Minimum Degree Thresholds for Bipartite Graph Tiling

Albert Bush and Yi Zhao

DEPARTMENT OF MATHEMATICS AND STATISTICS
GEORGIA STATE UNIVERSITY, ATLANTA
E-mail: yzhao6@gsu.edu

Received April 1, 2010; Revised January 17, 2011

Published online 14 April 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/jgt.20605

Abstract: Given a bipartite graph H and a positive integer n such that $v(H)$ divides $2n$, we define the minimum degree threshold for bipartite H-tiling, $\delta_2(n, H)$, as the smallest integer k such that every bipartite graph G with n vertices in each partition and minimum degree $\delta(G) \geq k$ contains a spanning subgraph consisting of vertex-disjoint copies of H. Zhao, Hladký-Schacht, Czygrinow-DeBiasio determined $\delta_2(n, K_{s,t})$ exactly for all $s \leq t$ and sufficiently large n. In this article we determine $\delta_2(n, H)$, up to an additive constant, for all bipartite H and sufficiently large n. Additionally, we give a corresponding minimum degree threshold to guarantee that G has an H-tiling missing only a constant number of vertices. Our $\delta_2(n, H)$ depends on either the chromatic number $\chi(H)$ or the critical chromatic number $\chi_{cr}(H)$, while the threshold for the almost perfect tiling only depends on $\chi_{cr}(H)$. These results can be viewed as bipartite analogs to the results of Kuhn and Osthus

Current address of Albert Bush is School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332.

Contract grant sponsor: NSA; Contract grant numbers: H98230-07-1-0019; H98230-10-1-0165 (to Y. Z.).

[Combinatorica 29 (2009), 65–107] and of Shokoufandeh and Zhao [Rand Struc Alg 23 (2003), 180–205]. J Graph Theory 70: 92–120, 2012

Keywords: *graph packing; tiling; critical chromatic number; Regularity Lemma; Blow-up Lemma*

1. INTRODUCTION

Let G be a graph on n vertices and H be a graph on h vertices. The *tiling* (sometimes called *packing*) problem in extremal graph theory is to find in G as many vertex-disjoint copies of H as possible. Researchers are interested in finding a tight minimum degree condition for G to contain an H-factor—a subgraph which consists of $\lfloor n/h \rfloor$ copies of H. This is also sometimes called a perfect H-tiling or H-packing. Dirac's theorem on Hamilton cycles [4] is one of the earliest tiling results. It implies that every n-vertex graph G with minimum degree $\delta(G) \geq n/2$ contains a perfect matching (K_2-factor). The seminal result of Hajnal and Szemerédi [6] determines the minimum degree threshold for a K_r-factor for all integers r. By applying Szemerédi's Regularity Lemma [19], Alon and Yuster [1] found minimum degree conditions that guarantee an H-factor for an arbitrary H. Komlós et al. [11] improved Alon–Yuster's result, giving a tight minimum degree for H with equal-sized color classes. Instead of using the chromatic number $\chi(H)$ as in [1, 11], Komlós [9] introduced the critical chromatic number $\chi_{cr}(H)$ and showed that it played a critical role in graph tiling (his result was improved by Shokoufandeh and Zhao [18]). Kühn and Osthus [13] finally determined exactly when the critical chromatic number or the chromatic number was the appropriate parameter. In order to accurately state their result, we need the following definitions.

Let H be a graph on h vertices with $\chi(H) = \ell$. The *critical chromatic number* $\chi_{cr}(H)$ is defined as $((\ell-1)h)/(h-\sigma(H))$, where $\sigma(H)$ is the size of the smallest color class over all proper ℓ-colorings of H. It is easy to see that $\ell - 1 < \chi_{cr}(H) \leq \ell$ with equality if and only if all ℓ-colorings of H are *balanced*, namely, all color classes have the same size. Suppose H has connected components $C_1, \ldots, C_{k_c}$. We define $hcf_c(H)$ as $hcf(|C_1|, \ldots, |C_{k_c}|)$, the *highest common factor* of integers $|C_1|, \ldots, |C_{k_c}|$. Given an ℓ-coloring C of H with $x_1 \leq x_2 \leq \cdots \leq x_\ell$ as the sizes of the color classes, let $D(C) = \{x_{i+1} - x_i | i = 1, \ldots, \ell - 1\}$. Let $D(H) = \cup D(C)$ where the union ranges over all ℓ-colorings of H. Define $hcf_\chi(H)$ as the highest common factor of $D(H)$. In particular, we set $hcf_\chi(H) = \infty$ if $D(H) = \{0\}$. Lastly, we say that

$$H \text{ is in } \textit{Class } 1 \text{ if } \begin{cases} hcf_\chi(H) = 1 & \text{when } \chi(H) \neq 2, \\ hcf_\chi(H) \leq 2 \text{ and } hcf_c(H) = 1 & \text{when } \chi(H) = 2, \end{cases}$$

otherwise H is in *Class* 2. (The authors of [13] used $hcf(H) = 1$ to denote the case when H is in Class 1.)

Theorem 1.1 (Kühn and Osthus [13]). *For every graph H on h vertices, there exist integers C and m_0 such that for all integers $m \geq m_0$, if G is a graph on $n = mh$ vertices* *then the following holds. If*

$$\delta(G) \geq \begin{cases} (1-1/\chi_{cr}(H))n + C & \text{if } H \text{ is in Class } 1, \\ (1-1/\chi(H))n + C & \text{if } H \text{ is in Class } 2, \end{cases}$$

then G contains an H-factor.

It was also shown in [13] that Theorem 1.1 is best possible up to the constant C. Other results and methods for tiling problems can be found in a recent survey of Kühn and Osthus [14].

Rather than working with an arbitrary graph G, one may restrict G to be r-partite and tile it with some r-partite graph H. Although it sounds like a special case, multipartite tiling is stronger than general tiling in the following sense. First, a result on multipartite tiling does not follow from the corresponding general result. For example, an arbitrary graph G of order n contains a perfect matching if $\delta(G) \geq n/2$ [4], while a bipartite graph B with two partition sets of size $n/2$ contains a perfect matching if $\delta(B) \geq n/4$ [7]. Second, a result on multipartite tiling often implies one for general tiling. For example, suppose we know that every bipartite graph with two partition sets of size $n/2$ and minimum degree at least $n/4$ contains a perfect matching (assumed that n is even). Let G be an arbitrary graph G with $\delta(G) \geq n/2 + \varepsilon n$ for some $\varepsilon > 0$. By taking a random balanced bipartition of G, we get a bipartite spanning subgraph B with $\delta(B) \geq \delta(G)/2 - o(n) \geq n/4$ (as $n \to \infty$). Then B contains a perfect matching, which is also a perfect matching of G.

In this article we consider tiling in a balanced bipartite graph, where an r-partite graph is *balanced* if all partition sets have the same size. Zhao [20] determined the minimum degree threshold for a $K_{s,s}$-factor in a balanced bipartite graph for all s. Hladký and Schacht [8] and Czygrinow and DeBiasio [3] later determined the minimum degree threshold for a $K_{s,t}$-factor for $s < t$. Given any bipartite H of order h, since $K_{h,h}$ contains an H-factor, the result in [20] gives a sufficient condition for an H-factor.

Theorem 1.2 (Zhao [20]). *Let H be a bipartite graph of order h. Suppose that n is sufficiently large and divisible by h. If G is a balanced bipartite graph on* $2n$ *vertices such that* $\delta(G) \geq n/2 + 3h/2 - 2$, *then G contains an H-factor.*

We first show that Theorem 1.2 is best possible (up to an additive constant) when H is in Class 2.

Proposition 1.3. *Let H be a bipartite graph on h vertices. We assume G to be a balanced bipartite graph on* $2n = mh$ *vertices where* $m \in \mathbb{N}$.

1. *If H is in Class* 2, *then there exists a G such that* $\delta(G) = \lceil n/2 \rceil - 1$ *and G does not contain an H-factor.*
2. *If H is in Class* 1, *then there exists a G such that*

$$\delta(G) = \left(1 - \frac{1}{\chi_{cr}(H)}\right) n - 1$$

and G does not contain an H-factor.

Zhao [20] asked about the minimum degree condition for H-factors in bipartite graphs and suggested using either $\chi(H) = 2$ or $\chi_{cr}(H)$ as in Theorem 1.1. The main result of this article answers this affirmatively; it can be viewed as a bipartite analog of Theorem 1.1.

Theorem 1.4. *Let H be a bipartite graph in Class 1 with h vertices. There exist positive integers m_0 and $c_1(H) \le 4h^3$ such that the following holds for all integers $m \ge m_0$. If G is a balanced bipartite graph on $2n = mh$ vertices such that*

$$\delta(G) \ge \left(1 - \frac{1}{\chi_{cr}(H)}\right) n + c_1(H),$$

then G contains an H-factor.

Proposition 1.3, Part 2, shows that Theorem 1.4 is best possible up to the value of $c_1(H)$. Our constant $c_1(H)$ is on the order of h^3, and its exact value is specified in (8) of Theorem 4.9. It is much smaller than the constant C in Theorem 1.1, which depends on the Regularity Lemma. Nevertheless, we are unable to determine the best possible value of $c_1(H)$ as in [3, 8, 20].

In other words, we determine the minimum degree threshold for bipartite tiling as follows. Given a bipartite graph H of order $h \ge 2$, let $\delta_2(n,H)$ denote the smallest integer k such that every balanced bipartite graph G of order $2n$, which is divisible by h, with $\delta(G) \ge k$ contains an H-factor. Proposition 1.3 and Theorem 1.4 together imply that

$$\delta_2(n,H) = \begin{cases} (1 - 1/\chi_{cr}(H))n + O(1) & \text{if } H \text{ is in Class 1,} \\ (1 - 1/\chi(H))n + O(1) & \text{if } H \text{ is in Class 2.} \end{cases}$$

Zhao [20] also asked for the minimum degree threshold for an almost perfect H-tiling. Komlós [9] showed that for any graph H, every graph G with n vertices and $\delta(G) \ge (1 - 1/\chi_{cr}(H))n$ contains an H-tiling that covers all but at most $o(n)$ vertices. Shokoufandeh and Zhao [18] improved $o(n)$ to a constant, $O(h^2)$, where h is the order of H. In this article we prove a similar result for bipartite tiling.

Theorem 1.5. *Let H be a bipartite graph with h vertices. There exist integers n_0 and $c_2(H) < 8h^2$ such that every bipartite graph G with $n \ge n_0$ vertices in each partition set contains an H-tiling that covers all but at most $c_2(H)$ vertices if $\delta(G) \ge (1 - 1/\chi_{cr}(H))n$.*

It is important to note that Kühn and Osthus [13] started their proof of Theorem 1.1 with the result of Komlós (alternatively they could use the one of Shokoufandeh and Zhao), which gives an almost-tiling of G, and then modified it into a perfect tiling under the assumption that H is in Class 1. While proving Theorem 1.4, since there is no Komlós theorem available, we first find an almost-tiling (which leaves $o(n)$ vertices uncovered) by using the approach in [18]. If H is in Class 1, then we modify it into a perfect H-tiling, otherwise we modify it into an H-tiling that leaves only $O(h^2)$ vertices uncovered.

The structure of the article is as follows. We prove Proposition 1.3 in Section 2. In Section 3, we lay some groundwork for our proofs: we state bipartite versions of the Regularity Lemma and Blow-up Lemma. Section 4 provides the proof of Theorem 1.4, which is divided into the nonextremal case and the extremal case. Section 5 gives the proof of Theorem 1.5 based on the one of Theorem 1.4. In the last section, we give concluding remarks, including a conjecture on r-partite tiling.

