotically tight.

1 Introduction

All graphs considered here are finite and have no loops, multiple edges or isolated vertices. For the standard terminology used the reader is referred to [3]. Let $\mathcal{F}$ be any fixed finite or infinite family of graphs. For a graph G, the *$\mathcal{F}$-packing number*, denoted $\nu_{\mathcal{F}}(G)$, is the maximum number of pairwise edge-disjoint copies of elements of $\mathcal{F}$ in G. Let $\binom{G}{\mathcal{F}}$ denote the set of copies of elements of $\mathcal{F}$ in G. A function ψ from $\binom{G}{\mathcal{F}}$ to $[0,1]$ is a *fractional $\mathcal{F}$-packing* of G if $\sum_{e \in H \in \binom{G}{\mathcal{F}}} \psi(H) \leq 1$ for each $e \in E(G)$. For a fractional $\mathcal{F}$-packing ψ, let $w(\psi) = \sum_{H \in \binom{G}{\mathcal{F}}} \psi(H)$. The *fractional $\mathcal{F}$-packing number*, denoted $\nu^*_{\mathcal{F}}(G)$, is defined to be the maximum value of $w(\psi)$ over all fractional packings ψ. Notice that, trivially, $\nu^*_{\mathcal{F}}(G) \geq \nu_{\mathcal{F}}(G)$. If $\mathcal{F}$ consists of a single graph H_0 we shall denote the parameters above by $\nu_{H_0}(G)$ and $\nu^*_{H_0}(G)$.

Since computing $\nu^*_{\mathcal{F}}(G)$ amounts to solving a linear program, it can be computed in polynomial time for every finite $\mathcal{F}$. On the other hand, it was proved by Dor and Tarsi in [4] that computing

*e-mail: raphy@research.haifa.ac.il World Wide Web: http://research.haifa.ac.il/~raphy

$\nu_{H_0}(G)$ is NP-Hard for every H_0 with a component having at least three edges. Thus, it is interesting to determine when $\nu^*_{\mathcal{F}}(G)$ and $\nu_{\mathcal{F}}(G)$ are "close", thereby getting a polynomial time approximating algorithm for an NP-Hard problem. The following result was proved by Haxell and Rödl in [9].

Theorem 1.1 *If H_0 is a fixed graph and G is a graph with n vertices, then $\nu^*_{H_0}(G) - \nu_{H_0}(G) = o(n^2)$.*

The proof of Theorem 1.1 presented in [9] is difficult. The major difficulty lies in the fact that their method requires proving that there is a fractional packing which is only slightly less than optimal, and which assigns to every copy of H_0 either 0 or a value greater than τ for some $\tau > 0$ which is only a function of H_0. Their method of proof also supplies a polynomial time algorithm that *finds* a set of $\nu_{H_0}(G) - o(n^2)$ edge-disjoint copies of H_0 in G.

In this paper we present a simpler and more general proof of Theorem 1.1. Our proof method enables us to generalize Theorem 1.1 to the "family" case.

Theorem 1.2 *If $\mathcal{F}$ is a fixed family of graphs and G is a graph with n vertices, then $\nu^*_{\mathcal{F}}(G) - \nu_{\mathcal{F}}(G) = o(n^2)$.*

Notice that Theorem 1.2 immediately yields a polynomial time algorithm for approximating $\nu_{\mathcal{F}}(G)$ to within an additive term of ϵn^2 for every $\epsilon > 0$. Furthermore, if $\mathcal{F}$ is finite, the degree of the polynomial depends only on $\mathcal{F}$, and not on $1/\epsilon$. Our proof also supplies a randomized polynomial time algorithm that *finds* a set of $\nu_{\mathcal{F}}(G) - o(n^2)$ edge-disjoint copies of elements of $\mathcal{F}$ in G. Our proof relies heavily on probabilistic arguments, and is consequently simpler than the proof in [9]. However, as noted above, the proof in [9] yields a deterministic algorithm.

We also prove that the $o(n^2)$ error term in Theorem 1.2, cannot, in general, be improved.

