DIMACS Technical Report 2001-36
October 2001

On a Tiling Conjecture of Komlós for 3-Chromatic Graphs

by

Ali Shokoufandeh
Dept. of Mathematics and
Computer Science
Drexel University
Philadelphia, PA 19104

Yi Zhao
Dept. of Mathematics
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Research Institute and Telcordia Technologies (formerly Bellcore).

DIMACS was founded as an NSF Science and Technology Center, and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

Given two graphs G and H, an H-*matching* of G (or a *tiling* of G with H) is a subgraph of G consisting of vertex-disjoint copies of H. For an r-chromatic graph H on h vertices, we write $u = u(H)$ for the smallest possible color-class size in any r-coloring of H. The *critical chromatic number* of H is the number $\chi_{cr}(H) = (r-1)h/(h-u)$. A conjecture of Komlós states that for every graph H, there is a constant K such that if G is any n-vertex graph of minimum degree at least $(1 - (1/\chi_{cr}(H)))\, n$, then G contains an H-matching that covers all but K vertices of G. In this paper we prove that the conjecture holds for all sufficiently large values of n, when H is a 3-chromatic graph.

1 Introduction

All graphs considered in this paper are finite, undirected and simple. If H is a graph on h vertices and host graph G is a graph on n vertices, in *tiling* problems the objective is to find many vertex disjoint copies of H in G, or even a complete tiling (or H-factor) of G with $\lfloor n/h \rfloor$ copies of H. Perhaps, one of the earliest tiling results in extremal graph theory is Dirac's theorem on Hamilton paths [5] that solves the 1-factor problem. The case of triangle-factors is due to Corrádi-Hajnal [4], and finally the celebrated result of Hajnal-Szemerédi settles the K_r-factor problem for all r:

Theorem 1 (Hajnal-Szemerédi 1970 [7]). *Let G be an n-graph with minimum degree*

$$\delta(G) \geq \left(1 - \frac{1}{r}\right) n$$

then G has a K_r-factor.

During the 90's Alon and Yuster extended the Hajnal-Szemerédi theorem to arbitrary H in various ways:

Theorem 2 (Alon-Yuster 1992 [2]). *For every $\varepsilon > 0$ and for every integer h there exists an $n_0 = n_0(\varepsilon, h)$ such that for every graph H on h vertices with chromatic number $\chi(H)$, and for any graph G with $n > n_0$ vertices and with minimum degree*

$$\delta(G) \geq \left(1 - \frac{1}{\chi(H)}\right) n, \tag{1}$$

contains at least $(1-\varepsilon)n/h$ vertex disjoint copies of H.

Theorem 3 (Alon-Yuster 1996 [3]). *For every $\varepsilon > 0$ and for every integer h there exists an $n_0 = n_0(\varepsilon, h)$ such that for every graph H on h vertices and for every $n > n_0$, any graph G with hn vertices and with minimum degree*

$$\delta(G) \geq \left(1 - \frac{1}{\chi(H)} + \varepsilon\right) n, \tag{2}$$

has an H-factor.

They conjectured that two error terms εn in Theorem 2 could be released to a constant and a zero. In [2] they also remarked that this is essentially best possible. These conjectures have recently been proven by Komlós, Sárközy and Szemerédi :

Theorem 4 (Komlós, Sárközy and Szemerédi 1995 [12]). *For every graph H there is a constant K such that, if G is an n-graph satisfying*

$$\delta(G_n) \geq \left(1 - \frac{1}{\chi(H)}\right) n, \tag{3}$$

then it has an H-matching that covers all but at most K vertices.

Theorem 5 (Komlós, Sárközy and Szemerédi 1995 [12]). *Given the condition of Theorm 4, if*

$$\delta(G) \geq \left(1 - \frac{1}{\chi(H)}\right) n + K, \tag{4}$$

then G has an H-factor.

Let us use the notation $TT(n,H) = \min\{t : \delta(G) \geq t \text{ implies that } n\text{-graph } G \text{ has an } H\text{-factor}\}$ and define $TT(n,H,M)$ to be the minimum number m such that, if G is an n-graph with minimum degree $\delta(G) \geq m$, then there is an H-matching covering at least M vertices in G. Then, the asymptotic behavior of $TT(n,H)$ seems to be completely understood: the limit of $TT(n,H)/n$ seems to be $1 - 1/\chi(H)$. This suggests that the chromatic number of H probably is the relevant quantity for tiling problems. Of course, it is still possible to fine-tune the known estimates for $TT(n,H)$, e.g. by characterizing those graphs H for which the error term K in Theorem 5 is actually 0, that is, $TT(n,H) \leq (1 - 1/\chi(H))n$. Dirac's theorem and the Hajnal-Szemerédi theorem show that paths and cliques are like that. In fact, for cliques the inequality is an equality. But a full characterization of these graphs is probably hard.
However, in [8] Komlós presented an improved form of Theorem 4 (but not of Theorem 5) in terms of *critical chromatic number* of H. For an r-chromatic graph H of order h, we write $u = u(H)$ for the smallest possible color-class size in any r-coloring of H. The critical chromatic number of H is the number $\chi_{cr}(H) = (r-1)h/(h-u)$. It is easy to see that $\chi(H) - 1 < \chi_{cr}(H) \leq \chi(H)$, and $\chi_{cr}(H) = \chi(H) = r$ if and only if every r-coloring of H has equal color-class sizes.

Theorem 6 (Komlós 2000 [8]). *Let H have parameters $\chi = \chi(H)$ and $\chi_{cr} = \chi_{cr}(H)$. Then, for all $0 < M \leq n$,*

$$TT(n,H,M) \geq M\left(1 - \frac{1}{\chi_{cr}}\right) + (n-M)\left(1 - \frac{1}{\chi - 1}\right). \tag{5}$$

in particular, $TT(n,H) \geq \left(1 - \frac{1}{\chi_{cr}}\right) n$.

He also proved the following result for existence of an H-matching in an n-graph G, with minimum degree of G depending only on $\chi_{cr}(H)$

Theorem 7 (Komlós [8]). *For every graph H and $\varepsilon > 0$ there is a threshold $n_0 = n_0(H,\varepsilon)$ such that, if $n \geq n_0$ and an n-graph G satisfies the degree condition*

$$\delta(G) \geq \left(1 - \frac{1}{\chi_{cr}(H)}\right) n, \tag{6}$$

then G contains an H-matching that covers all but at most εn vertices.

He also posed the following conjecture:

Conjecture 1 (Komlós [8]). *For every graph H there is a constant $K = K(H)$ such that, if G is an n-graph satisfying* (6)*, then G contains an H-matching that covers all but at most K vertices.*
This is best possible for every H. Hence,

$$TT(n, H, n-K) = \left(1 - \frac{1}{\chi_{cr}(H)}\right) n.$$

In this paper we will show that the conjecture holds for all sufficiently large values of n, when H is a 3-chromatic graph.

Theorem 8. *For any* 3*-chromatic graph H of order h with $u = u(H)$, there exists an n_0 such that for all $n \geq n_0$, if G is any n vertex graph with*

$$\delta(G) \geq \left(1 - \frac{1}{\chi_{cr}(H)}\right) n, \tag{7}$$

G contains an H-matching that covers all but at most $\frac{6(h-u)h}{u}$ vertices of G.

In the proof we will use the concept of *bottle-graphs*. A bottle-graph of chromatic number r is a complete r-partite graph with color-class sizes $(u, w, w, \ldots, w)$, where $u = \alpha w$, for some $\alpha \leq 1$. Clearly, the critical chromatic number of this graph is $r - 1 + \alpha$. The vector $(\frac{\alpha}{r-1+\alpha}, \frac{1}{r-1+\alpha}, \ldots, \frac{1}{r-1+\alpha})$ is called the *color-vector* of the bottle-graph. The parameters u and w are the *neck* and the *width* of the bottle-graph, respectively.
Given an r-chromatic graph H of order h with $u = u(H)$, we say that a graph $\mathcal{B} = \mathcal{B}(H)$ is the bottle-graph of H if $\mathcal{B}$ is the smallest bottle-graph that contains an H-factor, with the color-vector $\underline{\beta} = (s, t, \ldots, t)$, where $s = u/h$, and $t = (1-s)/(r-1)$. Observet that , $\chi_{cr}(\mathcal{B}) = \chi_{cr}(H) = (r-1)/(1-s)$. For example, let $u, u_1, u_2, \ldots, u_{r-1}$ denote the color-class sizes in an r-coloring of H, we can always construct a bottle-graph using $r-1$ vertex disjoint copies of H, with i-th copy of H having $u, u_i, u_{i+1}, \ldots, u_{r-1}, u_1, \ldots, u_{i-1}$ vertices in color class $1, 2, \ldots, r$ of bottle-graph. Consequently, the order of $\mathcal{B}(H)$ is at most $(r-1)h$. Therefore, it is sufficient to prove Theorem 8 bottle-graphs:

Theorem 9 (Main Theorem). *For every bottle-graph H with chromatic number* 3*, neck u, and width w, there exists an n_0 such that for all $n \geq n_0$ if G is any n-graph with*

$$\delta(G) \geq \frac{\alpha+1}{\alpha+2} n = \left(\frac{1}{2} + \gamma\right) n, \tag{8}$$

where $\alpha = \frac{u}{w}$ and $\gamma = \frac{\alpha}{2(\alpha+2)}$, then G contains an H-matching that covers all but at most $\frac{3}{\gamma} w$ vertices.