Notation. Fix a graph. For two vertices x,y, we write $x\sim y$ if x is adjacent to y. Let $\Gamma(x)$ denote the set of neighbors of x and $\deg(x)=|\Gamma(x)|$. For a vertex set S, let $\Gamma(x,S)=\Gamma(x)\cap S$ and $\deg(x,S)=|\Gamma(x,S)|$. A bipartite graph $G[X,Y]$ means a bipartite graph with partition sets X and Y. When G is given and $A,B\subseteq V(G)$ are two disjoint sets, we use $G[A,B]$ to denote the bipartite subgraph induced on $A\cup B$ and its size is denoted by $e(A,B)$. The *density* of A and B is the ratio $d(A,B)=e(A,B)/(|A|\cdot|B|)$. We will use the notation $\delta(X,Y)$ to denote the minimum degree of a vertex in X into a set Y. In other words, $\delta(X,Y)=\min_{x\in X}\deg(x,Y)$. Note that in general $\delta(X,Y)\neq\delta(Y,X)$.

Throughout this article, we assume that H is a bipartite graph on h vertices such that $\sigma(H)=u$ and $h-\sigma(H)=w$. Let $C_1,\ldots,C_{k_c}$ be its connected components. Then each component C_i has a unique 2-coloring $\{U_i,W_i\}$ with $|W_i|\geq|U_i|$. Let $c_i=|C_i|=|W_i|+|U_i|$ and $d_i=|W_i|-|U_i|$. Recall that $hcf_c(H)=hcf(c_1,\ldots,c_{k_c})$. We now define $\mathbf{hcf}_{\chi,c}(\mathbf{H})$ as $hcf(d_1,\ldots,d_{k_c})$.

Given integers $a_1,\ldots,a_k$ with $hcf(a_1,\ldots,a_k)=d$, it is well known that there are integers $b_1,\ldots,b_k$ such that $a_1b_1+\cdots+a_kb_k=d$. They are called the *Bézout numbers* of $a_1,\ldots,a_k$. We need the following fact on the Bézout numbers.

Fact 1.6. For any $k\geq 2$ positive integers $a_1,\ldots,a_k$ with $hcf(a_1,\ldots,a_k)=d$, we may find the Bézout numbers $b_1,\ldots,b_k$ such that $\max_{1\leq i\leq k}|b_i|\leq\max_{1\leq i\leq k}a_i/d$.

Proof. It suffices to prove the fact for $d=1$. In fact, if $hcf(a_1,\ldots,a_k)=d>1$, then $hcf(a'_1,\ldots,a'_k)=1$, where $a'_i=a_i/d$. If $b_1,\ldots,b_k$ are the Bézout numbers of $a'_1,\ldots,a'_k$ such that $\max_{1\leq i\leq k}|b_i|\leq\max_{1\leq i\leq k}a'_i$, then they are the Bézout numbers of $a_1,\ldots,a_k$ with $\max_{1\leq i\leq k}|b_i|\leq\max_{1\leq i\leq k}a_i/d$.

Suppose that $hcf(a_1,\ldots,a_k)=1$. We first show that there exist the Bézout numbers $b_1,\ldots,b_k$ such that $\max_{2\leq i\leq k}|b_i|\leq a_1$. In fact, let $b_1,\ldots,b_k$ be the Bézout numbers with the minimum $\sum_{i=2}^k|b_i|$. We claim that $|b_i|\leq a_1$ for all $i>1$. Suppose instead, there exists $i>1$ such that $|b_i|>a_1$. If $b_i>0$, then define $b'_i=b_i-a_1$ and $b'_1=b_1+a_i$; otherwise, let $b'_i=b_i+a_1$ and $b'_1=b_1-a_i$. Let $b'_j=b_j$ for $j>1$ and $j\neq i$. Then $a_1b'_1+\cdots+a_kb'_k=a_1b_1+a_2b_2+\cdots+a_kb_k=1$. Since $\sum_{i=2}^k|b'_i|=\sum_{i=2}^k|b_i|-a_1<\sum_{i=2}^k|b_i|$, we obtain a contradiction.

Next, among all the Bézout numbers $b_1,\ldots,b_k$ satisfying $\max_{2\leq i\leq k}|b_i|\leq a_1$, we assume that $b_1,\ldots,b_k$ have the minimum $|b_1|$. We claim that this $|b_1|\leq\max_{2\leq i\leq k}a_i$ and consequently $b_1,\ldots,b_k$ are the desired Bézout numbers. Suppose instead that $|b_1|>\max_{2\leq i\leq k}a_i$. Since $a_1b_1+\cdots+a_kb_k=1$ and $|b_1|>1$, there exists $i>1$ such that b_i has the opposite sign of b_1. Let $b'_1=b_1-a_i$ and $b'_i=b_i+a_1$ if $b_1>0$ (thus $b_i<0$); otherwise, let $b'_1=b_1+a_i$ and $b'_i=b_i-a_1$. Let $b'_j=b_j$ for $j>1$ and $j\neq i$. Then $a_1b'_1+\cdots+a_kb'_k=a_1b_1+\cdots+a_kb_k=1$ and $b'_1,\ldots,b'_k$ are the Bézout numbers with $|b'_1|<|b_1|$ and $|b'_j|\leq a_1$ for $j>1$, a contradiction. ■

We call the Bézout numbers with minimum $\max_{1\leq i\leq k}|b_i|$ the *smallest Bézout numbers*.

Definition 1.7. Let H be a bipartite graph with connected components $C_1,\ldots,C_{k_c}$. Suppose that $C_i=C_i[U_i,W_i]$ with $|W_i|\geq|U_i|$. Let $c_i=|W_i|+|U_i|$ and $d_i=|W_i|-|U_i|$.

Recall that $hcf_c(H) = hcf(c_1, \ldots, c_{k_c})$ and $hcf_{\chi,c}(H) = hcf(d_1, \ldots, d_{k_c})$.

1. We define $\zeta(H) = \max_{1 \le i \le k_c} |\zeta_i|$, where $\zeta_1, \ldots, \zeta_{k_c}$ are the smallest Bézout numbers of $c_1, \ldots, c_{k_c}$.
2. We define $\beta(H) = \max_{1 \le i \le k_c} |\beta_i|$, where $\beta_1, \ldots, \beta_{k_c}$ are the smallest Bézout numbers of $d_1, \ldots, d_{k_c}$.

Given H as in Definition 1.7, Fact 1.6 implies that

$$\zeta(H) \le \max_{1 \le i \le k_c} c_i \le h \quad \text{and} \quad \beta(H) \le \max_{1 \le i \le k_c} d_i \le w - u. \tag{1}$$

2. PROOF OF PROPOSITION 1.3

We first observe connections among $hcf_c(H)$, $hcf_\chi(H)$, and $hcf_{\chi,c}(H)$.

Lemma 2.1. *Let H be any bipartite graph.*

1. *Then $hcf_{\chi,c}(H) \le hcf_\chi(H) \le 2 \cdot hcf_{\chi,c}(H)$.*
2. *If $hcf_{\chi,c}(H) = 2$, then $hcf_c(H) \ge 2$.*
3. *Suppose $hcf_c(H) = 1$. Then $hcf_\chi(H) \le 2$ if and only if $hcf_{\chi,c}(H) = 1$.*

Proof. Suppose that H has k_c connected components $C_1[U_1, W_1], \ldots, C_{k_c}[U_{k_c}, W_{k_c}]$. Let $c_i = |C_i|$ and $d_i = |W_i| - |U_i|$.

Part 1. We have $hcf_\chi(H) = hcf(A)$, where $A = \{\sum_{i=1}^{k_c} e_i d_i : e_i \in \{-1, 1\}\}$ is the set of all combinations of adding and subtracting $d_1, \ldots, d_{k_c}$. Therefore, it suffices to show that

$$hcf(d_1, \ldots, d_{k_c}) \le hcf(A) \le 2 \cdot hcf(d_1, \ldots, d_{k_c}).$$

In fact, letting $d = hcf(d_1, \ldots, d_{k_c})$ and $q = hcf(A)$, we have $d \le q$ because d divides every element of A. On the other hand, for any i, q divides $d_1 + \cdots + d_{k_c}$ and $d_1 + \cdots + d_{i-1} - d_i + d_{i+1} + \cdots + d_{k_c}$ and thus q divides $2d_i$. Therefore, $q \le hcf(2d_1, \ldots, 2d_{k_c}) = 2d$.

Part 2. Suppose that $hcf_{\chi,c}(H) = 2$. Then for each component C_i of H, d_i is even. This means $|U_i|$ and $|W_i|$ have the same parity and c_i is even for all i. This implies that $hcf_c(H) \ge 2$.

Part 3. If $hcf_\chi(H) \le 2$, then by Part 1, $hcf_{\chi,c}(H) \le 2$. If $hcf_{\chi,c}(H) = 2$, then by Part 2, $hcf_c(H) \ge 2$ contradicting our assumption. Therefore $hcf_{\chi,c}(H) = 1$. On the other hand, if $hcf_{\chi,c}(H) = 1$, then $hcf_\chi(H) \le 2$ directly follows from Part 1. ■

We now prove Proposition 1.3 by using Lemma 2.1 and four constructions.

Proof of Proposition 1.3. The proof consists of four (mutually disjoint) cases. The first three cases together prove the existence of a graph G with $\delta(G) = \lceil n/2 \rceil - 1$ but containing no H-factor when H is in Class 2. The last case provides a graph G with $\delta(G) = [1 - (1/\chi_{cr}(H))]n - 1$ but containing no H-factor when H is in Class 1.

Case 1. $hcf_c(H) \ge 3$. Let $G = K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor + 1} \cup K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil - 1}$. Since $hcf_c(H) \ge 3$, and any component of H must fit entirely into one of the two connected components of G, we can deduce the following. The size of the components of G differs by 2; but the size of the components of H differs by multiples of $hcf_c(H)$ which is at least 3. Thus, there is no way to arrange the components nor the copies of H to even out the sizes of the components of G. So G contains no H-factor.

Case 2. $hcf_c(H)=2$. Then each component of H has an even size. If n is odd, let $G=K_{\lceil n/2\rceil,\lfloor n/2\rfloor}\cup K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$. If n is even, let $G=K_{n/2,n/2+1}\cup K_{n/2,n/2-1}$. In either case, since every component of G has an odd size, G does not contain an H-factor.

Case 3. $hcf_c(H)=1$ and $hcf_\chi(H)\geq 3$. Let $G=K_{\lfloor n/2\rfloor+1,\lceil n/2\rceil-1}\cup K_{\lceil n/2\rceil-1,\lfloor n/2\rfloor+1}$. It is an immediate consequence of Lemma 2.1 that if $hcf_\chi(H)\geq 3$ and $hcf_c(H)=1$, then $hcf_{\chi,c}(H)\geq 3$. (Note that this does not imply $hcf_{\chi,c}(H)\geq hcf_\chi(H)$.) Now, the sizes of the color classes of the connected components of G differ by 1 or 2. Since $hcf_{\chi,c}(H)\geq 3$, we can only adjust the relative sizes of the color classes of the connected components of G by multiples of $hcf_{\chi,c}(H)$; so we can never get an H-factor.

Case 4. $hcf_c(H)=1$ and $hcf_\chi(H)\leq 2$, namely, H is in Class 1. Recall that $|H|=h$, $u=\sigma(H)$, $w=h-\sigma(H)$, and $1-1/\chi_{cr}(H)=u/h$. Let $G=K_{nu/h-1,nw/h+1}\cup K_{nw/h+1,nu/h-1}$. Then $\delta(G)=[1-(1/\chi_{cr}(H))]n-1$. Let $C_1,C_2,\dots,C_{k_c}$ be the components of H. By contradiction, suppose G has an H-factor. The color class of G with size $\sigma(G)$ thus contains one color class from each of the mk_c packed components of H. Thus

$$\sigma(G)\geq m\sum_{i=1}^{k_c}\sigma(C_i)=mu.$$

However, it is easy to see that $\sigma(G)=mu-2$ by simply placing the 2 components of size $nu/h-1=mu/2-1$ in the same color class. This is a contradiction. So G contains no H-factor. ■

3. REGULARITY LEMMA AND OTHER TOOLS

The Regularity Lemma [19] and the Blow-up Lemma [10] are the backbone of our proof. They allow us to gain convenient structural properties from an arbitrary graph G. Before stating the lemmas, we define ε-regularity, and (ε,δ)-super-regularity.