Proposition 1.3 *For every $\epsilon > 0$ there exist $k = k(\epsilon)$ and $N = N(\epsilon)$ such that for all $n > N$ there exists a graph G with n vertices for which $\nu^*_{K_k}(G) - \nu_{K_k}(G) > n^{2-\epsilon}$.*

2 Tools used in the main result

As in [9], a central ingredient in our proof of the main result is Szemerédi's regularity lemma [11]. Let $G = (V, E)$ be a graph, and let A and B be two disjoint subsets of $V(G)$. If A and B are non-empty, let $E(A, B)$ denote the set of edges between them, and put $e(A, B) = |E(A, B)|$. The *density of edges* between A and B is defined as

$$d(A, B) = \frac{e(A, B)}{|A||B|}.$$

For $\gamma > 0$ the pair (A, B) is called *γ-regular* if for every $X \subset A$ and $Y \subset B$ satisfying $|X| > \gamma|A|$ and $|Y| > \gamma|B|$ we have

$$|d(X, Y) - d(A, B)| < \gamma.$$

An *equitable partition* of a set V is a partition of V into pairwise disjoint classes $V_1, \dots, V_m$ whose sizes are as equal as possible. An equitable partition of the set of vertices V of a graph G into the classes $V_1, \dots, V_m$ is called *γ-regular* if $|V_i| < \gamma|V|$ for every i and all but at most $\gamma\binom{m}{2}$ of the pairs (V_i, V_j) are γ-regular. The regularity lemma states the following:

Lemma 2.1 *For every $\gamma > 0$, there is an integer $M(\gamma) > 0$ such that for every graph G of order $n > M$ there is a γ-regular partition of the vertex set of G into m classes, for some $1/\gamma < m < M$.* ■

Let H_0 be a fixed graph with the vertices $\{1, \dots, k\}$, $k \geq 3$. Let W be a k-partite graph with vertex classes $V_1, \dots, V_k$. A subgraph J of W with ordered vertex set $v_1, \dots, v_k$ is *partite-isomorphic* to H_0 if $v_i \in V_i$ and the map $v_i \to i$ is an isomorphism from J to H_0.

The following lemma is almost identical to the proof of Lemma 15 in [9] (which, in turn, relies on a result from [7]) and hence the proof is omitted.

Lemma 2.2 *Let δ and ζ be positive reals. There exist $\gamma = \gamma(\delta, \zeta, k)$ and $T = T(\delta, \zeta, k)$ such that the following holds. Let W be a k-partite graph with vertex classes $V_1, \dots, V_k$ and $|V_i| = t > T$ for $i = 1, \dots, k$. Furthermore, for each $(i,j) \in E(H_0)$, (V_i, V_j) is a γ-regular pair with density $d(i,j) \geq \delta$ and for each $(i,j) \notin E(H_0)$, $E(V_i, V_j) = \emptyset$. Then, there exists a spanning subgraph W' of W, consisting of at least $(1-\zeta)|E(W)|$ edges such that the following holds. For an edge $e \in E(W')$, let $c(e)$ denote the number of subgraphs of W' that are partite isomorphic to H_0 and that contain e. Then, for all $e \in E(W')$, if $e \in E(V_i, V_j)$ then*

$$\left| c(e) - t^{k-2} \frac{\prod_{(s,p) \in E(H_0)} d(s,p)}{d(i,j)} \right| < \zeta t^{k-2}.$$

■

Finally, we need to state the seminal result of Frankl and Rödl [5] on near perfect coverings and matchings of uniform hypergraphs. Recall that if x, y are two vertices of a hypergraph then $deg(x)$ denotes the degree of x and $deg(x,y)$ denotes the number of edges that contain both x and y (their *co-degree*). We use the version of the Frankl and Rödl Theorem due to Pippenger (see, e.g., [6]).

Lemma 2.3 *For an integer $r \geq 2$ and a real $\beta > 0$ there exists a real $\mu > 0$ so that: If the r-uniform hypergraph L on q vertices has the following properties for some d:*
(i) $(1-\mu)d < deg(x) < (1+\mu)d$ holds for all vertices,
(ii) $deg(x,y) < \mu d$ for all distinct x and y,
then L has a matching of size at least $(q/r)(1-\beta)$. ■

3 Proof of the main result

Let $\mathcal{F}$ be a family of graphs, and let $\epsilon > 0$. To avoid the trivial case we assume $K_2 \notin \mathcal{F}$. We shall prove there exists $N = N(\mathcal{F}, \epsilon)$ such that for all $n > N$, if G is an n-vertex graph then $\nu^*_{\mathcal{F}}(G) - \nu_{\mathcal{F}}(G) < \epsilon n^2$.