2 Notations

$V(G)$ and $E(G)$ denote the vertex-set and the edge-set of the graph G, and we write $v(G) = |V(G)|$ (order of G) and $e(G) = |E(G)|$ (size of G). G_n denotes n-graphs. $N(v, X)$ is the set of neighbors of $v \in V$ in the set $X \subset V$. We use $N(v)$ to denote $N(v,V)$. Hence $|N(v,X)| = \deg(v,X) = \deg_G(v,X)$ is the degree of v in set X, and $\deg(v) = \deg(v,V)$. In a directed graph D, we use $N_{out}(v)$ to denote $\{u \in V(D)|\ (v,u) \in E(D)\}$ (the *out* neighborhood of v), and $\deg_{out}(v) = |N_{out}(v)|$. $\delta(G)$ stands for the minimum, and $\Delta(G)$ for the maximum degree in G. $\nu_i(G)$ denotes the size of a maximum set of vertex disjoint i-stars (stars with i leaves) in G. In particular, $\nu_1(G) = \nu(G)$ is the size of a maximum matching in G. We write $\chi(G)$ and $\chi_{cr}(G)$ for the chromatic number andcritical chromatic number of G, respectively. For an r-chromatic graph H of order h, we write $u = u(H)$ for the smallest possible color-class size in any r-coloring of H. When A and B are disjoint subsets of $V(G)$, we use $\deg(A,B)$ to denote the number of edges in $E(G)$ with one endpoint in A and the other in B. A bipartite graph G with color classes A and B and edge set E will be denoted by $G = (A,B,E)$, with $E \subseteq A \times B$. $K(n_1,n_2,\dots,n_r)$ is the complete r-partite graph with color classes of sizes $n_1,n_2,\dots,n_r$. The *density* between disjoint sets X and Y is defined as:

$$d(X,Y) = \frac{\deg(X,Y)}{|X||Y|}.$$

In the proof of Main Theorem Szemerédi's *Regularity Lemma* [13] plays a pivotal role. We will need the following definition to state the regularity lemma:

Definition 1 (Regularity Condition). *Let $\varepsilon > 0$. A pair (A,B) of disjoint vertex-sets in G is ε-**regular** if for every $X \subset A$ and $Y \subset B$, satisfying*

$$|X| > \varepsilon|A|,\ |Y| > \varepsilon|B|,$$

we have

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

This definition implies that regular pairs are highly uniform bipartite graphs; namely the density of any reasonably large subgraph is almost the same as the density of a regular pair. We will use the following form of the Regularity Lemma:

Lemma 10 (Degree Form). *For every $\varepsilon > 0$ there is an $M = M(\varepsilon)$ such that if $G = (V,E)$ is any graph and $d \in [0,1]$ is any real number, then there is a partition of the vertex set V into $\ell+1$ clusters $V_0, V_1, \dots, V_\ell$, and there is a subgraph G' of G with the following properties:*

- $\ell \le M$,
- $|V_0| \le \varepsilon|V|$,
- *all clusters V_i, $i \ge 1$, are of the same size $L \le \varepsilon\lceil|V|\rceil$,*
- $\deg_{G'}(v) > \deg_G(v) - (d+\varepsilon)|V|$ *for all $v \in V$,*
- $G'|_{V_i} = \emptyset$ *(V_i is an independent set in G'), for all i,*

- *all pairs* (V_i, V_j), $1 \le i < j \le \ell$, *are* ε*-regular, each with density either 0 or greater than* d, *in* G'.

A stronger one-sided property of regular pairs is the super-regularity:

Definition 2 (Super-Regularity condition). *Given a graph* G *and two disjoint subsets of its vertices* A *and* B, *the pair* (A, B) *is* (ε, d)-super-regular, *if it is* ε*-regular and,*

$$\deg(a) > d|B| \ \forall a \in A, \ \deg(b) > d|A| \ \forall b \in B.$$

We also use the Blow-up Lemma (see [11]):

Lemma 11. *Given a graph* R *of order* r *and positive parameters* δ *and* Δ, *there exists an* $\varepsilon = \varepsilon(\delta, \Delta, r) > 0$ *such that the following holds. Let* $n_1, n_2, \ldots, n_r$ *be arbitrary positive integers, and let us replace the vertices* $v_1, v_2, \ldots, v_r$ *of* R *with pairwise disjoint sets* $V_1, V_2, \ldots, V_r$ *of sizes* $n_1, n_2, \ldots, n_r$ *(blowing up. we construct two graphs on the same vertex-set* $V = \cup_i V_i$. *The first graph* R_b *is obtained by replacing each edge* $\{v_i, v_j\}$ *of* R *with the complete bipartite graph between the corresponding vertex-sets* V_i *and* V_j. *A sparser* G *is constructed by replacing each edge* $\{v_i, v_j\}$ *of* R *arbitrarily with some* (ε, δ)*-super-regular pair between* V_i *and* V_j. *If a graph* H *with* $\Delta(H) \le \Delta$ *is embeddable into* R_b *then it is already embeddable into* G.

3 Number of Left-over Vertices

In Theorem 9, K is a constant depended on the test graph H. El-Zahar conjecture [6] (recently proved by Sarmad Abbasi [1]) claimed the $K(C_l) = 0$ where C_l denotes a cycle in l vertices. In genereal, K may not be zero. The following example justifies the tightness on the number of leftover vertices in Theorem 9.

Let G be an n-graph and suppose the vertex set $V(G)$ can be partitioned to $A \cup B$, with $|A| = \left(\frac{1}{2} - \gamma\right) n$ and $|B| = \left(\frac{1}{2} + \gamma\right) n$. We will assume G is complete between A and B and A is an independent set. We will assume B contains $s = \left\lfloor \frac{|B|}{4\gamma n + 2w - 2} \right\rfloor$ connected components, each of which is a complete bipartite graph (L_i, R_i). We will assume that for $1 \le i \le s - 1$ the color classes of the component (L_i, R_i) satisfy $|L_i| = |R_i| = 2\gamma n + w - 1$. Without loss of generality we will assume $w + u$ divides $2\gamma n$. It is easy to see that G satisfies the degree condition in (8).

Let $\mathcal{H}$ denote an H-matching in G. In every copy of H in $\mathcal{H}$, two of its color classes have to come from an (L, R) component of B, and the third color class should lay in A. Let us assume every copy of H in $\mathcal{H}$ has one of its w-classes in A, i.e. restriction of $\mathcal{H}$ on B is a union of $K(w, u)$ graphs, then there are at least $2w - 2$ vertices left uncovered in each component (L_i, R_i) of B, for $i = 1, \ldots, s - 1$. In fact, let $t = \frac{2\gamma n}{w+u}$, then we can find $2t$ copies of $K(w, u)$ in each (L, R) component by placing t copies of w- and u-sets in each side (thus $2w - 2$ vertices will be left uncovered). Assume to the contrary that $2t + 1$ copies of $K(w, u)$

could be placed in an (L, R) component of B, i.e. either L or R contains $t + i$ w-sets and $t - i + 1$ u-sets, $(i \geq 1)$. But this is impossible, because

$$(t+i)w + (t-i+1)u = tw + tu + i(w-u) + u \geq t(w+u) + w > 2\gamma n + w - 1.$$

Finally, our assumption that all copies of H have one of their w-sets in A is necessary in achieving the maximum number of copies of H in $\mathcal{H}$. In fact, embedding $K(w, w)$ in the (L, R) of B will result in fewer than $2t$ copies of H in (L, R, A). Therefore, in any H-matching of G, there will be at least $2w - 2$ uncovered vertices for each component $(L_i, R_i) \in B$, $1 \leq i \leq s - 1$, leaving a total of at least $(s-1)(2w-2)$ vertices uncovered in B. Consequently, there will be at least $(s-1)(2w-2)\frac{|A|}{|B|}$ vertices left uncovered in A. Together, the number of uncovered vertices in any H-matching of G is at least

$$(s-1)(2w-2)\frac{n}{|B|} = \left\lceil \frac{|B|}{4\gamma n + 2w - 2} \right\rceil \frac{n}{|B|}(2w-2) \approx \frac{w}{2\gamma}.$$

4 Proof of Main Theorem

Our test graph H is a 3-chormatic bottle-graph with neck u and width w. We will assume $\chi_{cr} < \chi$ (that is, $u < w$ or $\alpha < 1$), since otherwise the result follows from Theorem 4. Throughout the paper, we will assume that n is sufficiently large and will use the following main parameters:

$$\varepsilon \ll d \ll \sigma \ll \mu \ll 1 - \alpha. \tag{9}$$

For simplicity we do not compute the actual dependencies among these parameters.
We first apply Lemma 10 to G, with ε and d as in (9), and will get a partition of $V(G)$ into clusters $V_0, V_1, \ldots, V_\ell$. Without loss of generality we assume L, the size of V_i, $1 \leq i \leq \ell$, is divisible by $2u(w-u)(w+u)(2w+u)$, otherwise we can move constant number of vertices from each V_i to V_0 to achieve this condition. We let $G'' \subset G'$ to be the induced graph on $V(G') \setminus V_0$ and define the *reduced graph* R as follows:
The vertices of R are the clusters V_i, $1 \leq i \leq \ell$, and there exists an edge between two vertices of R if the corresponding clusters form an ε-regular pair in G' with density exceeding d. Since $\delta(G') \geq (\frac{1}{2} + \gamma - (d + \varepsilon))n$, and $\varepsilon \ll d$, an easy computation shows that:

$$\delta(R) \geq \left(\frac{1}{2} + \gamma - 2d\right)\ell. \tag{10}$$

4.1 Outline of the proof

Throughout the proof we will use two classes of tripartite graphs. Recall that L is the size of clusters. The *unbalanced* triangle-graph $H^*(t)$ is a tripartite graph, with three color-classes of size $t, t, \alpha t$, where $t = cL$, for some constant $0 < c \leq 1$. Moreover, every two color-classes of $H^*(t)$ form an ε-regular pair. The *balanced* triangle-graph $H^{**}(t)$ is defined similar to $H^*(t)$, expect all its color classes are of size t.