Definition 3.1. Let $\varepsilon,\delta>0$. Let G be a graph with disjoint vertex sets X and Y.

(1) We say the pair (X,Y) is ε-regular if for every $A\subseteq X$ and $B\subseteq Y$ satisfying $|A|>\varepsilon|X|$, $|B|>\varepsilon|Y|$ we have $|d(A,B)-d(X,Y)|<\varepsilon$.
(2) The pair (X,Y) is (ε,δ)-super-regular if (X,Y) is ε-regular and $\deg(x,Y)>\delta|Y|$ for every $x\in X$ and $\deg(y,X)>\delta|X|$ for every $y\in Y$.

The next two lemmas follow from the definition of ε-regularity easily; their proofs can be found in the survey [12].

Lemma 3.2 (Slicing Lemma). *Let $\varepsilon,d>0$ be constants. Let (X,Y) be an ε-regular pair with density d. For any $\gamma>\varepsilon$, if $X'\subset X, Y'\subset Y$ and $|X'|\geq\gamma|X|, |Y'|\geq\gamma|Y|$, then (X',Y') is an ε'-regular pair with density d' where $|d-d'|<\varepsilon$ and $\varepsilon'=\max\{2\varepsilon,\varepsilon/\gamma\}$.*

Lemma 3.3 (Embedding Lemma). *Let $d\gg\varepsilon>0$. If (X,Y) is an ε-regular pair with density d, then for any positive integers a,b, there exists an n_0 such that if $|X|,|Y|\geq n_0$, then $K_{a,b}\subset(X,Y)$.*

Now we are ready to state the bipartite form of Szemerédi's Regularity Lemma (see [12] for various forms of the Regularity Lemma).

Lemma 3.4 (Regularity Lemma—Bipartite form). *For every $\varepsilon>0$, there exists an $M\in\mathbb{R}^+$ such that if $G=(X,Y;E)$ is any bipartite graph with $|X|=|Y|=n$, and $d\in[0,1]$ is any real number, then there is a partition of X into clusters $X_0,X_1,\dots,X_k$, a partition of Y into $Y_0,Y_1,\dots,Y_k$, and a spanning subgraph $G'=(X,Y;E')$ with the following properties*:

- $k\le M$,
- $|X_0|=|Y_0|\le\varepsilon n$,
- $|X_i|=|Y_j|\le\varepsilon n$ *for all* $1\le i,j\le k$,
- $\deg_{G'}(v)>\deg_G(v)-(d+\varepsilon)n$ *for all* $v\notin X_0\cup Y_0$,
- *All pairs* (X_i,Y_j), $1\le i,j\le k$, *are* ε*-regular in* G', *each with density either* 0 *or greater than* d.

The Blow-up Lemma is very useful for graph tiling, especially when combined with the Regularity Lemma as it essentially says that, when embedding a graph of bounded maximum degree, an (ε,δ)-super-regular pair behaves like a complete bipartite graph. We only need the bipartite form of this lemma.

Lemma 3.5 (Blow-up Lemma—Bipartite form). *For every $\delta,\Delta>0$, there exists an $\varepsilon>0$ such that the following holds. Let (X,Y) be an (ε,δ)-super-regular pair. If a bipartite graph H with $\Delta(H)\le\Delta$ can be embedded in $K_{|X|,|Y|}$, then H can be embedded in (X,Y).*

We now give a sufficient condition for a complete bipartite graph to contain an H-factor.

Lemma 3.6. *Let H be a bipartite graph on h vertices such that $hcf_{\chi,c}(H)=1$. Suppose that $\beta=\beta(H)$, $u=\sigma(H)$, and $w=h-u$. Let $G=K_{mu+t,mw-t}$ such that $t=q(w-u)+r$ and $r<w-u$ for non-negative integers m,t,q,r. If $m\ge r\beta+q$ and $q\ge r\beta$, then G contains an H-factor.*

Proof. $K_{mu,mw}$ has a natural H-factor with all copies of H having their smallest color classes on one side and the largest color classes on the other side. We will show how to transform this into an H-factor of G.

First, since $m\ge q$ we can take q copies of H and swap their sides (here swapping means switching the sides of the color classes). This now results in a spanning subgraph of $K_{mu+t-r,mw-t+r}$. Let us call the part of this tiling that was not swapped as G_1 and the part that was swapped as G_2. Since $hcf_{\chi,c}(H)=1$, there exist integers $\beta_1,\dots,\beta_{k_c}$ as in Definition 1.7. Let us say that $\beta_1,\dots,\beta_i$ are nonnegative and $\beta_{i+1},\dots,\beta_{k_c}$ are all negative. Now, in G_1 swap $r\beta_j$ copies of C_j for all $j=1,\dots,i$. Note that since $m-q\ge r\beta$, we have enough copies of each component to perform the aforementioned swaps. In G_2, swap $-r\beta_j$ copies of C_j for all $j=i+1,\dots,k_c$. We can perform this swap because $q\ge r\beta$. So, the left side gains

$$r=r\beta_1d_1+\cdots+r\beta_id_i+r\beta_{i+1}d_{i+1}+\cdots+r\beta_{k_c}d_{k_c}$$

vertices. Similarly, the right side loses r vertices, and we now have a spanning subgraph of $K_{mu+t,mw-t}=G$. ■

We will use the following corollary of Lemma 3.6 in Section 4A, which is slightly stronger than the bipartite version of Lemma 12 in [13].

Corollary 3.7. *Let H be a bipartite graph in Class 1 with h vertices. Let $u=\sigma(H)$, $w=h-\sigma(H)$, and $0<\gamma<(w-u)/u$. Suppose that $G[X,Y]$ is a complete bipartite graph on mh vertices for some sufficiently large integer m such that $(1+\gamma)(u/w)\leq|X|/|Y|\leq 1$. Then G contains an H-factor.*

Proof. We will prove that G satisfies the conditions of Lemma 3.6 in order to get an H-factor. First, since $|X|+|Y|$ is divisible by h, we may write $G=K_{mu+t,mw-t}$ where $m=(|X|+|Y|)/h$ and t is some integer. Further, write $t=q(w-u)+r$ for some integers q,r such that $0\leq r<w-u$. Let $\beta=\beta(H)$ and

$$m\geq\frac{(w-u)^2(h+u\gamma)\beta}{uw\gamma}. \tag{2}$$

We must prove $m\geq r\beta+q$ and $q\geq r\beta$. Since $q=\lfloor t/(w-u)\rfloor\leq t/(w-u)$, it is sufficient to prove that (i) $m\geq r\beta+t/(w-u)$ and (ii) $t/(w-u)\geq r\beta$. Since

$$\frac{|X|}{|Y|}=\frac{mu+t}{mw-t}\geq(1+\gamma)\frac{u}{w},$$

we have that $th+tu\gamma\geq mwu\gamma$ or $t\geq(uw/(h+u\gamma))m\gamma$. Now, by (2), we have $(uw/(h+u\gamma))m\gamma\geq(w-u)^2\beta$, which implies that $t\geq(w-u)^2\beta>(w-u)r\beta$ thus proving (ii). On the other hand, $|X|/|Y|\leq 1$ implies that $mu+t\leq mw-t$, or $2t\leq m(w-u)$. Since $t\geq(w-u)r\beta$, we have $m(w-u)\geq(w-u)r\beta+t$, which gives (i). ∎

4. PROOF OF THEOREM 1.4

Let H be a bipartite graph on h vertices with positive integers $u=\sigma(H)$ and $w=h-u$. We assume that $u<w$, otherwise $\chi_{cr}(H)=2$ and Theorem 1.2 gives the proof. We thus have $w\geq 2$, and $h\geq 3$.

The proof of our main theorem consists of two parts: the nonextremal case and the extremal case. Roughly speaking, a balanced bipartite graph with $2n=mh$ vertices is in the extremal case if it is relatively similar to $K_{nu/h-1,nw/h+1}\cup K_{nw/h+1,nu/h-1}$, the construction we gave in Case 4 of the proof of Proposition 1.3.

A. Nonextremal Case

In this subsection we prove the following theorem, which covers the nonextremal case.

Theorem 4.1. *Let H be a bipartite graph on h vertices such that H is in Class 1. Let $u=\sigma(H)$ and $w=h-\sigma(H)$. For every $\alpha>0$ there exist $\gamma>0$ and a positive integer m_0 such that if $m\geq m_0$ and $G[X,Y]$ is a balanced bipartite graph on $2n=mh$ vertices with*

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

then G either contains an H-factor or there exist sets $A\subset X$, $B\subset Y$ such that $|A|=|B|=\lfloor wn/h\rfloor$ and $d(A,B)\leq\alpha$.

We say that a bipartite graph $G[X,Y]$ is in the *extremal case* with parameter α if there exist sets $A\subset X$, $B\subset Y$ such that $|A|=|B|=\lfloor wn/h\rfloor$ and $d(A,B)\le\alpha$.

The proof of Theorem 4.1 is divided into two lemmas. The first lemma puts most vertices of G into super-regular pairs such that the ratio of the sizes between the pairs is slightly larger than u/w. Having a ratio slightly larger than u/w allows us to remove a small amount of vertices from the super-regular pair, yet its remaining vertices can be tiled by H perfectly by applying Corollary 3.7 and Lemma 3.5. We make this precise by the following definition.

Definition 4.2. Given $0<\varepsilon<d<1$ and positive integers p,q,N, let $G[X,Y]$ be a balanced bipartite graph. A partition of $V(G)=X_0\cup Y_0\cup P_1\cup Q_1\cup\cdots P_k\cup Q_k$ is called an almost (ε,d,p,q,N)-cover of G if

- $X_0\subset X$, $Y_0\subset Y$ and $|X_0|,|Y_0|\le\varepsilon n$,
- For all i, $|P_i|/p=|Q_i|/q\ge N$ and either $P_i\subset X$ and $Q_i\subset Y$, or $P_i\subset Y$ and $Q_i\subset X$,
- For all i, (P_i,Q_i) is (ε,d)-super-regular.

Lemma 4.3. *Let $w>u$ be positive integers and $h=w+u$. For every $\alpha>0$ and integer N, there exists a positive integer n_0 and constants $0<\varepsilon\ll d\ll\gamma\ll\alpha$ such that if $G[X,Y]$ is a balanced bipartite graph on $2n$ vertices with $n\ge n_0$, and $\delta(G)\ge(u/h-\gamma)n$, then either G is in the extremal case with parameter α or G contains an almost (ε,d,p,q,N)-cover, where $p=w+u/\gamma$ and $q=w/\gamma$ are integers.*

There are two reasons why we cannot immediately apply Corollary 3.7 to each (P_i,Q_i) in the cover. First, we need to get rid of the *exceptional sets* X_0 and Y_0. Second, we may not have $|P_i|+|Q_i|$ divisible by h. Achieving these two additional properties is the content of Lemma 4.4, in which we also assume H is in Class 1. By definition, if H is in Class 1, then $hcf_c(H)=1$ and $hcf_\chi(H)\le 2$. By Part 3 of Lemma 2.1, this implies that $hcf_c(H)=1$ and $hcf_{\chi,c}(H)=1$. The condition of $hcf_c(H)=1$ is used for achieving the divisibility of $|P_i|+|Q_i|$. The condition of $hcf_{\chi,c}(H)=1$ is needed for Corollary 3.7.