Let k_∞ denote the maximum order of a graph in $\mathcal{F}$ (it may be that $k_\infty = \infty$). Let $k_0 = \min\{k_\infty, \lceil 20/\epsilon \rceil\}$. Let $\delta = \beta = \epsilon/4$. For all $r = 2, \ldots, {k_0}^2$, let $\mu_r = \mu(\beta, r)$ be as in Lemma 2.3, and put $\mu = \min_{r=2}^{k_0^2}\{\mu_r\}$. Let $\zeta = \mu \delta^{{k_0}^2}/2$. For $k = 3, \ldots, k_0$, let $\gamma_k = \gamma(\delta, \zeta, k)$ and $T_k = T(\delta, \zeta, k)$ be as in Lemma 2.2. Let $\gamma = \min_{k=3}^{k_0}\{\gamma_k\}$. Let $M = M(\gamma\epsilon/(25{k_0}^2))$ be as in Lemma 2.1. Finally, we shall define N to be a sufficiently large constant, depending on the above chosen parameters, and for which various conditions stated in the proof below hold (it will be obvious in the proof that all these conditions indeed hold for N sufficiently large). Thus, indeed, $N = N(\mathcal{F}, \epsilon)$.

Fix an n-vertex graph G with $n > N$ vertices. Fix a fractional $\mathcal{F}$-packing ψ with $w(\psi) = \nu^*_{\mathcal{F}}(G)$. We may assume that ψ assigns a value to each *labeled* copy of an element of $\mathcal{F}$ simply by dividing the value of ψ on each nonlabeled copy by the size of the automorphism group of that element. If $\nu^*_{\mathcal{F}}(G) < \epsilon n^2$ we are done. Hence, we assume $\nu^*_{\mathcal{F}}(G) = \alpha n^2 \geq \epsilon n^2$.

We apply Lemma 2.1 to G and obtain a γ'-regular partition with m' parts, where $\gamma' = \gamma\epsilon/(25{k_0}^2)$ and $1/\gamma' < m' < M(\gamma')$. Denote the parts by $U_1, \ldots, U_{m'}$. Notice that the size of each part is either $\lfloor n/m' \rfloor$ or $\lceil n/m' \rceil$. For simplicity we may and will assume that n/m' is an integer, as this assumption does not affect the asymptotic nature of our result. For the same reason we may and will assume that $25{k_0}^2/\epsilon$ and $n/(25m'{k_0}^2/\epsilon)$ are integers.

We randomly partition each U_i into $25{k_0}^2/\epsilon$ equal parts of size $n/(25m'{k_0}^2/\epsilon)$ each. All m' partitions are independent. We now have $m = 25m'{k_0}^2/\epsilon$ *refined* vertex classes, denoted $V_1, \ldots, V_m$. Suppose $V_i \subset U_s$ and $V_j \subset U_t$ where $s \neq t$. We claim that if (U_s, U_t) is a γ'-regular pair then (V_i, V_j) is a γ-regular pair. Indeed, if $X \subset V_i$ and $Y \subset V_j$ have $|X|, |Y| > \gamma n/(25m'{k_0}^2/\epsilon)$ then $|X|, |Y| > \gamma' n/m'$ and so $|d(X,Y) - d(U_s, U_t)| < \gamma'$. Also $|d(V_i, V_j) - d(U_s, U_t)| < \gamma'$. Thus, $|d(X,Y) - d(V_i, V_j)| < 2\gamma' < \gamma$.

Let H be a labeled copy of some $H_0 \in \mathcal{F}$ in G. If H has k vertices and $k \leq k_0$ then the expected number of pairs of vertices of H that belong to the same vertex class in the refined partition is clearly at most $\binom{k}{2}\epsilon/(25{k_0}^2) < \epsilon/50$. Thus, the probability that H has two vertices in the same vertex class is also at most $\epsilon/50$. We call H *good* if it has $k \leq k_0$ vertices and its k vertices belong to k distinct vertex classes of the refined partition. By the definition of k_0 we observe that, if H has $k > k_0$ vertices, then $k_\infty \geq k > k_0$ and consequently $k > 20/\epsilon$. Since graphs with k vertices have at least $k/2$ edges, the contribution of graphs with $k > k_0$ vertices to $\nu^*_{\mathcal{F}}(G)$ is at most $\binom{n}{2}/(10/\epsilon) < \epsilon n^2/20$. Hence, if ψ^{**} is the restriction of ψ to good copies (the bad copies having $\psi^{**}(H) = 0$) then the expectation of $w(\psi^{**})$ is at least $(\alpha - \epsilon/50 - \epsilon/20)n^2$. We therefore *fix* a partition $V_1, \ldots, V_m$ for which $w(\psi^{**}) \geq (\alpha - 0.07\epsilon)n^2$.