The outline of proof is as follows:
We will show that, except for a special case, the vertices of G'' can be tiled by $H^*(L_1)$ and $H^{**}(L_2)$ in a way that positive percentage of vertices of G'' are in copies of $H^{**}(t_2)$. Here $L_2 = (2+\alpha)L_1 = cL$ for some constant $0 < c \leq 1$. Next, the vertices in V_0 will be inserted into appropriate (cluster) triangles (H^*'s or H^{**}') in a way that after we remove all copies of H containing new vertices, each of remaing triangles still has a H-tiling leaving at most constant ($4w$) vertices. The connection among clusters will finally reduce the number of left-over vertices to a constant depended only on w and u.

In the special case, we will show that there exists an almost-independent vertex set A in G of size $\frac{n}{2+\alpha}$. Furthermore, $B = V(G) \setminus A$ can be either partitioned into B_1 and B_2 such that $|B_1| = \frac{n}{2+\alpha}, |B_2| = \frac{\alpha n}{2+\alpha}$, and $d(B_1, B_2)$ is almost one, or tiled by $K(u, w)$ except for constant many vertices. In either case, we can conclude that except for a constant number of vertices, G can be tiled by H.

4.2 The tiling of G'' with cluster triangles

To obtain a tiling of G'' with H^* and H^{**}, we need the following concept from pure graph theory.

4.2.1 The maximal clique cover of a graph

Given a graph $\mathcal{G}$, a k-clique cover $\Phi = \{\Phi_k, \Phi_{k-1}, \ldots, \Phi_1\}$ is a collection of disjoint cliques, where Φ_i corresponds to a set of cliques of order i, for $i \in \{k, k-1, \ldots, 1\}$, and $V(\mathcal{G}) = \cup_{i=1}^{k} V(\Phi_i)$. A k-clique cover $\Phi = \{\Phi_k, \Phi_{k-1}, \ldots, \Phi_1\}$ is siad to be *maximal* if for any other k-clique cover $\Phi' = \left\{\Phi'_k, \Phi'_{k-1}, \ldots, \Phi'_1\right\}$, if for some $1 \leq i \leq k$, $|\Phi'_i| > |\Phi_i|$ then there is a $i < j \leq k$ such that $|\Phi_j| > |\Phi'_j|$.

Fix a maximal 3-clique cover $\Phi = \{\Phi_3, \Phi_2, \Phi_1\}$ in the graph $\mathcal{G}$. Let K and K' be two cliques in Φ of sizes i and j, with $i \leq j$. We say K and K' are:

- *well-connected* (or $K \hookrightarrow K'$), if $\deg(v, K') = j - 1$, $\forall v \in K$,
- *over-connected* (or $K \overset{>}{\hookrightarrow} K'$), if $\deg(K, K') \geq i(j-1)$ and $K \not\hookrightarrow K'$,
- *under-connected* (or $K \overset{<}{\hookrightarrow} K'$), if $\deg(K, K') < i(j-1)$.

The following propositions are imediate consequence of maximality of Φ:

Proposition 1.

1. $\deg(c, e) \leq 1$, $\deg(e, e') \leq 2$, $\deg(\{c, c'\}, \tau) \leq 4$ and $\deg(\{e, e'\}, \tau) \leq 8$, for any c, c' in Φ_1, e, e' in Φ_2, and τ in Φ_3.
2. Fix an edge $e \in \Phi_2$, and label its end vertices by $Top(e)$ and $Bot(e)$. If another $e' \in \Phi_2$ satisfies $e \hookrightarrow e'$, then the vertices of e' can be labeled as $Top(e')$ and $Bot(e')$ such that

$\{Top(e), Top(e')\} \times \{Bot(e), Bot(e')\}$ form a complete bipartite graph. Moreover, if $\mathcal{E} = \{e' \in \Phi_2 : e \hookrightarrow e'\}$, then $\{Top(e') : e' \in \mathcal{E}\}$ and $\{Bot(e') : e' \in \mathcal{E}\}$ form two independent sets.

3. If $\deg(\{c_1, c_2\}, \tau) = 4$, for $c_1, c_2 \in \Phi_1$ and $\tau \in \Phi_3$, then $c_1 \hookrightarrow \tau$ and $c_2 \hookrightarrow \tau$, and $N(c_1, \tau) = N(c_2, \tau)$. Let $Top_c(\tau)$ denotes the vertex of τ that is not contained in $N(c, \tau)$ when $c \hookrightarrow \tau$, then $Top_{c_1}(\tau) = Top_{c_2}(\tau)$ (we can simply use $Top(\tau)$ to denote this vertex, as it is independent of the choice of vertex $c \in \Phi_1$). It is obvious that $Top(\tau)$ plays the same role as the elements of Φ_1 and τ can be replaced by $c_1 \cup N(c_2, \tau)$. Moreover, if $\mathcal{S}_1 = \{\tau \in \Phi_3 : c_1 \hookrightarrow \tau, c_2 \hookrightarrow \tau\}$, then $\{Top(\tau) : \tau \in \mathcal{S}_1\}$ is an independent set.

4. If $\deg(e_1, \tau) = \deg(e_2, \tau) = 4$, for $e_1, e_2 \in \Phi_2$ and $\tau \in \Phi_3$, then $e_1 \hookrightarrow \tau$ and $e_2 \hookrightarrow \tau$, and the same vertex of τ is adjacent to both ends of e_1 and e_2. We will use $Tip(\tau)$ to denote this vertex and it is independent of the choice of edges e_1 and e_2. Further, we can label the other two vertices of τ as $Top(\tau)$ and $Bot(\tau)$, and the end points of e_i, $i = 1, 2$ as $Top(e_i)$ and $Bot(e_i)$ such that $\{Top(e_1), Top(e_2), Top(\tau)\}$ and $\{Bot(e_1), Bot(e_2), Bot(\tau)\}$ are independent sets. Hence, the pair $\{Top(\tau), Bot(\tau)\}$ plays the same role as e_1 or e_2. Finally, let $\mathcal{S}_2 = \{\tau \in \Phi_3 : e_1 \hookrightarrow \tau, e_2 \hookrightarrow \tau\}$, then $\{Top(\tau) : \tau \in \mathcal{S}_2\}$ and $\{Bot(\tau) : \tau \in \mathcal{S}_2\}$ form two independent sets in R.

5. A $\tau \in \Phi_3$ can be over-connected to at most one element in Φ_1 or Φ_2. Moreover, if τ is over-connected to one element, it will be under-connected to any other element of Φ_1 and Φ_2.

As a result of above properties we will have:

Lemma 12. *Suppose $\Phi = \{\Phi_1, \Phi_2, \Phi_3\}$ is a maximal 3-clique cover of a graph $\mathcal{G}$, with*

$$\max(|\Phi_2|, |\Phi_1|) \geq 2.$$

Then

$$|\Phi_3| \geq |\Phi_1| + 2\delta(\mathcal{G}) - v(\mathcal{G}).$$

Proof. First assume that $|\Phi_1| \geq 2$. Since Φ is maximal clique cover, for $c, c' \in \Phi_1$ we have:

$$\deg(\{c, c'\}, \Phi_1) = 0,$$

$$\deg(\{c, c'\}, \Phi_2) \leq 2|\Phi_2|,$$

$$\deg(\{c, c'\}, \Phi_3) \leq 4|\Phi_3|.$$

Together we have

$$2\delta(\mathcal{G}) \leq \deg(\{c, c'\}, V(\mathcal{G})) \leq 2\left(2|\Phi_3| + |\Phi_2|\right).$$

Combining with the fact that $3|\Phi_3| + 2|\Phi_2| + |\Phi_1| = v(\mathcal{G})$, the claim will follow. Next, assume $|\Phi_2| \geq 2$. For $e, e' \in \Phi_2$ we have

$$\deg(\{e, e'\}, \Phi_1) \leq 2|\Phi_1|,$$

$$\deg\left(\{e,e'\},\Phi_2\setminus\{e,e'\}\right)\le 4\left(|\Phi_2|-2\right),$$
$$\deg(\{e,e'\},\Phi_3)\le 8|\Phi_3|,$$

and consequently

$$4\delta(\mathcal{G})-8\le\deg\left(\{e,e'\},V(\mathcal{G})\setminus\{e,e'\}\right)\le 2|\Phi_1|+4\left(|\Phi_2|-2\right)+8|\Phi_3|.$$

Again, using $3|\Phi_3|+2|\Phi_2|+|\Phi_1|=v(\mathcal{G})$ the claim will follow. □

4.2.2 The maximal clique cover in R

We find a maximal 3-clique cover $\Phi=\{\Phi_3,\Phi_2,\Phi_1\}$ in the reduced graph R and use φ_i to denote the normalized size $|\Phi_i|/\ell$, for $1\le i\le 3$. For convenience, the families of cluster cliques in G'' corresponding to Φ_3,Φ_2,Φ_1 will also be denoted by Φ_3,Φ_2,Φ_1.