Lemma 4.4. *Let H be a bipartite graph with $hcf_c(H)=1$ and $hcf_{\chi,c}(H)=1$. Let $u=\sigma(H)$ and $w=h-u$. Let G be a balanced bipartite graph on $2n=mh$ vertices such that $\delta(G)\ge(1-1/\chi_{cr}(H)-\gamma)n$. Suppose that G contains an almost (ε,d,p,q,N)-cover for some positive $\varepsilon\ll d\ll\gamma\ll 1$, integers p, q satisfying $p/q=(1+\gamma)u/w$, and sufficiently large N. Then G contains an H-factor.*

Proof of Lemma 4.3. In the proof, we will omit the floor function when it does not affect our calculations. Assume n is large. Without loss of generality, assume $\alpha\ll 1$. We choose parameters ε_0, d_0, γ so that they satisfy the following relations:

$$\varepsilon_0\ll d_0\ll\gamma=\frac{1}{z}\ll\alpha \tag{3}$$

for some integer z. Let $p=uz+w$ and $q=wz$ be two integers. Then p and q have the following property:

$$\frac{u}{w}<\frac{p}{q}=\frac{u}{w}+\gamma\le 1. \tag{4}$$

We apply the Regularity Lemma (Lemma 3.4) with parameters ε_0 and d_0 to G. We obtain an integer $k_0\le M(\varepsilon_0)$ and a spanning subgraph G' consisting of *clusters*

$X_1, Y_1, \ldots, X_{k_0}, Y_{k_0}$ of size $N_0 \leq \varepsilon_0 n$ and *exceptional sets* X_0 and Y_0 of size at most $\varepsilon_0 n$. Every pair of clusters (X_i, Y_j) is ε_0-regular, with density either 0 or greater than d_0. The degrees of the vertices in G' are very close to their degrees in G:

$$\deg_{G'}(v) > \deg_G(v) - (d_0 + \varepsilon_0)n = \left(\frac{u}{h} - \gamma - d_0 - \varepsilon_0\right) n.$$

Let R be the reduced graph of G' where each vertex corresponds to a cluster in $G' - (X_0 \cup Y_0)$, and we say there is an edge between X_i and Y_j if the density $d(X_i, Y_j) > d_0$, written as $X_i \sim Y_j$. Note that we use the same notation for a cluster in G' and a vertex in R; we clearly say whether it is a cluster of G' or a vertex of R when this is not clear from the context. In order to bound $\delta(R)$, we consider an arbitrary X_i and an arbitrary vertex $x \in X_i$. We have

$$\left(\frac{u}{h} - \gamma - d_0 - 2\varepsilon_0\right) n \leq \deg_{G'}(x) - |Y_0| \leq \sum_{Y_j \sim X_i} |Y_j| = \deg_R(X_i) N_0. \tag{5}$$

Using (3) and $k_0 N_0 \leq n$, we derive that $\deg_R(X_i) \geq (u/h - 2\gamma)k_0$. The same holds for any cluster in Y. Thus we have

$$\delta(R) \geq \left(\frac{u}{h} - 2\gamma\right) k_0. \tag{6}$$

We need a simple fact on the size of a maximum matching in bipartite graphs; for completeness, we include a proof.

Fact 4.5. If $G[X, Y]$ is a bipartite graph with minimum degree δ such that $|X| \leq |Y|$, then G has a matching of size at least $\min\{2\delta, |X|\}$.

Proof. Let $M = \{x_1 y_1, \ldots, x_t y_t\}$ be a maximum matching in G. Assume $t < |X|$. Then, there exists a vertex $x \in X - \{x_1, \ldots, x_t\}$. Since $|Y| \geq |X|$, there also exists $y \in Y - \{y_1, \ldots, y_t\}$. Let $I = \{1 \leq i \leq t : y_i \in \Gamma(x)\}$ and $J = \{1 \leq j \leq t : x_j \in \Gamma(y)\}$. Then $|I|, |J| \geq \delta$. Since M is a maximum matching, we have $I \cap J = \emptyset$ and $|I|, |J| \geq \delta$ (otherwise we may extend the matching). This implies that $t \geq |I| + |J| \geq 2\delta$. ■

Let M be a maximum matching in the reduced graph R. Since $2u < w + u = h$, by Fact 4.5, we have $|M| \geq 2\delta(R) \geq 2(u/h - 2\gamma)k_0$. Denote by U_1 and U_2 the set of unmatched clusters from X and Y, respectively. Then $|U_1|, |U_2| \leq ((w-u)/h + 2\gamma)k_0$.

The next part of the proof will be decomposing clusters to get pairs of ratio p/q. We first prove that we can find two disjoint subgraphs P_1 and P_2 of R that satisfy the following properties (Fig. 1). The subgraph P_1 will have vertex sets U_1 and $\Gamma(U_1) := \bigcup_{X_i \in U_1} \Gamma(X_i)$. Moreover, for any vertex $X_i \in U_1$, $\deg_{P_1}(X_i) = p$, and for any vertex

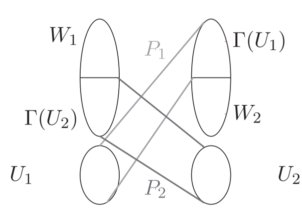

FIGURE 1. Finding P_1 and P_2.

$Y_j \in \Gamma(U_1)$, $\deg_{P_1}(Y_j) \leq q-p$. The subgraph P_2 will have vertex sets U_2 and $\Gamma(U_2)$; for any $Y_j \in U_2$, $\deg_{P_2}(Y_j) = p$, and for any $X_i \in \Gamma(U_2)$, $\deg_{P_2}(X_i) \leq q-p$. Note that since M is maximal, $\Gamma(U_1), \Gamma(U_2) \subset V(M)$ and no edge of M has one end in $\Gamma(U_1)$ and the other end in $\Gamma(U_2)$.

Let $\alpha' = \alpha/12$. We prove the following claim:

Claim 4.6.

(a) *If* $|U_1|, |U_2| \leq ((w-u)/h - \alpha')k_0$, *then there exist two disjoint subgraphs* P_1 *and* P_2 *with the above properties.*
(b) *If* $|U_1|, |U_2| > ((w-u)/h - \alpha')k_0$, *then* G *is in the extremal case with parameter* α.

Proof. We first prove (a). We will only prove that we can find P_1 because the proof for P_2 is the same. We will find P_1 by the greedy algorithm. Arbitrarily order the vertices in U_1. For each vertex in U_1, we find p neighbors in $\Gamma(U_1)$ with the restriction that we cannot choose any vertex in $\Gamma(U_1)$ more than $q-p$ times. When considering the ith vertex in U_1, suppose that there are t vertices in $\Gamma(U_1)$ that have been chosen $q-p$ times. Since $t \leq (i-1)p/(q-p) < |U_1|p/(q-p)$, it suffices to show that $\delta(R) \geq p + |U_1|p/(q-p)$. Using (6) and $|U_1| \leq ((w-u)/h - \alpha')k_0$, we have

$$\delta(R) - \frac{p}{q-p}|U_1| \geq \left(\frac{u}{h} - 2\gamma\right)k_0 - \frac{p}{q-p}\left(\frac{w-u}{h} - \alpha'\right)k_0.$$

From the Regularity Lemma, we know that $k_0 \geq 1/(2\varepsilon_0)$. Thus, it suffices to show that

$$\phi := \left(\frac{u}{h} - 2\gamma\right) - \left(\frac{w-u}{h} - \alpha'\right)\frac{p}{q-p} \geq 2\varepsilon_0 p.$$

In fact, the definition of p, q and the assumption $z \geq 2w/(w-u)$, which follows from $\gamma \ll 1$, give that

$$\frac{p}{q-p} - \frac{u}{w-u} = \frac{uz+w}{(w-u)z-w} - \frac{u}{w-u} = \frac{w^2}{((w-u)z-w)(w-u)} \leq \frac{2w^2}{(w-u)^2 z}.$$

By using (3), we obtain that

$$\phi \geq \left(\frac{u}{h} - 2\gamma\right) - \left(\frac{w-u}{h} - \alpha'\right)\left(\frac{u}{w-u} + \frac{2w^2}{(w-u)^2 z}\right) > -2\gamma - \frac{2w^2\gamma}{h(w-u)} + \frac{u}{w-u}\alpha' \geq 2\varepsilon_0 p.$$

Thus, the greedy algorithm is sufficient to find the subgraphs P_1 and P_2.

Now, we prove (b). We assume $|U_1|, |U_2| > ((w-u)/h - \alpha')k_0$. Let W_i be the neighbors of $\Gamma(U_i)$ in M for $i = 1, 2$. It is easy to see that the following four quantities must all be equal to 0 or we can extend the matching in G:

$$e(U_1, U_2) = e(U_1, W_2) = e(U_2, W_1) = e(W_1, W_2) = 0.$$

For example, if there exists an edge X_iY_j between W_1 and W_2, then we can extend the matching as follows. Let Y_i denote the matched neighbor of X_i, X_j denote the matched neighbor of Y_j, $X_{i'}$ denote a vertex in U_1 adjacent to Y_i, and $Y_{j'}$ denote a vertex in U_2 adjacent to X_j. Then we can enlarge the matching by replacing X_iY_i, X_jY_j by $X_{i'}Y_i$, X_iY_j, and $X_jY_{j'}$.

Now, letting $\mathcal{A}=U_1\cup W_1$, and $\mathcal{B}=U_2\cup W_2$, then $e_R(\mathcal{A},\mathcal{B})=0$. Moreover,

$$|\mathcal{A}|=|U_1|+|W_1|\geq|U_1|+\delta(R)\geq\left(\frac{w-u}{h}-\alpha'\right)k_0+\left(\frac{u}{h}-2\gamma\right)k_0=\left(\frac{w}{h}-\alpha'-2\gamma\right)k_0.$$

Let A' and B' be the sets of vertices of G in all the clusters of $\mathcal{A}$ and of $\mathcal{B}$, respectively. Since $k_0N_0\geq(1-\varepsilon_0)n$ and $\varepsilon_0\ll\gamma\ll\alpha'$, we derive that $|A'|\geq(w/h-2\alpha')n$. The same holds for $|B'|$. Since $e_{G'}(A',B')=0$, for any subset $S\subseteq A'$, we have

$$e_G(S,B')\leq e_{G'}(A',B')+|S|(d_0+\varepsilon_0)n\leq 2d_0n|S|.$$

Now, by adding at most $2\alpha' n$ vertices to A' and B', we get two sets A,B of size exactly $\lfloor wn/h\rfloor$; when $|A'|$ or $|B'|$ is greater than $\lfloor wn/h\rfloor$, we simply take a subset of size $\lfloor wn/h\rfloor$. Since each of the new vertices in A (or B) might be adjacent to all the vertices in B (or A), we have

$$d(A,B)\leq\frac{e_G(A'\cap A,B')+2\alpha' n|B|+2\alpha' n|A|}{|A||B|}=\frac{2d_0n+4\alpha' n}{|B|}\leq 12\alpha'=\alpha.$$

So, we are in the extremal case with parameter α. ■

We assume that G is not in the extremal case with parameter α, and thus Claim 4.6(a) holds. Now we use the structures of P_1 and P_2 to guide us to break up clusters. In order to evenly divide a cluster into small pieces, we ensure the size of all clusters is divisible by $pq(q^2-p^2)$ by moving at most $pq(q^2-p^2)-1$ vertices from each cluster to the exceptional set. This increases $|X_0|$ and $|Y_0|$ by a constant, less than $pq(q^2-p^2)k_0$. For simplicity, we still use N_0 for the size of the clusters.