Let G^* be the spanning subgraph of G consisting of the edges with endpoints in distinct vertex classes of the refined partition that form a γ-regular pair with density at least δ (thus, we discard edges inside classes, between non regular pairs, or between sparse pairs). Let ψ^* be the restriction of ψ^{**} to the labeled copies of elements of $\mathcal{F}$ in G^*. We claim that $\nu^*_{\mathcal{F}}(G^*) \geq w(\psi^*) > w(\psi^{**}) - 0.72\delta n^2 \geq (\alpha - 0.07\epsilon - 0.72\delta)n^2 = (\alpha-\delta)n^2$. Indeed, by considering the number of discarded edges we get (using $m' > 1/\gamma'$ and $\delta \gg \gamma'$)

$$w(\psi^{**}) - w(\psi^*) \leq |E(G) - E(G^*)|$$
$$< \gamma' \binom{m'}{2} \frac{n^2}{m'^2} + \binom{m'}{2} (\delta + \gamma') \frac{n^2}{m'^2} + m' \binom{n/m'}{2} < (\frac{\delta}{2} + 2\gamma')n^2 < 0.72\delta n^2.$$

Let R denote the m-vertex graph whose vertices are $\{1, \ldots, m\}$ and $(i,j) \in E(R)$ if and only if (V_i, V_j) is a γ-regular pair with density at least δ. We define a (labeled) fractional $\mathcal{F}$-packing ψ' of R as follows. Let H be a labeled copy of some $H_0 \in \mathcal{F}$ in R and assume that the vertices of H are $\{u_1, \ldots, u_k\}$ where u_i plays the role of vertex i in H_0. We define $\psi'(H)$ to be the sum of the values of ψ^* taken over all subgraphs of $G^*[V_{u_1}, \ldots, V_{u_k}]$ which are partite isomorphic to H_0, divided by n^2/m^2. Notice that by normalizing with n^2/m^2 we guarantee that ψ' is a proper fractional $\mathcal{F}$-packing of R and that $\nu^*_{\mathcal{F}}(R) \geq w(\psi') = m^2 w(\psi^*)/n^2 \geq m^2(\alpha - \delta)$.

We use ψ' to define a random coloring of the edges of G^*. Our "colors" are the labeled copies of elements of $\mathcal{F}$ in R. Let $d(i,j)$ denote the density of (V_i, V_j) and notice that $|E_{G^*}(V_i, V_j)| = d(i,j)n^2/m^2$. Let H be a labeled copy of some $H_0 \in \mathcal{F}$ in R, and assume that H contains the edge (i,j). Each $e \in E(V_i, V_j)$ is chosen to have the "color" H with probability $\psi'(H)/d(i,j)$. The choices made by distinct edges of G^* are independent. Notice that this random coloring is legal (in the sense that the sum of probabilities is at most one) since the sum of $\psi'(H)$ taken over all labeled copies of elements of $\mathcal{F}$ containing (i,j) is at most $d(i,j)$. Notice also that some edges might stay uncolored in our random coloring of the edges of G^*.

Let H be a labeled copy of some $H_0 \in \mathcal{F}$ in R, and assume that $\psi'(H) > m^{1-k_0}$. Without loss of generality, assume that the vertices of H are $\{1, \ldots, k\}$ where $i \in V(H)$ plays the role of $i \in V(H_0)$. Let r denote the number of edges of H. Notice that $r < k_0^2$. Let $W_H = G^*[V_1, \ldots, V_k]$ (in fact we only consider edges between pairs that correspond to edges of H_0). Notice that W_H is a subgraph of G^* which satisfies the conditions in Lemma 2.2, since $t = n/m > N\epsilon/(25{k_0}^2 M) > T_k$ (here we assume $N > 25{k_0}^2 M T_k/\epsilon$). Let W'_H be the spanning subgraph of W_H whose existence is guaranteed in Lemma 2.2. Let X_H denote the spanning subgraph of W'_H consisting only of the edges whose color is H. Notice that X_H is a random subgraph of W'_H. For an edge $e \in E(X_H)$, let $C_H(e)$ denote the set of subgraphs of X_H that contain e and that are partite isomorphic to H_0. Put $c_H(e) = |C_H(e)|$. A crucial argument is the following:

Lemma 3.1 *With probability at least* $1 - m^3/n$, *for all* $e \in E(X_H)$,

$$\left| c_H(e) - t^{k-2}\psi'(H)^{r-1} \right| < \mu \psi'(H)^{r-1} t^{k-2}. \qquad (1)$$

Proof: Let $C(e)$ denote the set of subgraphs of W'_H that contain e and that are partite isomorphic to H_0. Put $c(e) = |C(e)|$. According to Lemma 2.2, if $e \in E(V_i, V_j)$ then

$$\left| c(e) - t^{k-2} \frac{\prod_{(s,p)\in E(H_0)} d(s,p)}{d(i,j)} \right| < \zeta t^{k-2}. \tag{2}$$

Fix an edge $e \in E(X_H)$ belonging to $E(V_i, V_j)$. The probability that an element of $C(e)$ also belongs to $C_H(e)$ is precisely

$$\rho = \psi'(H)^{r-1} \cdot \frac{d(i,j)}{\prod_{(s,p)\in E(H_0)} d(s,p)}.$$

We say that two distinct elements $Y, Z \in C(e)$ are *dependent* if they share at least one edge other than e. Consider the dependency graph B whose vertex set is $C(e)$ and whose edges connect dependent pairs. Since two dependent elements share at least three vertices (including the two endpoints of e), we have that $\Delta(B) = O(t^{k-3})$. Hence, $\chi(B) = O(t^{k-3})$. Put $s = \chi(B)$. Let $C^1(e), \ldots, C^s(e)$ denote a partition of $C(e)$ into independent sets. Let $C^q_H(e) = C^q(e) \cap C_H(e)$, $c^q(e) = |C^q(e)|$ and $c^q_H(e) = |C^q_H(e)|$. Clearly, $c^1(e) + \cdots + c^s(e) = c(e)$ and $c^1_H(e) + \cdots + c^s_H(e) = c_H(e)$. The expectation of $c^q_H(e)$ is $\rho c^q(e)$. Consider some $C^q(e)$ with $c^q(e) > \sqrt{t}$. According to a large deviation inequality of Chernoff (cf. [2] Appendix A), for every $\eta > 0$, and in particular for $\eta = \mu/8$, if n (and hence t and hence $c^q(e)$) is sufficiently large,

$$\Pr[|c^q_H(e) - \rho c^q(e)| > \eta\rho c^q(e)] < e^{-\frac{(\eta\rho c^q(e))^2}{3c^q(e)}} = e^{-\eta^2\rho^2 c^q(e)/3} \ll t^{-k-1}.$$

Since $t = n/m$ and n is sufficiently large, it follows that with probability at least $1 - st^{-k-1} > 1 - t^{-3}$, for all $C^q(e)$ with $c^q(e) > \sqrt{t}$, $(1-\eta)\rho c^q(e) \le c^q_H(e) \le (1+\eta)\rho c^q(e)$ holds. Since the sum of $c^q(e)$ having $c^q(e) \le \sqrt{t}$ is $O(t^{k-2.5})$ and since $c(e) = \Theta(t^{k-2})$ we have that this sum is much less than $\rho\eta c(e)$. Thus, together with (2) and the fact that $\rho < \psi'(H)^{r-1}\delta^{-r}$ we have

$$c_H(e) = \sum_{q=1}^{s} c^q_H(e) \le \rho(1+\eta)(\sum_{q=1}^{s} c^q(e)) + \rho\eta c(e) = \rho(1+2\eta)c(e) \le \tag{3}$$

$$\rho(1+2\eta)t^{k-2}(\zeta + \frac{\prod_{(s,p)\in E(H_0)} d(s,p)}{d(i,j)}) = (1+2\eta)t^{k-2}(\psi'(H)^{r-1} + \zeta\rho) \le$$

$$t^{k-2}\psi'(H)^{r-1}(1+2\eta)(1+\zeta\delta^{-r}) \le t^{k-2}\psi'(H)^{r-1}(1+\mu/4)(1+\mu/2) \le (1+\mu)t^{k-2}\psi'(H)^{r-1}.$$