We will assume that at least one of $|\Phi_2|$ and $|\Phi_1|$ is bigger than one. If this is not the case, the G vertices in the clusters that are contained in Φ_1 and Φ_2 can be moved to V_0. As a result the vertices in $V(G)\setminus V_0$ are covered by copies of $H^{**}(L)$ (corresponding to cliques in Φ_3) and still $|V_0|\le 4\varepsilon n$.

When $\max(|\Phi_2|,|\Phi_1|)\ge 2$, we apply Lemma 12 to R and Φ, using the degree condition (10), we will get

$$\varphi_3\ge\varphi_1+2\gamma-4d. \tag{11}$$

Throughout Section 4.3, we will assume that $\varphi_3>\varphi_1+2\gamma+\sigma$ for some positive number σ defined in (9). The exceptional case that $\varphi_3\le\varphi_1+2\gamma+\sigma$ will be discussed at the end.

4.2.3 Covering G'' with H^* and H^{**}

Let $s=\varphi_3-\varphi_1-2\gamma$. Our assumption is $s>\sigma$. We will modify the clique cover Φ to a collection of $H^*(L_1)$ and $H^{**}(L_2)$ copies, where $L_2=(2+\alpha)L_1=CL$ for some $0<C\le 1$. Before stating the decomposition algorithm for G'' we give the following observation:

Proposition 2 *Let V_i and V_j , $1\le i,j\le\ell$ be two clusters in G'' that correspond to endpoints of an edge e in reduced graph R. If V_i and V_j are divided to p sub-clusters $\{V_i^1,\dots,V_i^p\}$, and $\{V_j^1,\dots,V_j^p\}$, such that the size of each sub-cluster is proposioanl to L, then the (V_i^r,V_j^t), $1\le r,t\le p$ are ε'-regular pairs, with $\varepsilon'\ge\varepsilon$.*

We will refer to above sub-clustering of V_i and V_j as p-division of clusters corresponding to edge e.

We first estimate the numbers of good triangles in Φ_3 for a typical edge and a typical singleton: Consider a typical $e\in\Phi_2$ and let $\lambda=|\{\tau\in\Phi_3:\deg(e,\tau)\le 3\}|/\ell$, then

$$2\left(\frac{1}{2}+\gamma-2d\right)\ell\le\deg(e,R)\le\left(4\varphi_3-t+2\varphi_2+\varphi_1\right)\ell+2b, \tag{12}$$

which implies $\lambda \leq \varphi_1 + s + 5d$, in turn, there are at least $(2\gamma - 5d)\ell$ good triangles in Φ_3 for e. Then consider a typical singleton c in Φ_1. This time let $\lambda = \left|\left\{\tau \in \Phi_3 : c \overset{\leq}{\hookrightarrow} \tau\right\}\right| / \ell$. Then

$$\left(\frac{1}{2} + \gamma - 2d\right)\ell \leq (2\varphi_3 - \lambda + \varphi_2)\ell + b,$$

$$\lambda \leq \frac{s}{2} + 3d.$$

Consequently there are at least $(\varphi_1 + 2\gamma + \frac{s}{2} - 3d)\ell$ good triangles for c.

Our algorithm of decomposing G'' contains two phases: Modifying $\mathcal{K}_2$ and Modifying Φ_1. In each phase we will first state the algorithm then verify it.

Phase I: Modifying Φ_2.

1. Recall that u is the neck of the bottle-graph H. We will u-divide each edge of Φ_2. Denote new size of clusters L'.

2. For each refined typical edge, (say, generated from e_0), we arbitrarily pick a "alive" (will be explained later) good triangle τ for e_0 in Φ_3. In τ, the cluster $Tip(\tau)$ is adjacent to both ends of e_0. We take $\alpha L'$ vertices of $Tip(\tau)$ and all the vertices in two clusters of e_0 to form a $H^*(L')$. A triangle τ could be chosen by more than one edges, and by Proposition 1.4, the cluster that we used in τ, $Tip(\tau)$ is same for all typical edges. The triangle τ is called "alive" unless only αL vertices are left in $Tip(\tau)$, i.e., τ has been chosen $\frac{(1-\alpha)L}{\alpha L'}$ times, (in this case we call it "dead").

3. By definition, for each atypical (original) edge $e \in \Phi_2$, there exist at least $\left\lceil \frac{1+\alpha}{1-\alpha} \right\rceil$ triangles which is over-connected to e. By Proposition 1.5, these triangles do not participate in Step 2. In each of these triangles,τ, there are at least two clusters which are adjacent to both ends of e, we label arbitrary one of them with $Tip(\tau)$. Thus we can associate refined edges generated from e with these over-connected triangles following the procedure of Step 2.

4. After modifying Φ_2, the resulting triangles in Φ_3 might have less than L vertices in one of its clusters (which we called $Tip(\tau)$). Suppose for some triangle, cluster sizes are (L, L, T) where $\alpha L \leq T \leq L$, $\alpha L' \mid L - T$. If we set $m = \frac{(1-\alpha)L}{\alpha L'}$, we can write $T = L - s\alpha L'$ where $0 \leq s \leq m$. We then divide this triangle into $H^*(\frac{L-T}{1-\alpha})$ and $H^{**}(\frac{T-\alpha L}{1-\alpha})$. Using $\frac{L-T}{1-\alpha} = s\frac{\alpha}{1-\alpha}L'$ and $\frac{T-\alpha L}{1-\alpha} = (m-s)\frac{\alpha}{1-\alpha}L'$, we further refine $H^*(\frac{L-T}{1-\alpha})$ to s copies of $H^*(\frac{\alpha}{1-\alpha}L')$, and refine $H^{**}(\frac{T-\alpha L}{1-\alpha})$ to $m-s$ copies of $H^{**}(\frac{\alpha}{1-\alpha}L')$.
Note, in our algorithm, each triangle in Φ_3, $H^{**}(L)$ which were used by s (refined) edges, becomes $H^{**}(L - s\frac{\alpha}{1-\alpha}L')$.

Only Step 2 and 3 need to be justified:

Let us see why Step 3 is feasible. Indeed, each triangle that is over-connected to the atypical (original) e will provide up to $(1-\alpha)L$ vertices from $Tip(\tau)$, i.e., it can used by $\frac{(1-\alpha)L}{\alpha L'}$ refined edges. Totally, vertices provided by $\left\lceil \frac{1+\alpha}{1-\alpha} \right\rceil$ triangles are more than what is needed by the refined edges from e:

$$\left\lceil \frac{1+\alpha}{1-\alpha} \right\rceil \frac{(1-\alpha)L}{\alpha L'} > \frac{L}{L'}.$$

In Step 2, since any typical edges has at least $(2\gamma - 5d)\ell$ good triangles, the algorithm can not continue only if all of these triangle become "dead". A triangle is considered as "dead" when it has been used for $\frac{(1-\alpha)L}{\alpha L'}$ refined edges. The algorithm is valid as long as we can show that

$$\varphi_2 \ell \frac{L}{L'} \le \frac{(1-\alpha)L}{\alpha L'}(2\gamma - 5d),$$

Or show that $2\gamma - 5d - \frac{\alpha}{1-\alpha}\varphi_2$ is positive.
In fact,

$$\begin{aligned} 2\gamma - 5d - \frac{\alpha}{1-\alpha}\varphi_2 &= 2\gamma - 5d - \frac{\alpha\,(1 - 3\varphi_3 - \varphi_1)}{2\,(1-\alpha)} \\ &= 2\gamma - 5d - \frac{\alpha}{1-\alpha}\left(\frac{1}{2} - \frac{3}{2}\,(\varphi_1 + 2\gamma + s) - \frac{1}{2}\varphi_1\right) \\ &= 2\gamma\left(1 + \frac{3\alpha}{2(1-\alpha)}\right) - \frac{\alpha}{2(1-\alpha)} + \frac{2\alpha}{1-\alpha}\varphi_1 + \frac{3\alpha}{2(1-\alpha)}s - 5d \\ &= \frac{2\alpha}{1-\alpha}\varphi_1 + s_1, \end{aligned} \tag{13}$$

for $s_1 = \frac{3\alpha}{2(1-\alpha)}s - 5d > \alpha\sigma \gg d > 0$.

After Phase I, Φ_2 and Φ_3 become disjoint copies of $H^*(L')$, $H^*(\frac{\alpha}{1-\alpha}L')$and $H^{**}(\frac{\alpha}{1-\alpha}L')$. We put the copies of H^* aside, and denote the new triangle family Φ_3', or $\Phi_3' = \left\{H^{**}(\frac{\alpha}{1-\alpha}L')\right\}$. Let $L'' = \frac{\alpha}{1-\alpha}L' = \frac{L}{w-u}$. By the note in Step 4, after handling $\varphi_2\ell$ original edges, the number of G vertices located in the balanced triangles are $3\left(\varphi_2\frac{\alpha}{1-\alpha}\right)n$ less. So $|\Phi_3'| = \left(\varphi_3 - \varphi_2\frac{\alpha}{1-\alpha}\right)\frac{L}{L''}$.