Now we only give the details on how to handle the clusters in $U_1\cup\Gamma(U_1)$. We evenly decompose every cluster $X_i\in U_1$ into p subclusters and adjoin each subcluster to a unique neighbor of X_i in P_1 (Fig. 2). Since $\deg_{P_1}(X_i)=p$ for each $X_i\in U_1$, this is possible. However, we do not adjoin each subcluster of X_i to the entire cluster. Instead, we adjoin it to a subcluster of size N_0/q. Thus, the ratio between two adjoining subclusters is p/q.

Let $Y_j\subset Y$ be a cluster covered by the matching M. We know that Y_j has degree $i\leq q-p$ in P_1 ($i=0$ when $Y_j\notin\Gamma(U_1)$). In total, iN_0/q vertices of Y_j are already us ed. We match up the remaining N_0-iN_0/q vertices in Y_j with its neighbor X_j in M forming at most 3 cluster pairs of ratio p/q as follows. First take $iN_0/(q-p)$ vertices

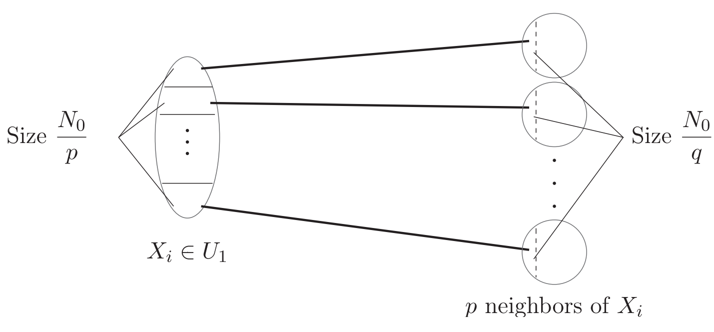

FIGURE 2. Decomposing one cluster in U_1.

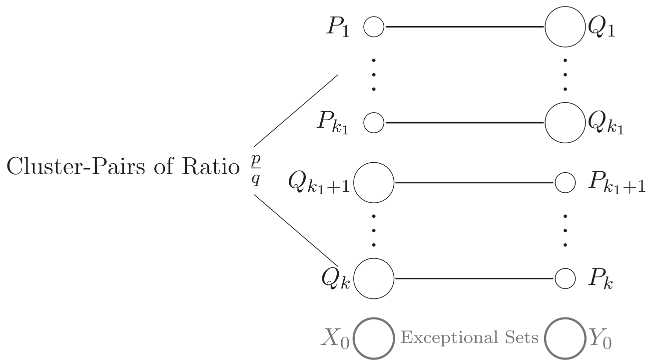

FIGURE 3. Graph G' after decomposition.

from X_j and match them with $ipN_0/(q(q-p))$ vertices from Y_j. This makes a cluster pair with ratio p/q. Now, the number of remaining vertices in X_j is $N_0 - iN_0/(q-p)$, while the number of remaining vertices in Y_j is $N_0 - iN_0/q - ipN_0/(q(q-p))$, also equal to $N_0 - iN_0/(q-p)$. Finally, we make two more cluster pairs with ratio p/q by pairing together $(N_0 - iN_0/(q-p))(p/(q+p))$ vertices from one cluster with $(N_0 - iN_0/(q-p))(q/(q+p))$ from the other.

In summary, we broke all the clusters into subclusters and group them into pairs with sizes

$$\left\{\frac{N_0}{p}, \frac{N_0}{q}\right\}, \quad \left\{\frac{iN_0}{q-p}, \frac{ipN_0}{q(q-p)}\right\}, \quad \left\{\frac{q-p-i}{q-p}\frac{q}{p+q}N_0, \frac{q-p-i}{q-p}\frac{p}{p+q}N_0\right\}, \tag{7}$$

where $0 \le i \le q-p$.

Let $\gamma' = \min\{1/q, p/(q^2-p^2)\}$ (then $\gamma' > d > \varepsilon_0$ by (3)). The size of any subcluster is at least $\gamma' N_0$, which is larger than the given integer N because $N_0 \ge (1-2\varepsilon_0)(n/k_0)$ is sufficiently large. Let $(P_1, Q_1), \ldots, (P_k, Q_k)$ denote these cluster pairs. After relabeling, we may assume that the first k_1 of them have P_i in X and Q_i in Y (see Fig. 3). We have $k \le 2pk_0$ because each cluster in $U_1 \cup U_2$ generates at most p pairs, while each cluster covered by M generates at most 3 pairs, and $p \ge 3$. The ε_0-regularity between the original clusters implies that all (P_i, Q_i) have density within ε_0 of d_0. Lemma 3.2 further guarantees that all (P_i, Q_i) are ε_1-regular with $\varepsilon_1 = \varepsilon_0/\gamma'$.

In order to obtain super-regularity for each (P_i, Q_i), we now remove vertices with small degree into the opposite cluster to the exceptional sets X_0, Y_0. Suppose that, for example, $P_i \subset X$ and $Q_i \subset Y$. We move any vertex $x \in P_i$ such that $\deg(x, Q_i) < (d(P_i, Q_i) - \varepsilon_1)|Q_i|$ to X_0, and any vertex $y \in Q_i$ such that $\deg(y, P_i) < (d(P_i, Q_i) - \varepsilon_1)|P_i|$ to Y_0. The ε_1-regularity between P_i and Q_i guarantees that we move at most $\varepsilon_1|C|$ vertices from each $C \in \{P_i, Q_i\}$. In order to maintain the ratio to be exactly p/q, we may have to move more vertices from P_i to X_0 and from Q_i to Y_i such that, in total, P_i loses at most $p\lceil \varepsilon_1|P_i|/p\rceil \le \varepsilon_1|P_i| + p \le 2\varepsilon_1|P_i|$ vertices while Q_i loses at most $q\lceil \varepsilon_1|Q_i|/q\rceil \le \varepsilon_1|Q_i| + q \le 2\varepsilon_1|Q_i|$ vertices.

We still denote the resulting clusters by P_i and Q_i. Since the original P_i has at least $\gamma' N_0$ vertices, the modified P_i has at least $(1-2\varepsilon_1)\gamma' N_0$ vertices. By Lemma 3.2, the modified (P_i, Q_i) is $2\varepsilon_1$-regular. Since the density between the original P_i and Q_i is at least $d_0 - \varepsilon_0$, the modified (P_i, Q_i) satisfies $\deg(x, Q_i) \ge (d_0 - \varepsilon_0 - 2\varepsilon_1)|Q_i|$ for any vertex $x\in P_i$, and $\deg(y,P_i)\ge(d_0-\varepsilon_0-2\varepsilon_1)|P_i|$ for any vertex $y\in Q_i$. Let $\varepsilon=2\varepsilon_1$ and $d=d_0-\varepsilon_0-2\varepsilon_1$. Then all (current) (P_i,Q_i) are (ε,d)-super-regular.

In total, we moved at most $\sum_C(\varepsilon_1|C|+q)\le\varepsilon_1 n+kq$ vertices to X_0 where the sum ranges over all current clusters contained in X. As a result, $|X_0|\le\varepsilon_0 n+pq(q^2-p^2)k_0+\varepsilon_1 n+kq\le\varepsilon n$. The same holds for $|Y_0|$. ∎

Proof of Lemma 4.4. Let $X_0,Y_0,P_1,Q_1,\ldots,P_k,Q_k$ be the given almost (ε,d,p,q,N)-cover of G. As before, we call X_0,Y_0 exceptional sets, and $P_i,Q_i,i=1,\ldots,k$, clusters. We know that $|X_0|,|Y_0|\le\varepsilon n$, all pairs (P_i,Q_i) are (ε,d)-super-regular with $|P_i|/|Q_i|=p/q=u/w+\gamma$. Our first goal will be to take vertices in $X_0\cup Y_0$ and find disjoint copies of $K_{u,w}$ (a supergraph of H) for each of them. ∎

Claim 4.7. *We may remove $|X_0\cup Y_0|$ disjoint copies of $K_{u,w}$ from G, each of which contains exactly one vertex from $X_0\cup Y_0$, such that each cluster $C\in\{P_i,Q_i\}$ loses at most $(d/3)|C|$ vertices.*

Proof. We say that a vertex v is adjacent to a cluster C (written as $v\sim C$) if $\deg(v,C)|\ge d|C|$. Following an arbitrary order of X_0 and Y_0, we associate each vertex $x\in X_0\cup Y_0$ with a cluster C that x is adjacent to. We also say that x is associated with the cluster pair (P_i,Q_i) if $C\in\{P_i,Q_i\}$. First assume that $C=P_i$. By Lemma 3.2, $(\Gamma(x,P_i),Q_i)$ is ε/d-regular and by Lemma 3.3, $(\Gamma(x,P_i),Q_i)$ contains a copy of $K_{u,w-1}$ with $w-1$ vertices in Q_i. We then remove this copy of $K_{u,w-1}$ together with x (they form a copy of $K_{u,w}$). When $C=Q_i$, we remove a copy of $K_{u-1,w}$ from $(P_i,\Gamma(x,Q_i))$ with $u-1$ vertices in P_i. Together with x, the removed vertices form a copy of $K_{u,w}$.

To ensure that each cluster C loses at most $(d/3)|C|$ vertices, we associate at most $(d/3w)|Q_i|$ vertices of $X_0\cup Y_0$ to any pair (P_i,Q_i). Then Q_i loses at most $(d/3)|Q_i|$ vertices because each associated vertex of $X_0\cup Y_0$ makes Q_i lose at most w vertices. On the other hand, P_i loses at most u vertices for each associated vertex. Since $|Q_i|/w\le|P_i|/u$, P_i loses at most $u(d/3w)|Q_i|\le(d/3)|P_i|$ vertices.

We need to prove that under this restriction, there are enough clusters for all the vertices in the exceptional sets. First, we give a lower bound for $\sum_{x\sim C}|C|$ for all $x\in X_0\cup Y_0$. Fix $x\in X_0$ (the case when $x\in Y_0$ is similar). By the minimum degree condition and the definition of $x\sim C$,

$$\left(\frac{u}{h}-\gamma\right)n\le d_G(x)\le|Y_0|+\sum_{x\sim C}|C|+\sum_{C\subset Y:x\not\sim C}d|C|\le\varepsilon n+dn+\sum_{x\sim C}|C|,$$

which implies that $\sum_{x\sim C}|C|\ge(u/h-2\gamma)n$ by using $\varepsilon\ll d\ll\gamma$. For a cluster $C\in\{P_i,Q_i\}$ with $x\sim C$, if we have associated $(d/3w)|Q_i|\ge(d/3w)|C|$ exceptional vertices with (P_i,Q_i), then we cannot associate x with C. If all the clusters C adjacent to x cannot be used, then the number of exceptional vertices that have been considered is at least

$$\sum_{x\sim C}\frac{d}{3w}|C|\ge\frac{d}{3w}\left(\frac{u}{h}-2\gamma\right)n>2\varepsilon n,$$

a contradiction. ∎

Other than a small number of copies of $K_{u,w}$, the graph G now consists of cluster pairs (P_i,Q_i) with ratio near p/q. In order to apply Corollary 3.7 to these (P_i,Q_i), we want $|P_i|+|Q_i|$ to be divisible by h. We use the fact that $hcf_c(H)=1$ and let $\zeta=\zeta(H)$.

Claim 4.8. *We may remove at most $2\zeta hk$ disjoint copies of H such that each cluster $C\in\{P_i,Q_i\}$ loses at most ζh^2 vertices, and all $|P_i|+|Q_i|$ are divisible by h.*

Proof. Recall that $\sum_{i=1}^{k_c}\zeta_i c_i=1$ and $\zeta=\max_{1\le i\le k_c}|\zeta_i|$, where $c_1,\dots,c_{k_c}$ are the sizes of the components of H. In order to ensure that the size of each cluster pair is divisible by h, we show how to increase or decrease the size of a cluster pair by 1 modulo h. Let G_1 and G_2 denote the subgraphs induced by two cluster pairs (P_i,Q_i) and (P_j,Q_j), respectively. We will decrease the order of G_1 by 1 modulo h and increase the order of G_2 by 1 modulo h. To do this, we remove 2ζ copies of H by selectively choosing where the components of H come from. Since the cluster pairs are regular, we can find these copies of H by Lemma 3.3.