Similarly,

$$c_H(e) \ge \rho(1-\eta)c(e) - \rho\eta c(e) = \rho(1-2\eta)c(e) \ge \tag{4}$$

$$\rho(1-2\eta)t^{k-2}(\frac{\prod_{(s,p)\in E(H_0)} d(s,p)}{d(i,j)} - \zeta) = (1-2\eta)t^{k-2}(\psi'(H)^{r-1} - \zeta\rho) \ge$$

$$t^{k-2}\psi'(H)^{r-1}(1-2\eta)(1-\zeta\delta^{-r}) \ge t^{k-2}\psi'(H)^{r-1}(1-\mu/4)(1-\mu/2) \ge (1-\mu)t^{k-2}\psi'(H)^{r-1}.$$

Combining (3) and (4) we have that (1) holds for a fixed $e \in E(X_H)$ with probability at least $1 - t^{-3}$. As $|E(X_H)| < n^2$ we have that (1) holds for all $e \in E(X_H)$ with probability at least $1 - n^2/t^3 = 1 - m^3/n$. ■

We also need the following lemma that gives a lower bound for the number of edges of X_H.

Lemma 3.2 *With probability at least* $1 - 1/n$,

$$|E(X_H)| > (1-2\zeta)r\frac{n^2}{m^2}\psi'(H).$$

Proof: We use the notations from Lemma 3.1 and the paragraph preceding it. For $(i,j) \in E(H_0)$, the expected number of edges of $E(V_i, V_j)$ that received the color H is precisely $d(i,j)\frac{n^2}{m^2}\frac{\psi'(H)}{d(i,j)} = \frac{n^2}{m^2}\psi'(H)$. Summing over all r edges of H_0, the expected number of edges of W_H that received the color H is precisely $r\frac{n^2}{m^2}\psi'(H)$. As at most $\zeta|E(W_H)|$ edges belong to W_H and do not belong to W'_H we have that the expectation of $|E(X_H)|$ is at least $(1-\zeta)r\frac{n^2}{m^2}\psi'(H)$. As ζ, r, m are constants and as $\psi'(H)$ is bounded from below by the constant m^{1-k_0}, we have, by the common large deviation inequality of Chernoff (cf. [2] Appendix A), that for $n > N$ sufficiently large, the probability that $|E(X_H)|$ deviates from its mean by more than $\zeta r\frac{n^2}{m^2}\psi'(H)$ is exponentially small in n. In particular, the lemma follows. ■

Since R contains at most $O(m^{k_0})$ labeled copies of elements of $\mathcal{F}$ with at most k_0 vertices, we have that with probability at least $1 - O(m^{k_0}/n) - O(m^{k_0+3}/n) > 0$ (here we assume again that N is sufficiently large) *all* labeled copies H of elements of $\mathcal{F}$ in R with $\psi'(H) > m^{1-k_0}$ satisfy the statements of Lemma 3.1 and Lemma 3.2. We therefore fix a coloring for which Lemma 3.1 and Lemma 3.2 hold for all labeled copies H of elements of $\mathcal{F}$ in R having $\psi'(H) > m^{1-k_0}$.

Let H be a labeled copy of some $H_0 \in \mathcal{F}$ in R with $\psi'(H) > m^{1-k_0}$, and let r denote the number of edges of H. We construct an r-uniform hypergraph L_H as follows. The vertices of L_H are the edges of the corresponding X_H from Lemma 3.1. The edges of L_H correspond to the edge sets of the subgraphs of X_H that are partite isomorphic to H_0. We claim that our hypergraph satisfies the conditions of Lemma 2.3. Indeed, let q denote the number of vertices of L_H. Notice that Lemma 3.2 provides a lower bound for q. Let $d = t^{k-2}\psi'(H)^{r-1}$. Notice that by Lemma 3.1 *all* vertices of L_H have their degrees between $(1-\mu)d$ and $(1+\mu)d$. Also notice that the co-degree of any two vertices of L_H is at most t^{k-3} as two edges cannot belong, together, to more than t^{k-3} subgraphs of X_H that are partite isomorphic to H_0. In particular, for N sufficiently large, $\mu d > t^{k-3}$. By Lemma 2.3 we have at least $(q/r)(1-\beta)$ edge-disjoint copies of H_0 in X_H. In particular, we have at least

$$(1-\beta)(1-2\zeta)\frac{n^2}{m^2}\psi'(H) > (1-2\beta)\psi'(H)\frac{n^2}{m^2}$$

such copies. Recall that $w(\psi') \geq m^2(\alpha - \delta)$. Since there are at most $O(m^{k_0})$ labeled copies H of elements of $\mathcal{F}$ in R with $0 < \psi'(H) \leq m^{1-k_0}$, their total contribution to $w(\psi')$ is at most $O(m)$.