Phase II: Modifying Φ_1.

5. For each typical singleton c, we choose $L/(\frac{1-\alpha}{1+\alpha}L'')$ good triangles in Φ_3'. Suppose in these triangles τ, $V_1(\tau), V_2(\tau)$ denote the two clusters of τ connected to c, and $Top(\tau)$ be the third cluster of τ. We will form two copies of $H^*\left(\frac{L''}{2}\right)$ in τ using $\frac{1+\alpha}{2}L''$ vertices of the clusters $V_1(\tau)$ and $V_2(\tau)$, and the whole cluster $Top(\tau)$. The remaining vertices in $V_1(\tau), V_2(\tau)$ and $\frac{1-\alpha}{1+\alpha}L''$ vertices of the cluster c will form two copies of $H^*\left(\frac{1-\alpha}{2(1+\alpha)}L''\right)$. This work is repeated for all $L/(\frac{1-\alpha}{1+\alpha}L'')$ chosen triangles.

6. Every atypical singleton c has at least $\left\lceil \frac{1+\alpha}{1-\alpha} \right\rceil$ over-connected triangles in Φ_3 originally. By Proposition 1.5, these triangles were not used in Phase I, Step 2 and 3 (but $w-u$-refined in Step 4 and have not be used by other singletons either. Thus there are

$\left\lceil\frac{1+\alpha}{1-\alpha}\right\rceil(w-u) \geq L/(\frac{1-\alpha}{1+\alpha}L'')$ over-connected triangles in Φ_3' and we can follow the above procedure.

We only need to justify that there are enough good triangles even for the singleton that is considered at last. In fact, in Φ_3, there are at least $(\varphi_1 + 2\gamma + \frac{s}{2} - 3d)\ell$ good triangles for any typical c. In Φ_3', this number becomes $(\varphi_1 + 2\gamma + \frac{s}{2} - 3d - \frac{\alpha}{1-\alpha}\varphi_2)\ell\frac{L}{L''}$. To guarantee that each singleton has a disjoint set of $L/(\frac{1-\alpha}{1+\alpha}L'')$ good triangles, we need to verify that:

$$(\varphi_1 + 2\gamma + \frac{s}{2} - 3d - \frac{\alpha}{1-\alpha}\varphi_2)\ell\frac{L}{L''} - \frac{L}{\frac{1-\alpha}{1+\alpha}L''}\varphi_1\ell > 0,$$

or

$$\varphi_1 + 2\gamma + \frac{s}{2} - 3d - \frac{\alpha}{1-\alpha}\varphi_2 - \left(\frac{1+\alpha}{1-\alpha}\right)\varphi_1 > 0.$$

In fact, we have known that, $2\gamma - \frac{\alpha}{1-\alpha}\varphi_2 = \frac{2\alpha}{1-\alpha}\varphi_1 + s_1 + 5d$, so the left-hand side of this inequality is

$$\frac{s}{2} + s_1 + 2d > \frac{\sigma}{2} > 0. \tag{14}$$

Consequently, following the decomposition algorithm G'' will be covered by disjoint copies of H^* and H^{**}, and by (14), at least $\frac{\sigma}{2}n$ of its vertices are covered by copies of H^{**}. It shoudl be mentioned that in this decomposition there are four possible sizes for copies of H^*; $H^*(\frac{L}{u})$, $H^*(\frac{L}{(w-u)})$, $H^*(\frac{L}{2(w-u)})$ and $H^*(\frac{L}{2(w+u)})$, while all the balanced triangles will be of the form $H^{**}(\frac{L}{w-u})$. Using Proposition 2 we can replace this cover with one containg copies of $H^*(L_1)$ and $H^{**}((2+\alpha)L_1)$, for $L_1 = \frac{L}{2u(w-u)(w+u)(2w+u)}$. Hence,

Lemma 13. *If $\varphi_3 - \varphi_1 - 2\gamma > \sigma$, the G'' can be covered by vertex disjoint copies $H^*(L_1)$ and $H^{**}(L_2)$, for $L_2 = (2+\alpha)L_1 = CL$ for some $0 < C \leq 1$. Moreover, at least $\frac{\sigma}{2}n$ of vertices in $V(G)$ will belong to copies of H^{**}.* □

4.3 Handling of Exceptional Vertices

We prove a more general lemma:

Lemma 14 *Asume $G = G_1 \cup G_2 \cup V_0$ satisfying:*

1. $|V(G_2)| = \rho n, |V_0| = \theta n, \rho \gg \theta \geq \varepsilon$
2. *G_1 is tiled by $H^*(L_1)$ and G_2 is tiled by $H^{**}(L_2)$, where $L_2 = (2+\alpha)L_1 = CL$ for some constant C between 0 and 1,*

Then G has H-matching which leaves at most $\frac{3}{\gamma}w$ uncovered.

Proof
Let $G'' = G_1 \cup G_2$ and $\mathcal{H}$ as the family of H^* and H^{**}. The families of H^* and H^{**} are denoted by $\mathcal{H}^*$ and $\mathcal{H}^{**}$, respectively. Let $h_1 = |\mathcal{H}^*|, h_1 = |\mathcal{H}^{**}|$. By definition,

$$(2+\alpha)L_1h_1 + 3(2+\alpha)L_1h_2 = |V(G'')| = n - \theta n \tag{15}$$

We let $\rho \leq \frac{1}{2}$, since we can always decrease the number of H^{**} in $\mathcal{H}$ by dividing one copy of H^{**} $((2+\alpha)L')$ to three copies of $H^*(L')$.

For $\tau = \{U_1, U_2, U_3\} \in \mathcal{H}^*$, we always denote the cluster which contains αL_1 vertices by U_3, and refer it to a *small* cluster.

We will denote two complete tripartite graphs $K(w+u, w+u, 2w)$ and $K(h,h,h)$ by H_2 and H_3 respectively. It is obvious that both H_2 and H_3 are H-factors. For convenience, we also use "defect" graphs $H^- = K(w, w, u-1), H_3^- = K(h, h, h-1)$.

We will use a corollary of Key Lemma in [10]:

Lemma 15 *If $\tau = \{V_1, V_2, V_3\}$ is an original cluster triangle in G'', i.e. $|V_i| = L$ and (V_i, V_j) is an ε-regular edge. Assume $\tau' = \{U_1, U_2, U_3\}$ is generated from τ satisfying $U_i \subset V_i, |U_i| > d_1L, i = 1,2,3,$ where $d_1 \gg \varepsilon$. Then $H_3 \subset \tau'$, in particular, $H \subset \tau'$, with its small color class in any of U_1, U_2, U_3.*

□

For each vertex $v \in V_0$, we will find certain vertices from G'' to construct one or three copies of H with v. The following notation will be used: for a cluster U, we write $v \sim U$ if $\deg(v, U) \geq d\,|U|$. There are two ways to make up H copies including v:

- Suppose $\tau = \{U_1, U_2, U_3\} \in \mathcal{H}^{**}$ satisfies: $\exists i_0 \in \{1,2,3\}$, such that $v \sim U_j\ \forall j \neq i_0$. U_i is called a *host* cluster of v. Applying Lemma 15, we can find a H_3^- in τ such that $H_3^- \cup \{v\}$ forms an H_3. In fact, when $j \neq i_0$, we simply use $N(v, U_j)$ to replace U_j in Lemma 15. After removing this H_3, U_{i_0} has one more vertices than other two clusters.

- If $\tau = \{U_1, U_2, U_3\} \in \mathcal{H}^*$ satisfies $v \in V_0$ satisfies $v \sim U_1$ and $v \sim U_2$, U_3 is called a *small host* of v. We can remove a H^- from τ with $u-1$ vertices in U_3, such that $H^- \cup \{v\}$ forms an H. This removal causes that in the remaining triangle U_3 has one more vertices than α times $|U_1|$ or $|U_2|$.

We will consider the vertices of V_0 one by one and remove H or H_3 copies whenever there exists a host. Although sets G'', U_i have less vertices after removal, we still use the same notations for them. To prevent any cluster from losing too many vertices, we stop conisdering any cluster triangle when any of its clusters has lost $\sqrt{\theta}L_1$ vertices, (assume $\sqrt{\theta} \ll \rho$. We call this as *usage constrain*. The following lemma guarantees that this algorithm can be proceeded for all of V_0 :

Claim 16. *Suppose v is the m^{th} vertex in V_0. There exist hosts for v, who have not lost $\sqrt{\theta}L_1$ vertices.*

Proof of Claim. After we remove $m-1$ copies of H or H_3, G'' lost at most $(m-1)(3h-1)$ vertices. First we estimate the number of hosts (large or small) then conclude that some of them are always available even under the usage constraint.