From G_1 we remove $\zeta-\zeta_i$ copies of C_i for $1\le i\le k_c$. Totally, G_1 loses $\sum_{i=1}^{k_c}(\zeta-\zeta_i)c_i=\zeta h-1$ vertices. From G_2 we remove $\zeta+\zeta_i$ copies of C_i for $1\le i\le k_c$. Then G_2 loses $\sum_{i=1}^{k_c}(\zeta+\zeta_i)c_i=\zeta h+1$ vertices. Since it is impossible that all the removed $\zeta h+1$ vertices come from one of P_j and Q_j, each of P_j,Q_j loses at most ζh vertices.

Let r_i be the remainder of $|P_i|+|Q_i| \bmod h$ for $i=1,\dots,k$. Suppose that r_i is the smallest nonzero remainder and r_j is the largest remainder. By applying the procedure above at most $\min\{r_i,h-r_j\}$ times, we either reduce r_i to 0 or enlarge r_j to h. Repeat this process at most $k-1$ times and obtain $r_i\equiv 0 \bmod h$ for all $i=1,\dots,k$ (note that $\sum r_i\equiv 0 \bmod h$ all the time). The total number of the removed copies of H is at most $2\zeta(h-1)(k-1)<2\zeta hk$, and each cluster loses at most $\zeta h(h-1)<\zeta h^2$ vertices.

Pairing (P_i,Q_i) and (P_j,Q_j) together and performing this process until either $r_i\equiv 0 \bmod h$ or $r_j\equiv 0 \bmod h$, it is easy to see that one may apply this procedure totally at most $(h-1)\sum_{i=1}^{k} r_i$ times to ensure that $|P_i|+|Q_i|$ is divisible by h for all $i=1,\dots,k$. ■

Fix $i=1,\dots,k$. Let P_i',Q_i' denote the clusters obtained from P_i,Q_i after applying Claim 4.7 and Claim 4.8. We observe that $|P_i'|,|Q_i'|$ are large and $(1+\gamma/2)(u/w)\le|P_i|/|Q_i|\le 1$. In fact, by Claims 4.7 and 4.8, each cluster C loses at most $d|C|/3+\zeta h^2\le d|C|/2$ vertices, and consequently $|C'|\ge(1-d/2)|C|$. Since $d\ll\gamma\ll 1$, we derive that

$$\left(1+\frac{\gamma}{2}\right)\frac{u}{w}\le\left(1-\frac{d}{2}\right)\left(\frac{u}{w}+\gamma\right)=\frac{\left(1-\dfrac{d}{2}\right)|P_i|}{|Q_i|}\le\frac{|P_i'|}{|Q_i'|}\le\frac{|P_i|}{\left(1-\dfrac{d}{2}\right)|Q_i|}=\frac{\dfrac{u}{w}+\gamma}{1-\dfrac{d}{2}}<1.$$

By Corollary 3.7, the complete bipartite graph $K_{|P_i'|,|Q_i'|}$ contains an H-factor. If we can show that (P_i',Q_i') is super-regular, then the Blow-up Lemma implies that $G[P_i',Q_i']$ also contains an H-factor. In fact, since (P_i,Q_i) is (ε,d)-super-regular, we have $|\Gamma(x,Q_i')|\ge d|Q_i|-d|Q_i|/2\ge d|Q_i|/2$ for all $x\in P_i'$ and similarly $|\Gamma(y,P_i')|\ge d|P_i'|/2$ for all $y\in Q_i'$. By the Slicing Lemma, (P_i',Q_i') is $(2\varepsilon,d/2)$-super-regular.

Note that $V(G)\setminus\bigcup_{i=1}^{k}(P_i\cup Q_i)$ consists of disjoint copies of H. We thus obtain the desired H-factor of G. ■

B. The Extremal Case

We now prove that we can tile G in the extremal case. More precisely, we prove the following theorem:

Theorem 4.9. *Let H be a bipartite graph with H is in Class 1, $u=\sigma(H)$, $w=h-\sigma(H)$, $\zeta=\zeta(H)$, and $\beta=\beta(H)$. Let*

$$c_1(H):=\zeta h^2+\beta(w-u)^2+\frac{h}{2}(w-u)+w. \tag{8}$$

Then, there exist $\alpha>0$ and an integer m_0 such that for any $m\geq m_0$, if $G[X,Y]$ is a balanced, bipartite graph on $2n=mh$ vertices such that (i) G has minimum degree

$$\delta(G)\geq\left(\frac{u}{h}\right)n+c_1(H),$$

and (ii) there are subsets $A\subset X$, $B\subset Y$, where $|A|=|B|=\lfloor wn/h\rfloor$ with $d(A,B)\leq\alpha$, then G contains an H-factor.

By (1) and (8), we derive that $c_1(H)\leq 4h^3$ and thus complete the proof Theorem 1.4.

To prove Theorem 4.9, let us start with a simple corollary of the Blow-up Lemma. Recall that $\delta(X,Y)$ denotes $\min_{x\in X}\deg(x,Y)$.

Lemma 4.10. *For every positive integer Δ, there exists a positive number $\rho=\rho(\Delta)<1$ such that the following holds. For any bipartite graph F, if $\Delta(F)\leq\Delta$ and F can be embedded into $K_{|X|,|Y|}$, then F can be embedded into every bipartite graph $G[X,Y]$ with*

$$\delta(X,Y)\geq(1-\rho)|Y|,\quad \delta(Y,X)\geq(1-\rho)|X|. \tag{9}$$

Proof. We first prove that for any $0<\rho<1$, every bipartite graph $G[X,Y]$ satisfying (9) is $\sqrt{\rho}$-regular. In fact, consider subsets $A\subseteq X$, $B\subseteq Y$ with $|A|=\gamma_1|X|$ and $|B|=\gamma_2|Y|$ for some $\gamma_1,\gamma_2>\sqrt{\rho}$. By (9), we have $\delta(A,Y)\geq|Y|-\rho|Y|$ and consequently $\delta(A,B)\geq|B|-\rho|Y|=(\gamma_2-\rho)|Y|$. The density between A and B satisfies

$$d(A,B)\geq\frac{\delta(A,B)|A|}{|A||B|}\geq\frac{(\gamma_2-\rho)|Y|}{|B|}=\frac{\gamma_2-\rho}{\gamma_2}>1-\frac{\rho}{\sqrt{\rho}}=1-\sqrt{\rho}.$$

Since $1-\sqrt{\rho}<d(A,B)\leq 1$ and in particular, $1-\sqrt{\rho}<d(X,Y)\leq 1$, we have $|d(A,B)-d(X,Y)|<\sqrt{\rho}$.

Now assume that $K_{|X|,|Y|}$ contains a copy of F and let ε be given by the Blow-up Lemma (Lemma 3.5) with $\delta=\frac{1}{2}$ and $\Delta(F)=\Delta$. Let $\rho=\min\{\varepsilon^2,\frac{1}{2}\}$ and $G[X,Y]$ be a bipartite graph satisfying (9). Then G is $(\varepsilon,\frac{1}{2})$-super-regular and thus contains a copy of F. ■

Proof of Theorem 4.9. Recall that $A\subset X$ and $B\subset Y$ are sets of size $\lfloor wn/h\rfloor$ with $d(A,B)\leq\alpha$. Let $A^c=X-A$ and $B^c=Y-B$. Then $|A^c|=|B^c|=\lceil un/h\rceil$.

We define the following subsets:

$$A_1=\{x\in X:\deg(x,B)<\alpha^{\frac{1}{3}}|B|\},\quad B_1=\{y\in Y:\deg(y,A)<\alpha^{\frac{1}{3}}|A|\},$$

$$A_2=\{x\in X:\deg(x,B)>(1-\alpha^{\frac{1}{3}})|B|\},\quad B_2=\{y\in Y:\deg(y,A)>(1-\alpha^{\frac{1}{3}})|A|\},$$

$$A_0=X-A_1-A_2,\quad B_0=Y-B_1-B_2.$$

Clearly $A_1 \cup A_2 \cup A_0$ is a partition of X and $B_1 \cup B_2 \cup B_0$ is a partition of Y. We claim that A_1, B_1, A_2, B_2 are very close to A, B, A^c, B^c, respectively (so A_0 and B_0 are fairly small) and subgraphs $G[A_1, B_2]$ and $G[A_2, B_1]$ are almost complete. ■

Claim 4.11. *Assume that $\alpha^{1/3} < \frac{1}{2}$ and $\delta(G) \geq (u/h)n$ (so $c_1(H)$ is unnecessary here).*

1. $(1-\alpha^{2/3})|A| \leq \{|A_1|, |B_1|\} \leq (1+\alpha^{2/3})|A|$ *and* $|A^c| - \alpha^{2/3}|A| \leq \{|A_2|, |B_2|\} \leq |A^c| + \alpha^{2/3}|A|$.
2. $\delta(B_2, A_1) \geq (1-2\alpha^{1/3})|A_1|$, $\delta(A_2, B_1) \geq (1-2\alpha^{1/3})|B_1|$ *and* $\delta(A_1, B_2) \geq (1-2\alpha^{1/3}(w/u))|B_2|$, $\delta(B_1, A_2) \geq (1-2\alpha^{1/3}(w/u))|A_2|$.
3. $\Delta(B_1, A_1), \Delta(A_1, B_1) \leq |A|(\alpha^{2/3} + \alpha^{1/3})$.
4. $|A_0|, |B_0| \leq 2\alpha^{2/3}|A|$ *and* $\delta(A_0, B_1), \delta(B_0, A_1) \geq (\alpha^{1/3} - \alpha^{2/3})|A|$.

Proof.
Part 1. We only prove bounds for $|A_1|$ and $|A_2|$; the calculations for $|B_1|$ and $|B_2|$ are exactly the same. By definition of A_1,

$$e(A-A_1, B) \geq \delta(A-A_1, B)|A-A_1| \geq \alpha^{\frac{1}{3}}|B||A-A_1|.$$

On the other hand,

$$e(A-A_1, B) \leq e(A, B) \leq \alpha|A||B|.$$

Together they imply that $|A-A_1|\alpha^{1/3}|B| \leq \alpha|A||B|$ or $|A-A_1| \leq \alpha^{2/3}|A|$. Since $|A| - |A_1| \leq |A-A_1|$, it follows that $|A_1| \geq (1-\alpha^{2/3})|A|$.

In order to derive an upper bound for $|A^c - A_2|$, we need the minimum degree condition $\delta(G) \geq (u/h)n$. Since $\delta(G)$ is an integer, we actually have $\delta(G) \geq \lceil (u/h)n \rceil$. Then

$$e(B, A^c) = e(B, X) - e(A, B) \geq \left\lceil \frac{u}{h} n \right\rceil |B| - \alpha|A||B|.$$

Let $\bar{e}(B, A^c)$ denote the size of the bipartite complement of G on $[B, A^c]$. Since $\lfloor (w/h)n \rfloor + \lceil (u/h)n \rceil = n$, we have

$$\bar{e}(B, A^c) = |B||A^c| - e(B, A^c) \leq |B|\left(n - \left\lfloor \frac{w}{h} n \right\rfloor\right) - \left(\left\lceil \frac{u}{h} n \right\rceil |B| - \alpha|A||B|\right) = \alpha|A||B|.$$

By definition of A_2,

$$e(A^c - A_2, B) \leq (1-\alpha^{\frac{1}{3}})|B||A^c - A_2|.$$

Therefore,

$$\bar{e}(A^c - A_2, B) \geq |A^c - A_2||B| - (1-\alpha^{\frac{1}{3}})|B||A^c - A_2| = \alpha^{\frac{1}{3}}|B||A^c - A_2|.$$

The upper and lower bounds for $\bar{e}(A^c - A_2, B)$ together imply that $\alpha^{1/3}|B||A^c - A_2| \leq \alpha|A||B|$, which gives $|A^c - A_2| \leq \alpha^{2/3}|A|$. We thus deduce that $|A_2| \geq |A^c| - \alpha^{2/3}|A|$. Since $|A_0| + |A_1| + |A_2| = n = |A| + |A^c|$, we further have $|A_0| + |A_1| \leq |A| + \alpha^{2/3}|A|$. Together with $|A_1| \geq (1-\alpha^{2/3})|A|$, it yields that $|A| - \alpha^{2/3}|A| \leq |A_1| \leq |A| + \alpha^{2/3}|A|$. The lower bound for $|A_1|$ also implies that $|A_2| \leq |A^c| + \alpha^{2/3}|A|$. Together with $|A_2| \geq |A^c| - \alpha^{2/3}|A|$, we thus obtain desired bounds for $|A_2|$.