Hence, summing the last inequality over all H with $\psi'(H) > m^{1-k_0}$ we have at least

$$(1-2\beta)m^2(\alpha-\delta-O(\frac{1}{m}))\frac{n^2}{m^2} > n^2(\alpha-\epsilon)$$

edge disjoint copies of elements of $\mathcal{F}$ in G. It follows that $\nu_{\mathcal{F}}(G) \geq n^2(\alpha-\epsilon)$. As $\nu^*_{\mathcal{F}}(G) = \alpha n^2$, Theorem 1.2 follows. ■

The proof of Theorem 1.2 implies an $O(n^{poly(k_0)})$ time algorithm that produces a set of $n^2(\alpha-\epsilon)$ edge-disjoint copies of elements of $\mathcal{F}$ in G with probability at least, say, 0.99. Indeed, Lemma 2.1 can be implemented in $o(n^3)$ time using the algorithm of Alon et al. [1]. Lemma 2.3 has a polynomial running time implementation due to Grable [8]. Since we only need to compute ψ^{**}, rather than ψ, we can do this in $O(n^{poly(k_0)})$ time using any polynomial time algorithm for LP. The other ingredients of the proof are easily implemented in polynomial time.

4 The gap between integral and fractional packings

Theorem 1.2 shows that the integer and fractional packing differ by at most $o(n^2)$. Thus, it is interesting to determine how large the gap between them can be. Proposition 1.3 shows that, in general, the gap is essentially this large.

Proof of Proposition 1.3: Let $\epsilon > 0$ be given. Let $k = k(\epsilon)$ be a constant to be chosen sufficiently large as a function of ϵ. It is well known (see, e.g., [10]) that for $N = N(k)$ sufficiently large, K_n contains more than, say, $n^2/(2k^2)$ edge-disjoint copies of K_{k+1}. Hence, let G^* be an n-vertex graph with $r(k+1)k/2$ edges where $r > n^2/(2k^2)$ and which consists of a set R^* of r edge-disjoint copies of K_{k+1}.

Consider a copy H of K_k in G^*. We say that H is *good* if it is contained entirely in some K_{k+1} element of R^*. Other copies of K_k are called *bad*. A bad copy H corresponds to a decomposition of the edges of K_k into smaller cliques, corresponding to the intersection of H with elements of R^*. Let Z_k denote the family of all nontrivial clique decompositions of K_k (hence, e.g., $Z_4 = \{A, B\}$ where A is the decomposition of K_4 into 6 edges and B is the decomposition of K_4 into a triangle and three edges). Thus, each bad copy H has some *type* $t(H) \in Z_k$. For $X \in Z_k$ let $c(X)$ denote the number of elements of X (in the example above, $c(A) = 6$ and $c(B) = 4$). It is well known that if $X \in Z_k$ then $k(k-1)/2 \geq c(X) \geq k$. For $X \in Z_k$ let $s(X)$ denote the minimum number of elements of X incident with a vertex (in the example above, $s(A) = 3$ and $s(B) = 2$). Clearly, $2 \leq s(X) \leq k-1$.