Let m_1 denote the number of elements of $\mathcal{H}^*$ whose small clusters are hosts of v, and m_2 as the number of elements of $\mathcal{H}^{**}$ which contain at least one host of v. The degree condition implies:

$$\left(\frac{1}{2}+\gamma-3h\theta\right)n \leq \left(\frac{1}{2}+\gamma-\theta\right)n-(m-1)(3h-1) \leq \deg(v, V(G'')) \leq \\ (h_1-m_1)(1+\alpha+d)L_1+m_1(2+\alpha)L_1+(h_2-m_2)(1+2d)(2+\alpha)L_1+3m_2(2+\alpha)L_1$$

We divide two ends of inequality by $n-\theta n$ and let $t_1=\frac{m_1(2+\alpha)L_1}{n-\theta n}, t_2=\frac{m_2 3(2+\alpha)L_1}{n-\theta n}$. Recall that $\frac{1+\alpha}{2+\alpha}=\frac{1}{2}+\gamma$ and $\rho=\frac{3(2+\alpha)L_1h_2}{n-\theta n}$, also using (15), we obtain:

$$\begin{aligned} \frac{1}{2}+\gamma-3h\theta &\leq (1-\rho-t_1)\left(\frac{1+\alpha+d}{2+\alpha}\right)+t_1+\left(\frac{1+2d}{3}\right)(\rho-t_2)+t_2 \\ &\leq \left(\frac{1}{2}+\gamma+d\right)(1-\rho)+\left(\frac{1}{2}-\gamma\right)t_1+\left(\frac{1}{3}+d\right)\rho+\frac{2}{3}t_2, \end{aligned}$$

then

$$\left(\frac{1}{2}+\gamma-\frac{1}{3}\right)\rho-d-3h\theta \;\leq\; \frac{1}{2}t_1+\frac{2}{3}t_2 \;\leq\; \frac{7}{6}\max(t_1,t_2)$$

which implies $\max(t_1,t_2) \geq \left(\frac{1}{7}+\frac{6}{7\gamma}\right)\rho-d-3h\theta > \frac{1}{7}\rho$. Because $t_1=\frac{m_1(1-\rho)}{h_1}, t_2=\frac{m_2\rho}{h_2}$, we conclude that, either $m_1>\frac{1}{7}\rho h_1$ or $m_2>\frac{1}{7}h_1$. Because

$$\min\left(\frac{1}{7}\rho h_1, \frac{1}{7}h_2\right) > \sqrt{\theta}\frac{n}{L_1} = \frac{\theta n}{\sqrt{\theta}L_1}$$

We can always find a host who has not lost many vertices. □

To achieve the super-regularity, after V_0 becomes empty, we add a few more vertices to V_0: for a cluster $U_i \in \tau$, we remove all the vertices v for which there exists an $j \neq i$, such that $\deg(v, U_j) < (d-\varepsilon)|U_j|$. The ε-regularity condition guarantees that at most $2\varepsilon|U_i|$ such vertices exist in each cluster of τ. When $\tau \in \mathcal{H}^*$, we may also remove extra vertices such that $|U_1|=|U_2|, |U_3| \geq \alpha|U_1|$. For these new V_0 ($|V_0| \leq 6\varepsilon n$), we repeat the procedure described before. Finally the edges in the resulting triangles satisfies the super-regularity.

Thus, resulting "H^*" copies does not follow the original definition of H^*. We define ϵ-pseudo-$H^*(L)$ as a cluster triangle $\tau=\{U_1,U_2,U_3\}$ satisfying $L \geq |U_1|=|U_2| \geq (1-\epsilon)L, \alpha|U_1|+\epsilon L \geq |U_3| \geq \alpha|U_1|$ for some $\epsilon \ll 1$. Let $\theta_1=\sqrt{\theta}+2\varepsilon$. It is easy to see that resulting triangles are θ_1-pseudo-$H^*(L_1)$.

A descrepency might exist among clusters of a copy of "$\mathcal{H}^{**}$". Suppose each U_i had hosted x_i exceptional vertices, thus, for example, U_1 lost $x_1(h-1)+x_2h+x_3h$. Let $x=x_1+x_2+x_3$. Thus current U_i has size $(2+\alpha)L_1-xh+x_i$ for $i=1,2,3$. We will divide all triangles in $\mathcal{H}^{**}$ to three copies of pseudo-H^*, such that clusters sizes are $(L(x), L(x), \alpha L(x)+x_3), (L(x), \alpha L(x)+x_2, L(x)), (\alpha L(x)+x_1, L(x), L(x))$, respectively, where $L(x) = L_1 - wx$.

Note, although super-regularity existed for the edges of "H^{**}", it may not be inherited within new copies of "H^*". In this case, we will remove more vertices from three new copies of "H^*". For example, suppose $H^{**} \to \tau^1, \tau^2, \tau^3$, $\tau^i = \{U_1^i, U_2^i, U_3^i\}$, where U_i^i is the small cluster. If $v \in U_1^1$ has positive degree only in U_2^2 and U_3^3, we will remove a $H_3 = K(h,h,h)$ from U_1^1, U_2^2 and U_3^3. Set $a = \lceil 1/\alpha \rceil$. We next remove $2a$ copies of H_3 from remaining 6 clusters (ah vertices from each), thus remaining triangles are pseudo-H^* (pseudo means $|U_3| \geq \alpha|U_1|, |U_1| = |U_2|$.)

Since $xh - x_i < \sqrt{\theta} L_1$, it is easy to see that resulting pseudo-H^*s are also θ_1-pseudo-$H^*(L_1)$s.

Thus the resulting cover (still denoted by $\mathcal{H}$) contains only copies of θ_1-pseudo-$H^*(L_1)$.

G'' still denotes vertices of G in the cluster graph. Note, $G \setminus G''$ has been covered by copies of H. We know, $n - |G''| \leq C\theta n$ for some constant C. To tile G'' with H, we first consider the largest possible H-matching inside pseudo-H^*.

Let us still denote corresponding vertices of G in the $\mathcal{H}$, Before we describe the algorithm of covering the corresponding G vertices in $\mathcal{H}$, we will observe the possible H-matching inside the cluster triangle and then consider the connection among clusters in the different triangle.

Claim 17. *In each pseudo-H^*, we can find a H-matching that covers all but at most $4w$ vertices.*

Proof. Suppose $\tau = \{U_1, U_2, U_3\}$ is a triangle of $\mathcal{H}$, with $|U_1| = |U_2| = s, |U_3| = \alpha s + t$. We will remove i copies of H_2 from τ in the way that $2w$-vertex classes always come from U_3. New clusters are denoted by U_1', U_2', U_3', and we wish that $|U_3'| = \alpha|U_1'| (= \alpha|U_2'|)$. That is

$$\alpha s + t - 2iw = \alpha(s - i(w+u)),$$

or

$$i = \frac{t}{2w - \alpha(w+u)}.$$

This implies, we can achieve at least, $|U_3'| - \alpha|U_1'| < 2w - \alpha(w+u)$. On the other hand, if we apply Blow-up Lemma in $H^*(m)$, for sufficiently large m, we can find a H-tiling that covers all but $h-1$, since m may not be divisible by w. Together, there might be totally at most $h + 2w - \alpha(w+u) \leq 4w$ left-over vertices in τ. □

Corollary 18 *G'' has a H-matching that leaves only constant vertices*

If we simply apply Claim 17 to each member of $\mathcal{H}$, at most $4w\,|\mathcal{H}|$ vertices will be uncovered. Because $|\mathcal{H}| = \langle_\infty + \ni \langle_\in = \mathcal{O}(\ell)$ and $\ell \leq M(\varepsilon)$ from Lemma refr10, the corollary follows. □

Corollary 18 already proved Conjecture 1. In order to make the number of uncovered to a constant irrelevant to $\mathcal{H}$, let us now consider the connection between two clusters in different triangles. Our updated cluster graph is denoted by R, that is, the vertices of R' are clusters, large or small, $(X, Y) \in E(R)$ if and only if $d(X, Y) > d - \varepsilon$.

we define a directed graph $\mathcal{D}$ whose vertices are all the clusters (big or small) in the elements of $\mathcal{H}$ and define a directed edge from U to U', $U \to U'$, if U' belongs to some triangle $\tau \in \mathcal{H}$ such that $(U, U'') \in E(R)$ $\forall U'' \in \tau \setminus \{U'\}$.

Claim 19. *For any $C \in V(\mathcal{D})$, $\deg_{out}(C) \geq \frac{4\gamma}{3}|V(\mathcal{D})|$.*

Proof. Since less than $\theta_1 L_1$ vertices have been removed, before and after removal $E(R)$ was not changed. We can assume that each cluster has the same size L_1 and the total number of underlying G-vertices is $n - \theta n$. The average degree of a cluster is defined as

$$\deg(X) = \frac{1}{|X|}\sum_{v \in X} \deg_{G''}(v) = \sum_{(X,Y) \in E(R)} d(X, Y)|Y|$$

Fix a cluster $c \in V(R)$. Suppose $\tau = \{U_1, U_2, U_3\}$ in $\mathcal{H}$, again U_3 is the small cluster. We define t_1, as the percentage of the triangles $\tau \in \mathcal{H}$ in which both (c, U_1) and (c, U_2) are edges in $E(R)$, t_2 as the percentage of τ in which (c, U_3) and only one of (c, U_1) and (c, U_2) are in $E(R)$, t_3 as the percentage of τ in which $(c, U_i) \in E(R)$, for $i = 1, 2, 3$.