The proof above actually gives that

$$|A - A_1|, \quad |B - B_1|, \quad |A^c - A_2|, \quad |B^c - B_2| \leq \alpha^{2/3}|A|.$$

Part 2. Let us consider the minimum degree between A_1 and B_2 here; the same holds for the degree between B_1 and A_2. First $\delta(B_2,A_1) \geq \delta(B_2,A) - |A - A_1| \geq (1 - \alpha^{1/3} - \alpha^{2/3})|A|$. By using $\delta(G) \geq \lceil un/h \rceil = |B^c|$, we derive that

$$\delta(A_1,B_2) \geq \delta(A_1,B^c) - |B^c - B_2| \geq \delta(G) - \alpha^{\frac{1}{3}}|B| - |B^c - B_2| \geq |B^c| - (\alpha^{\frac{1}{3}} + \alpha^{\frac{2}{3}})|B|.$$

We now prove that $\delta(B_2,A_1)/|A_1| \geq 1 - 2\alpha^{1/3}$. By Part 1, $|A_1| \leq (1 + \alpha^{2/3})|A|$. Then

$$\frac{\delta(B_2,A_1)}{|A_1|} \geq \frac{(1 - \alpha^{\frac{1}{3}} - \alpha^{\frac{2}{3}})|A|}{(1 + \alpha^{\frac{2}{3}})|A|} \geq 1 - 2\alpha^{\frac{1}{3}}$$

because $\alpha^{1/3} > 2\alpha^{2/3}$.

Similarly we can prove $\delta(A_1,B_2)/|B_2| \geq 1 - 2\alpha^{1/3}(w/u)$ though we also need $|B| \leq (w/h)n \leq (w/u)|B^c|$:

$$\frac{\delta(A_1,B_2)}{|B_2|} \geq \frac{|B^c| - (\alpha^{\frac{1}{3}} + \alpha^{\frac{2}{3}})|B|}{|B^c| + \alpha^{\frac{2}{3}}|B|} \geq \frac{|B^c| - (\alpha^{\frac{1}{3}} + \alpha^{\frac{2}{3}})|B^c|\frac{w}{u}}{|B^c| + \alpha^{\frac{2}{3}}|B^c|\frac{w}{u}}.$$

By using $\alpha^{1/3} > 2\alpha^{2/3}$ again, we derive that $\delta(A_1,B_2)/|B_2| \geq 1 - 2\alpha^{1/3}(w/u)$.

Part 3. By using $|A_1 - A| \leq |A^c - A_2| \leq \alpha^{2/3}|A|$, we obtain $\Delta(B_1,A_1) \leq \Delta(B_1,A) + |A_1 - A| \leq (\alpha^{1/3} + \alpha^{2/3})|A|$. The same holds for $\Delta(A_1,B_1)$.

Part 4. Part 1 immediately implies that $|A_0|, |B_0| \leq 2\alpha^{2/3}|A|$. By definition of A_1, we have $\delta(A_0,B_1) \geq \alpha^{1/3}|B| - |B - B_1| \geq (\alpha^{1/3} - \alpha^{2/3})|B|$. The same holds for $\delta(B_0,A_1)$. ■

Recall that $2n = mh$. We now separate the proof into two parts, when m is even and when m is odd. We give all details in Part 1, including the exact values of α and n, and while reducing Part 2 to Part 1, we only justify the value of $c_1(H)$.

Part I: *m is even*. Let $\rho = \rho(w)$ be given as in Lemma 4.10. We define $\alpha > 0$ such that

$$\alpha^{\frac{1}{3}} = \min\left\{\frac{1}{5h^2}, \frac{\rho}{2h}\right\}, \tag{10}$$

Let $\zeta = \zeta(H)$. By choosing m_0 sufficiently large, we may assume $m \geq 2\zeta h^2/\alpha^{2/3}$ and consequently $n = mh/2$ satisfies

$$n\alpha^{\frac{2}{3}} \geq \zeta h^3. \tag{11}$$

Let $G_1 = G[A_1, B_2 \cup B_0]$ and $G_2 = G[B_1, A_2 \cup A_0]$ denote the induced subgraphs of G on $A_1 \cup B_2 \cup B_0$ and $B_1 \cup A_2 \cup A_0$, respectively. Our first step is to remove some copies of H so that the orders of G_1 and G_2 are divisible by h.

Suppose that $v(G_1) \equiv r \pmod h$ and accordingly $v(G_2) \equiv -r \pmod h$ for some $0 \leq r < h$.

Claim 4.12. *We may remove $2r\zeta$ copies of H from G where $r\zeta h + r$ vertices come from G_1 and $r\zeta h - r$ vertices come from G_2. On the other hand, $r\zeta h$ vertices are from each of X and Y.*

Proof. We first note that since $G_1[A_1, B_2]$ and $G_2[A_2, B_1]$ are almost complete, we may find many disjoint copies of H in them. In fact, since $|A_1|/|B_2|$ is about w/u, $K_{|A_1|,|B_2|}$ contains an H-tiling that covers most of its vertices. By Claim 4.11,

$\delta(B_2, A_1) \geq (1 - 2\alpha^{1/3})|A_1|$ and $\delta(A_1, B_2) \geq (1 - 2\alpha^{1/3}(w/u))|B_2|$. By (10), $2\alpha^{1/3}(w/u) \leq \rho$. Lemma 4.10 thus implies that $G_1[A_1, B_2]$ contains an H-tiling that covers most of its vertices.

We remove $2r\zeta$ copies of H as follows: from $G_1[A_1, B_2]$, remove $r(\zeta + \zeta_i)$ copies of C_i, and from $G_2[A_2, B_1]$, remove $r(\zeta - \zeta_i)$ copies of C_i for all $i = 1, \ldots, k_c$. Now fix an index i. Note that $r(\zeta + \zeta_i)$ and $r(\zeta - \zeta_i)$ have the same parity. If they are even, then we remove $r(\zeta + \zeta_i)/2$ copies of C_i from G_1 with the larger side in X, and the other $r(\zeta + \zeta_i)/2$ copies of C_i from G_2 with the smaller side in X. Similarly, remove $r(\zeta - \zeta_i)/2$ copies of C_i from G_2 with the larger side in X, and the other copies of H with the smaller side in X. Clearly X and Y lose the same number of vertices for each i. Since at the end X and Y together lose $2r\zeta h$ vertices, each of them loses $r\zeta h$ vertices. If $r(\zeta + \zeta_i)$ is odd, then remove $\lceil r(\zeta + \zeta_i)/2 \rceil$ copies of C_i from G_1 with the larger side in X and $\lfloor r(\zeta + \zeta_i)/2 \rfloor$ copies of C_i from G_1 with the smaller side in X (therefore X loses $w_i - u_i$ more vertices than Y). On the other hand, we remove $\lfloor r(\zeta - \zeta_i)/2 \rfloor$ copies of C_i from G_2 with the larger side in X and $\lceil r(\zeta - \zeta_i)/2 \rceil$ copies of C_i from G_2 with the smaller side in X (this makes Y lose $w_i - u_i$ more vertices than X). Thus X and Y again lose the same number of vertices: each loses $r\zeta h$ vertices at the end. The total number of vertices that G_1 loses is

$$r(\zeta + \zeta_1)c_1 + \cdots + r(\zeta + \zeta_{k_c})c_{k_c} = r\zeta(c_1 + \cdots + c_{k_c}) + r(\zeta_1 c_1 + \cdots + \zeta_{k_c} c_{k_c}) = r\zeta h + r.$$

A similar calculation shows that G_2 loses $r\zeta h - r$ vertices. ■

Denote the sets of the remaining vertices in X, Y, A_1, A_2, B_1, B_2 by $X', Y', A_1', A_2', B_1', B_2'$, respectively (Fig. 4). The difference between $|A_1|$ and $|A_1'|$ (similarly between $|B_2|$ and $|B_2'|$, etc.) is at most $r\zeta h$. Our choice (11) of n is equivalent to $wh^2\zeta \leq \alpha^{2/3}(w/h)n$. Since $r \leq h - 1$ and $|A| = (w/h)n$, we derive that

$$wr\zeta h \leq \alpha^{\frac{2}{3}}|A|. \tag{12}$$

Let $\tilde{A}_2 = A_2' \cup A_0$ and $\tilde{B}_2 = B_2' \cup B_0$. The current G_1, G_2 are $G_1[A_1', \tilde{B}_2]$ and $G_2[B_1', \tilde{A}_2]$, respectively. By Claim 4.12, both $v(G_1)$ and $v(G_2)$ are divisible by h. Let $m_1 = v(G_1)/h$, $m_2 = v(G_2)/h$, and write

$$|A_1'| = m_1 w + s, \quad |B_1'| = m_2 w + t, \quad |\tilde{A}_2| = m_2 u - t, \quad |\tilde{B}_2| = m_1 u - s$$

for some integers s and t. Since X' and Y' have equal number of vertices, we have $m_1 w + s + m_2 u - t = m_2 w + t + m_1 u - s$, which implies that

$$(m_1 - m_2)(w - u) = 2(t - s). \tag{13}$$

Without loss of generality, assume that $m_1 \geq m_2$. This implies $t \geq s$.

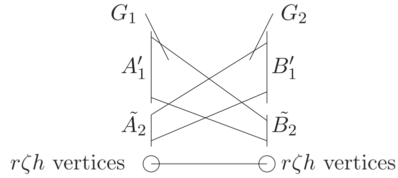

FIGURE 4. Graph G with sets $A_1', \tilde{A}_2, B_1', \tilde{B}_2$, and removed copies of H.

By using $\delta(G) \geq (u/h)n + c_1(H)$ and $m_2 \leq (m/2) = (n/h)$, we obtain a lower bound on $\delta(B_1', A_1')$:

$$\delta(B_1', A_1') \geq \delta(G) - |\tilde{A_2}| - r\zeta h = \frac{u}{h}n + c_1(H) - m_2 u + t - r\zeta h \geq c_1(H) + t - r\zeta h. \tag{14}$$

Now we use the assumption that m is even: $m - 2r\zeta = m_1 + m_2$ is even; thus. $m_1 - m_2$ is even. Then, by (13), we see that $w - u$ divides $t - s$.

We now separate the cases when $t \geq 0$ and when $t < 0$.