We construct a random spanning subgraph G of G^* by independently choosing each element of R^* with probability $n^{-\epsilon/2}$. Let $R \subset R^*$ denote the random subset chosen. Clearly, $E[|R|] = rn^{-\epsilon/2}$. The maximum number of **edge-disjoint** good copies of K_k in G is precisely $|R|$. For any $X \in Z_k$, let us estimate the number of bad copies H with $t(H) = X$ that survived in G. Clearly, the probability that H appears in G is precisely $n^{-(\epsilon/2)c(X)}$. How many such H appear in G^*? Trivially, every vertex v of G^* appears in $O(n)$ elements of R^*. Hence, v can play the role of the vertex of X yielding $s(X)$ at most $O(n^{s(X)})$ times. Summing this for each vertex of G^* we get that the number of H with $t(H) = X$ is $O(n^{s(X)+1})$. We consider two cases. If $s(X) < \epsilon k/4$ the expected number of surviving bad copies H with $t(H) = X$ is therefore at most

$$O(n^{-(\epsilon/2)c(X)}n^{s(X)+1}) \le O(n^{-(\epsilon/2)k}n^{(\epsilon/4)k+1}) = o(1).$$

If $s(X) \ge \epsilon k/4$ then consider the number of elements of X with at most $8/\epsilon$ vertices. By the definition of $s(X)$ each vertex in the X-decomposition of K_k appears in a least $s(X) - (k-1)\epsilon/8 \ge \epsilon k/8$ elements with at most $8/\epsilon$ vertices. It follows that $c(X) \ge k(\epsilon k/8)/(8/\epsilon) = k^2\epsilon^2/64$. Thus, the expected number of surviving bad copies H with $t(H) = X$ is at most

$$O(n^{-(\epsilon/2)c(X)}n^{s(X)+1}) \le O(n^{-(\epsilon/2)(k^2\epsilon^2/64)}n^k) = o(1).$$

Since $|Z_k|$ is a constant depending only on k we have shown that the expected number of bad copies in G is $o(1)$. Since $|R|$ is highly concentrated around its mean $rn^{-\epsilon/2}$ we have that with positive (in fact, high) probability, G has no bad copies and at least $0.5rn^{-\epsilon/2}$ good copies. Now, $\nu_{K_k}(G) = |R|$ and since $\nu^*_{K_k}(K_{k+1}) = (k+1)/(k-1)$ we have $\nu^*_{K_k}(G) = |R|(k+1)/(k-1)$. Thus, $\nu^*_{K_k}(G) - \nu_{K_k}(G) = 2|R|/(k-1) > n^{2-\epsilon}$. ■

It is interesting to determine the largest possible difference between integer and fractional packings of specific graphs. For K_3 an argument similar to the proof of Proposition 1.3 shows that there are n-vertex graphs G for which $\nu^*_{K_3}(G) - \nu_{K_3}(G) = \Theta(n^{1.5})$. Thus, it is interesting to determine whether $\nu^*_{K_3}(G) - \nu_{K_3}(G) = O(n^{1.5})$ holds for all graphs G. We note that the method of proof of Theorem 1.2 would probably not be adequate in order to prove such a result since the regularity lemma is not sensitive enough to establish sub-quadratic error terms.

Acknowledgments

The author thanks Jeff Kahn for his help in constructing the lower bound given in Proposition 1.3.

References

[1] N. Alon, R.A. Duke, H. Lefmann, V. Rödl and R. Yuster, *The algorithmic aspects of the Regularity Lemma*, Journal of Algorithms 16 (1994), 80–109.

[2] N. Alon and J. H. Spencer, **The Probabilistic Method**, *Second Edition*, Wiley, New York, 2000.

[3] B. Bollobás, *Extremal Graph Theory*, Academic Press, 1978.

[4] D. Dor and M. Tarsi, *Graph decomposition is NPC - A complete proof of Holyer's conjecture*, Proc. 20th ACM STOC, ACM Press (1992), 252–263.

[5] P. Frankl and V. Rödl, *Near perfect coverings in graphs and hypergraphs*, European J. Combinatorics 6 (1985), 317–326.

[6] Z. Füredi, *Matchings and covers in hypergraphs*, Graphs and Combinatorics 4 (1988), 115–206.

[7] R. Duke, H. Lefmann and V. Rödl, *A fast algorithm for computing frequencies in a given graph*, SIAM Journal of Computing 24 (1995), 598–620.

[8] D. Grable, *Nearly-perfect hypergraph packing is in NC*, Information Processing Letters 60 (1996), 295–299.

[9] P. E. Haxell and V. Rödl, *Integer and fractional packings in dense graphs*, Combinatorica 21 (2001), 13–38.

[10] V. Rödl, *On a packing and covering problem*, Europ. J. of Combin. 6 (1985), 69–78.

[11] E. Szemerédi, Regular partitions of graphs, in: *Proc. Colloque Inter. CNRS* 260, CNRS, Paris, 1978, 399–401.