Cosider the average degree of the cluster c, we have

$$\frac{1}{2}+\gamma-2d \leq \deg(c) \leq (1-2\gamma)t_1+\left(\frac{1}{2}+\gamma\right)t_2+t_3 = \left(\frac{1}{2}-\gamma\right)+\left(\frac{1}{2}-\gamma\right)t_1+2\gamma t_2+\left(\frac{1}{2}+\gamma\right)t_3$$

Since $\gamma = \frac{\alpha}{2(2+\alpha)} \leq \frac{1}{6}$, we have $(\frac{1}{2}+\gamma) \leq 2(\frac{1}{2}-\gamma)$. This in turn will imply

$$\begin{aligned} 2\gamma - 2d &\leq \left(\frac{1}{2}-\gamma\right)t_1 + 2\gamma t_2 + \left(\frac{1}{2}+\gamma\right)t_3 \\ &\leq \left(\frac{1}{2}-\gamma\right)(t_1 + t_2 + 2t_3). \end{aligned}$$

Therefore in $\mathcal{D}$,

$$\begin{aligned} \deg_{out}(c) &= (t_1 + t_2 + 3t_3)\,|\mathcal{H}| \\ &\geq \frac{2\gamma - 2d}{\frac{1}{2}-\gamma}\,\frac{|V(\mathcal{D})|}{3} \\ &> \frac{4\gamma}{3}|V(\mathcal{D})|. \qquad \square \end{aligned}$$

In a directed graph D the *source set* of v is defined as $\mathcal{W}(v) = \{u \in V(D) : \exists$ a directed path from u to $v\}$. It is easy to see that $|\mathcal{W}(u)| \geq |\mathcal{W}(v)|$, if $u \in N_{out}(v)$. A set $\mathcal{S}(D) \subseteq V(D)$ will be called a *sink* set if $V(D) = \bigcup_{v \in \mathcal{S}(D)} \mathcal{W}(v)$.

Lemma 20. *In a directed graph D, with $\delta = \min_{v \in V(D)} \deg_{out}(v)$, there is a sink set $\mathcal{S}$ of size at most $\frac{|V(D)|}{\delta+1}$.*

Proof. Let x_1 be a vertex in which $|\mathcal{W}(x_1)| = \max_{v \in V(D)} |\mathcal{W}(v)|$. Since $N_{out}(x_1) \in \mathcal{W}(x_1)$ (otherwise, some out-neighbor of x_1 produces a larger source set), we have $|\mathcal{W}(x_1)| \geq \delta + 1$. Let $D' = D \setminus \mathcal{W}(x_1)$, then for every vertex $v \in D'$, the set $N_{out}(v)$ in D' is the same as $= N_{out}(v)$ in D, since, if $v' \in N_{out}(v)$ in D and $v' \in \mathcal{W}(x_1)$, v should also be in $\mathcal{W}(x_1)$. Now, let x_2 be a vertex with the largest source set in D'. Similar to x_1, we have $\mathcal{W}(x_2) \geq \delta + 1$ in D'. The above procedure can be repeated at most $\frac{|V(D|)}{\delta+1}$ times and the proof will follow. $\square$

The algorithm after deleting the exceptional vertices and obtaining a H^*-cover $\mathcal{H}$ is as follows:

Using Claim 17, in each element of $\mathcal{H}$, we calculate the number of left-over vertices (called extra number) in each cluster, but not proceed the covering yet. Then in the (cluster) directed graph $\mathcal{D}$, we find a sink set of size at most $\frac{3}{4\gamma}$ by the Claim 19 and Lemma 20.

For any cluster C^0 whose extra number x is not zero, we will find a directed path $C^0, C^1, \ldots, C^t$ from C^0 to some cluster C^t in the sink set. Suppose that $C^i \in \tau_i$ and it takes u_i vertices of C^i to form a H in τ_i ($u_i = w$, or v). For $i = 1, \ldots, t$, we find x copies of H using x vertices of C^{i-1} and $hx - x$ vertices from τ_i (where $x(u_i - 1)$ vertices from C^i). Thus, the extra number of C^0 becomes zero and the extra number of C^t is increased by x, while the H-tiling of τ_i is not impacted.

Although one cluster might be in many paths to the sink, the fact that the total of extra numbers (at most) $4w|\mathcal{H}|$ is much smaller than dL_1, guarantees that above procedure can be continued for all clusters. Finally, we apply the Blow-up Lemma to find the maximal H-tiling in all the triangles of $\mathcal{H}$. Only those containing clusters of the sink will leave (at most) $4w$ uncovered by Claim 17. Therefore, the total of left-over vertices is at most $\frac{3w}{\gamma}$.

4.4 The Exceptional Cases

Recall $\Phi = \{\Phi_1, \Phi_2, \Phi_3\}$ is the maximal clique-cover of R. The condition for exceptional cases is

$$s = \varphi_3 - \varphi_1 - 2\gamma \leq \sigma. \tag{16}$$

Depending on relative size of Φ_1 we will consider two separate cases:

4.4.1 Exceptional Case (I): $\varphi_1 < \mu$.

In this case $\varphi_3 < 2\gamma + 2\mu$. Since $\mu \ll 1 - \alpha$, we have

$$\varphi_2 = \frac{1}{2}(1 - 3\varphi_3 - \varphi_1) > \frac{1}{2}(1 - 6\gamma - 7\mu) = \frac{1}{2}\left(\frac{2(1-\alpha)}{2+\alpha} - 7\mu\right) \gg \mu. \tag{17}$$

Recall that for any two edges $e, e' \in \Phi_2$, $\deg(\{e, e'\}, \Phi_2) \leq 4\varphi_2\ell$, and $\deg(\{e, e'\}, \Phi_3) \leq 8\varphi_3\ell$. Let $m = |\{\tau \in \Phi_3 : \tau \notin \mathcal{S}\}|/\ell$ and $m' = |\{e \in \Phi_2 : e \notin \mathcal{E}\}|/\ell$, then

$$4\left(\frac{1}{2} + \gamma - 2d\right) \leq 8\varphi_3 - m + 4\varphi_2 - m' + 2\varphi_1,$$

that implies

$$m + m' \leq 4\mu + 8d.$$

Together with Proposition 5 and (17), this shows that number of edges which are over-connected to some triangle of Φ_3 is much smaller than $|\Phi_2|$. Hence we can find two edges $e_1, e_2 \in \Phi_2$ such that $\deg(e_i, \tau) \leq 4, \forall \tau \in \Phi_3$, for $i = 1, 2$. Let $\mathcal{S} = \{\tau \in \Phi_3 : \deg(\{e_1, e_2\}, \tau) = 8\} = \{\tau \in \Phi_3 : e_1 \hookrightarrow \tau, e_2 \hookrightarrow \tau\}$ and $\mathcal{E} = \{e \in \Phi_2 : \deg(\{e_1, e_2\}, e) = 4\} = \{e \in \Phi_2 : e_1 \hookrightarrow e, e_2 \hookrightarrow e\}$, Repeat above computation, all but at most $4\mu + 8d$ elements of Φ_2 and Φ_3 belong to $\mathcal{S}$ and $\mathcal{E}$, respectively. By Proposition 1.2 and 1.4 the sets $A = \{Top(e), Top(\tau) : e \in \mathcal{E}, \tau \in \mathcal{S}\}$ and $B = \{Bot(e), Bot(\tau) : e \in \mathcal{E}, \tau \in \mathcal{S}\}$ are both independent sets, with $|A| = |B| = \varphi_2 + \varphi_3 - (4\mu + 8d) > \frac{1}{2} - \gamma - 6\mu$.
Let $\beta = 6\mu$, then by the degree condition in R for all $a \in A$ and $b \in B$ we have $\deg(a, R\backslash A) > (1 - \beta)|R\backslash A|$ and $\deg(b, R\backslash A) > (1 - \beta)|R\backslash B|$. Next, we will form a set C which consists of clusters $Tip(\tau)$ for all $\tau \in \mathcal{S}$ and all the remaining clusters of R. Clearly, $|C| = 2\gamma + \beta$.
¿From now on, A, B, and C will refer to their underlying vertex sets in G. First we will move all exceptional vertices (set V_0) to C. We will remove a vertex $v \in C$ to A (or respectively to B) if $\deg(v, A) < \beta_1|A|$ (or $\deg(v, B) < \beta_1|B|$) where $\beta \ll \beta_1 \ll 1$. Let us still denote the resulting sets by A, B, and C. In the ideal case, $|A| = |B| = \frac{1}{2+\alpha}n$ and $|C| = \frac{\alpha}{2+\alpha}n$. Then by super-regularity between every two class in $\{A, B, C\}$, using the Blow-up Lemma we can find the desired H-factor in G.
If $|A| < \frac{1}{2+\alpha}n$ and $|B| < \frac{1}{2+\alpha}n$, an argument similar to Claim 17 shows that all but $5w$ vertices of G can be covered by disjoint copies of H. Consequently, we can assume that $|A| > \frac{1}{2+\alpha}n$, and $|B| < \frac{1}{2+\alpha}n, |C| < 2\gamma n$ (the other cases are similar).
¿From A we will move a vertex v for which $\deg(v, A) > \beta_1|A|$ to one of the classes B or C with fewer vertices, without hurting super-regularity. Let A, B, and C still denote the resulting sets. After this step, either we can achieve the ideal case, or for all $v \in A, \deg(v, A) \leq \beta_1|A|$. Let us assume that the latter is the case, and set $t = |A| - \frac{1}{2+\alpha}n$, we will use the following fact:

Proposition 3.

$$\nu_i(G) \geq (\delta(G) - i + 1)\frac{n}{2(i+1)\Delta(G)}$$

To see this take a maximal set of i-stars, let m denote its size, and $\mathbf{E}$ be the number of edges between the stars and the remaining vertices of G, we have the following chain of inequalities which in turn proves this proposition:

$$(n - m(i+1))(\delta(G) - (i-1)) \leq \mathbf{E} \leq m(i+1)\Delta(G).$$

Proposition 3 implies that we can find t vertex disjoint w-stars in A and by moving their centers to either B or C we can immediately find t copies of H. The remaining sets A, B, and C have size ratio $(1, 1, \alpha)$ and satisfy super-regularity condition. Applying the Blow-up Lemma will complete the proof.

4.4.2 Exceptional Case (II): $\varphi_1 \geq \mu$

Fix two typical clusters $c, c' \in \Phi_1$. Define m and m' as $\left|\{\tau \in \Phi_3 : \deg(\{c, c'\}, \tau) \leq 3\}\right| / \ell$ and $\left|\{e \in \Phi_2 : \deg(\{c, c'\}, e) \leq 1\}\right| / \ell$, respectively. Then

$$2\left(\frac{1}{2} + \gamma - 2d\right) \leq 4\varphi_3 - m + 2\varphi_2 - m',$$

that implies

$$m + m' \leq s + 4d < 2\sigma.$$

Therefore, there are totally $\varphi_3 + \varphi_2 - 2\sigma$ good triangles and edges for c and c'. Using Proposition 1.3 for each $\tau \in \Phi_3$ and $e \in \Phi_2$ the clusters $Top(\tau)$ and $Top(e)$ play the same role as clusters c and c'. Together with the rest of the clusters in Φ_1, these clusters form an independent cluster set A with $|A| \geq (\varphi_3 + \varphi_2 + \varphi_1 - 2\sigma)\ell \geq \left(\frac{1}{2} - \gamma - \frac{\sigma}{2} - 2\sigma\right)\ell$. We will define B as the remaining clusters of these good triangles and edges, thus B is made up of edges and singletons. The vertices of those clusters not in A or B will be added to exceptional set V_0. Clearly, we will have $|V_0| < 6\sigma n$.

Let V_b be the union of corresponding G vertices in the clusters of B with V_0. Our objective is to cover V_b with copies of $K(w, u)$, then combine each copy of $K(w, u)$ with w vertices of A to obtain the desired H-matching. For simplicity we define H_0 as $K(w, u)$. To find the H_0-tiling of $B \cup V_0$, we almost repeat Section 4.2, 4.3. The only difference is that each each copy of H_0 should satisfy the property that its vertices share many common neighbors in A.

Because B is made up of edges and singletons, as the way to modify $\Phi_3 \cup \Phi_2$ to balanced and balanced cluster triangles, we modify B to balanced and balanced cluster edges E^*, E^{**}. It is possible, because

$$2\gamma - \frac{\alpha}{1-\alpha}\varphi_2 > \frac{2\alpha}{1-\alpha}\varphi_1 = C\mu.$$

¿From now on, A will be referred to its underlying G vertex set Let $\beta = 6\sigma$, we have

$$\left(\frac{1}{2} - \gamma - \beta\right) n \leq |A| \leq \left(\frac{1}{2} - \gamma\right) n,$$

and

$$\deg(v, V_b) \geq \left(\frac{1}{2} + \gamma - 3d\right) n \geq (1-\beta)|V_b|, \ \ \forall v \in A.$$

Let us move (at most $2\beta|V_b|$) vertices v from V_b to A if

$$\deg(v, A) < \beta_1 |A|, \tag{18}$$

with ($\beta \ll \beta_1 \ll \mu$). To finish the proof of Main Theorem, we finally consider following three cases:

Case (1): $|A| = (\frac{1}{2} - \gamma)n$

Let $\mathcal{V}$ be a subset of V_b, in which $\forall v \in \mathcal{V}$,

$$\deg(v, A) \leq (1 - \beta_2)|A|, \tag{19}$$

with $\beta \ll \beta_2 \ll \beta_1$.

Let us first consider one special case that $|\mathcal{V}| = 0$. We will find a H_0-matching that covers all but constant vertices of V_b. Then Since (19) does not hold for any vertex of V_b, for every copy of H_0 with vertices $\{v_1, \dots, v_{w+u}\}$, the common neighborhood of $v_i, i = 1, \dots, w+u$ will almost be the whole A. On the other hand, the common neighborhood of any w vertices in A is almost $|V_b|$. We combine all but constant number of vertices of A into w-vertex sets. By König-Hall Lemma, it is easy to see that there exists a perfect matching between H_0 copies from V_b and w-sets from A. This implies G contains disjoint copies of H except for constant vertices.

When there were no vertices in any clusters of B satisfying (18), tiling V_b with H_0 is exactly as the way in Section 4.3. When there are vertices that were moved from V_b to A because of (18), we need to control the impact of losing vertices to clusters. Assume an $E^*(L_1)$ copy (U_1, U_2) with $|U_2| = \alpha|U_1|$ experienced vertices loss. We will move some vertices to V_0 such that (I) both of resulting U_1, U_2 have more than $\frac{L_1}{2}$ vertices, (II) it satisfies $1 \geq \frac{|U_1|}{|U_2|} \geq \alpha$. For an damaged $H_0^{**}(L_2)$ copy, we will remove some vertices to V_0 such that it becomes $H_0^{**}(L_2')$ with some $L_2' \geq \frac{L_2}{2}$. It is easy to see that, after these steps, size of V_0 is smaller than $C_0\beta n$, where $C_0 = \frac{4}{\alpha}$. Next we follow Lemma 14 to find a H_0-matching that leaves only $\frac{\frac{1}{2}+\gamma}{\gamma}2w$ vertices. After combining with set A, the total number of uncovered (by H) vertices is bounded by

$$2w\frac{\frac{1}{2}+\gamma}{\gamma} \cdot \frac{1}{\frac{1}{2}+\gamma} = \frac{2w}{\gamma}.$$

When $|\mathcal{V}| > 0$, we will need an H_0-matching of $B \cup V_0$ in which each copy of H_0 contains at most one vertex of $\mathcal{V}$. Then the common degree of vertices of an H_0 copy in A is still big and the König-Hall Lemma still holds. Observe that in the special case the vertices of V_0 were distributed in different H_0's and $|\mathcal{V}| < \frac{2\beta}{\beta_2}n < \beta_2 n$. We will simply move the vertices of $\mathcal{V}$ to V_0 and follow the previous procedure.

Case 2 $|A| < (\frac{1}{2} - \gamma)n$

In Case 1 each copy of H will have a w set coming from A. When $|A| < (\frac{1}{2} - \gamma)n$ we will change the tiling procedure such that some of the copies of H have their u sets in A. The rest is similar to Case 1.

Case 3 $|A| > (\frac{1}{2} - \gamma)n$

Similar to exceptional Case (I) in Section 4.4.1 we will find $|A| - (\frac{1}{2} - \gamma)n$ copies of w-star in A, reducing the size of A, and then finish the tiling using an argument similar to Case 1. $\square$

References

[1] Sarmad Abbasi, The solution of the El-Zahar problem, Ph. D. Thesis, Rutgers University, 1998.

[2] N. Alon and R. Yuster. Almost h-factors in dense graphs. *Graphs and Combinatorics*, (8):95–107, 1992.

[3] N. Alon and R. Yuster. h-factors in dense graphs. *J. Combinatorial Theory*, (B66):269–282, 1996.

[4] H. Corrádi and A. Hajnal. On the maximal number of independent circuits in a graph. *Acta Math. Hung.*, (14):423–439, 1963.

[5] G. A. Dirac. Some theorem on abstract graphs. *Proc. London Math. Soc.*, (2):69–81, 1952.

[6] M. H. El-Zahar, On Circuits in Graphs, Discrete Math. 50 (1984), 227-230.

[7] A. Hajnal and E. Szemerédi. Proof of a conjecture of Erdös. In P. Erdös and V. T. Sós, editors, *Combinatorial Theory and Its Applications, II.* Colloquia Mathematica Societatis János Bolyai, North-Holland, Amsterdam/London, 1970.

[8] J. Komlós. Tiling turán theorems. Combinatorica 20 (2000), no. 2, 203–218.

[9] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1997), 109-123.

[10] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory (survey), Bolyai Society Mathematical Studies 2, Combinatorics, Paul ErdHos is Eighty (Volume 2) (D. Miklós, V. T. Sós and T. SzHonyi eds.), Keszthely (Hungary) (1993), Budapest (1996), 295-352.

[11] G.N. Sárközy J. Komlós and E. Szemerédi. An algorithmic version of blow-up lemma. *Random Structures and Algorithms*, (12):297–312, 1998.

[12] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture.

[13] E. Szemerédi. Regular partitions of graphs. *Colloques Internationaux C.N.R.S № 260 -problèmes Combinatoires et Théorie des Graphes*, pages 399–401, 1976.