Case 1. Assume $t \geq 0$. We claim that t is reasonably small. In fact, by Claim 4.12, $v(G_2) = |A_2| + |A_0| + |B_1| - (r\zeta h - r)$. From Claim 4.11, we know that $|A_2| + |B_1| \geq n - 2\alpha^{2/3}|A|$ and consequently $m_2 = v(G_2)/h \geq (n - 2\alpha^{2/3}|A| - r\zeta h)/h$. By definition,

$$t = |B_1'| - m_2 w \leq |A| + \alpha^{\frac{2}{3}}|A| - \frac{w}{h}n + 2\frac{w}{h}\alpha^{\frac{2}{3}}|A| + wr\zeta = \alpha^{\frac{2}{3}}|A| + 2\frac{w}{h}\alpha^{\frac{2}{3}}|A| + wr\zeta.$$

By (12), we have $wr\zeta \leq (1/h)\alpha^{2/3}|A|$ and thus $t \leq 3\alpha^{2/3}|A|$.

We want to move t vertices from A_1' to $\tilde{A_2}$ and t vertices from B_1' to $\tilde{B_2}$. To move these vertices, we will find t w-stars from B_1' to A_1' and t w-stars from A_1' to B_1' by the following lemma (Lemma 12 from [20]), and then move the centers of these stars.

Lemma 4.13 (Zhao [20]). *Let $1 \leq k \leq \delta \leq M$ be positive integers, and $0 < c < 1/(6k+7)$. Let $F[V_1, V_2]$ be a bipartite graph such that $||V_i| - M| \leq cM$ for $i = 1, 2$. If $\delta \leq \delta(V_1, V_2) \leq cM$ and $\Delta(V_2, V_1) \leq cM$, then F contains $2(\delta - k + 1)$ vertex-disjoint k-stars of which $\delta - k + 1$ are centered in V_1 and $\delta - k + 1$ are centered in V_2.*

By (8), we have $c_1(H) > r\zeta h + w - 1$. With (14), it implies that

$$\delta(B_1', A_1') - w + 1 \geq c_1(H) + t - r\zeta h - w + 1 > t. \tag{15}$$

On the other hand, $\delta(B_1', A_1'), \Delta(A_1', B_1') \leq (\alpha^{1/3} + \alpha^{2/3})|A|$ by Claim 4.11. From (10) and the fact that $w \geq 2$, we can derive that $2\alpha^{1/3} \leq 2/(15(w+1)) < 1/(6w+7)$. Thus, Lemma 4.13 provides t vertex-disjoint w-stars centered in A_1' and t vertex-disjoint w-stars centered in B_1'. We now move the centers of these stars from A_1' to $\tilde{A_2}$ and from B_1' to $\tilde{B_2}$. The resulting $A_1', \tilde{A_2}, B_1', \tilde{B_2}$ satisfy

$$|A_1'| = m_1 w + s - t, \quad |\tilde{B_2}| = m_1 u - s + t; \quad |B_1'| = m_2 w, \quad |\tilde{A_2}| = m_2 u.$$

Below we explain how to find an H-factor in G_1; the same procedure works for G_2.

The resulting G_1 contains $t \leq 3\alpha^{2/3}|A|$ disjoint w-stars centered at $\tilde{B_2}$. By definition, $B_0 \subset \tilde{B_2}$. We next find $|B_0|$ disjoint w-stars centered at B_0 from G_1 which are also disjoint from the existing w-stars. From Claim 4.11, we have $|B_0| < 2\alpha^{2/3}|A|$ and $\delta(B_0, A_1) \geq (\alpha^{1/3} - \alpha^{2/3})|A|$. Since $|A_1'| \geq |A_1| - r\zeta h - t$ and $r\zeta h \leq \alpha^{2/3}|A|$, we have

$$\delta(B_0, A_1') \geq \delta(B_0, A_1) - (t + r\zeta h) \geq (\alpha^{\frac{1}{3}} - \alpha^{\frac{2}{3}})|A| - 3\alpha^{\frac{2}{3}}|A| - \alpha^{\frac{2}{3}}|A| = (\alpha^{\frac{1}{3}} - 5\alpha^{\frac{2}{3}})|A|.$$

Since $\alpha^{1/3} \geq 5h\alpha^{2/3} \geq 5(w+1)\alpha^{2/3}$ by (10), we derive that

$$\delta(B_0, A_1') \geq (\alpha^{\frac{1}{3}} - 5\alpha^{\frac{2}{3}})|A| \geq 5w\alpha^{\frac{2}{3}}|A| \geq w(|B_0| + t).$$

We may therefore choose disjoint w-stars for the vertices of B_0 greedily.

Now, we have $t+|B_0|$ w-stars centered in $\tilde{B}_2$. For each star, we will find a copy of $K_{u,w}$ (a supergraph of H), such that $u-1$ vertices come from B_2', and the rest are from the w-star. Recall that $|B_2-B_2'|\le r\zeta h$. Suppose that a w-star has leaves $v_1,\dots,v_w$ in A_1'. We claim that $|\bigcap_{i=1}^{w}\Gamma(v_i,B_2')|\ge(u-1)(|B_0|+t)$; thus, we can greedily find a copy of $K_{u,w}$ for each star such that it is vertex disjoint from the existing copies of $K_{w,u}$. In fact, by Claim 4.11 and (12),

$$\left|\bigcap_{i=1}^{w}\Gamma(v_i,B_2')\right|\ge\left(1-w\frac{w}{u}2\alpha^{\frac{1}{3}}\right)|B_2|-r\zeta h\ge\left(1-\frac{2w^2}{u}\alpha^{\frac{1}{3}}\right)\left(1-\alpha^{\frac{2}{3}}\right)|B|-\alpha^{\frac{2}{3}}|B|$$

$$\ge\left(1-\frac{2w^2}{u}\alpha^{\frac{1}{3}}-2\alpha^{\frac{2}{3}}\right)|B|.$$

By (10), we have $5u\alpha^{2/3}<\alpha^{1/3}$ and $(2w^2/u+1)\alpha^{1/3}<2h^2\alpha^{1/3}<1$. Consequently

$$\left|\bigcap_{i=1}^{w}\Gamma(v_i,B_2')\right|-(u-1)(|B_0|+t)\ge\left(1-\frac{2w^2}{u}\alpha^{\frac{1}{3}}-2\alpha^{\frac{2}{3}}\right)|B|-(u-1)5\alpha^{\frac{2}{3}}|B|$$

$$>\left(1-\frac{2w^2}{u}\alpha^{\frac{1}{3}}-\alpha^{\frac{1}{3}}\right)|B|>0.$$

We remove these copies of $K_{w,u}$, and let A_1'' and B_2'' denote the set of remaining vertices in A_1' and $\tilde{B}_2$. We know that $A_1''\subseteq A_1$ and $B_2''\subseteq B_2$ satisfy

$$|A_1|\ge|A_1''|\ge|A_1|-r\zeta h-t-w(|B_0|+t),\quad |B_2|\ge|B_2''|\ge|B_2|-r\zeta h-(u-1)(|B_0|+t).$$

Furthermore, $|A_1''|=m_1'w+s-t$ and $|B_2''|=m_1'u-s+t$ for some large integer m_1'. Since $w-u$ divides $t-s$, we can apply Lemma 3.6 and obtain an H-factor of $K_{|A_1''|,|B_2''|}$. It remains to show that $G[A_1'',B_2'']$ satisfy the condition (9) of Lemma 4.10 (then Lemma 4.10 provides an H-factor of $G[A_1'',B_2'']$).

In fact, by Claim 4.11,

$$\delta(B_2'',A_1'')\ge(1-2\alpha^{\frac{1}{3}})|A_1|-r\zeta h-t-w(|B_0|+t)$$

$$\ge(1-2\alpha^{\frac{1}{3}})|A_1|-\alpha^{\frac{2}{3}}|A|-3\alpha^{\frac{2}{3}}|A|-w(5\alpha^{\frac{2}{3}}|A|).$$

By (10) and $w+1\le h$, we have $\alpha^{1/3}\ge5(w+1)\alpha^{2/3}$, which implies that, by Claim 4.11,

$$\alpha^{\frac{1}{3}}|A_1|\ge\alpha^{\frac{1}{3}}(1-\alpha^{\frac{2}{3}})|A|\ge(4\alpha^{\frac{2}{3}}+5w\alpha^{\frac{2}{3}})|A|.$$

Consequently $\delta(B_2'',A_1'')\ge(1-3\alpha^{1/3})|A_1|\ge(1-3\alpha^{1/3})|A_1''|$.

On the other hand,

$$\delta(A_1'',B_2'')\ge\left(1-\frac{2w}{u}\alpha^{\frac{1}{3}}\right)|B_2|-r\zeta h-(u-1)(|B_0|+t)$$

$$\ge\left(1-\frac{2w}{u}\alpha^{\frac{1}{3}}\right)|B_2|-\alpha^{\frac{2}{3}}|B|-(u-1)5\alpha^{\frac{2}{3}}|B|.$$

By (10), we have $\alpha^{13}(u/h) \geq 5u\alpha^{2/3}$. Together with $|B_2| \geq |B^c| - \alpha^{2/3}|B| \geq ((u/h) - \alpha^{2/3})|B|$, we have

$$\alpha^{\frac{1}{3}}|B_2| \geq \alpha^{\frac{1}{3}}\left(\frac{u}{h} - \alpha^{\frac{2}{3}}\right)|B| \geq \alpha^{\frac{2}{3}}|B| + 5(u-1)\alpha^{\frac{2}{3}}|B|.$$

Consequently, $\delta(A_1'', B_2'') \geq (1-(2w/u)\alpha^{1/3} - \alpha^{1/3})|B_2| \geq (1-2h\alpha^{1/3})|B_2''|$. By (10), we have $3\alpha^{1/3} \leq 2h\alpha^{1/3} \leq \rho$, and thus $\delta(B_2'', A_1'') \geq (1-\rho)|A_1''|$, and $\delta(A_1'', B_2'') \geq (1-\rho)|B_2''|$, as stated in (9).

Case 2. Assume $t<0$. Let $-t = q(w-u)+p$ for some nonnegative integers q and p such that $p<w-u$. Since $-s \geq -t$ and $w-u$ divides $t-s$, we may write $-s = q'(w-u)+p$ for some integer $q' \geq q$. Similar as in Case 1, we derive that $-s \leq 3\alpha^{2/3}|A|$.

First, assume that $q \geq p\beta$. Then by Lemma 3.6, $K_{|A_1'|,|\tilde{B}_2|}$ and $K_{|B_1'|,|\tilde{A}_2|}$ each contains an H-factor (here we need $n \gg -s, -t$). In order to obtain an H-factor in $G_1 = G[A_1', \tilde{B}_2]$ (similar for $G_2 = G[B_1', \tilde{A}_2]$), as in Case 1, we first find $|B_0|$ disjoint w-stars with centers at B_0 and leaves in A_1'. Then we extend these w-stars to (disjoint) copies of $K_{w,u}$ and finally apply Lemma 4.10 to find an H-factor covering the remaining part of G_1.

Second, assume that $q \leq p\beta - 1$. Our first step is to move $w-u-p$ vertices from A_1' to $\tilde{A}_2$, and $w-u-p$ vertices from B_1' to $\tilde{B}_2$. By (8), we have $c_1(H) > p\beta(w-u) + r\zeta h + w \geq (q+1)(w-u) + r\zeta h + w$. With (14), this implies that

$$\delta(B_1', A_1') - w + 1 > c_1(H) + t - r\zeta h - w + 1 \geq w - u - p. \tag{16}$$

Applying Lemma 4.13, we find $2(w-u-p)$ vertex-disjoint w-stars in $G[A_1', B_1']$ with $w-u-p$ of them centered at A_1' and the other $w-u-p$ stars centered at B_1'. After moving the centers of these stars from A_1' to $\tilde{A}_2$ and from B_1' to $\tilde{B}_2$, we have

$$|A_1'| = m_1 w + s - (w-u-p) = m_1 w - (q'+1)(w-u),$$

$$|\tilde{A}_2| = m_2 u - t + (w-u-p) = m_2 u + (q+1)(w-u),$$

$$|B_1'| = m_2 w - (q+1)(w-u), \quad |\tilde{B}_2| = m_1 u + (q'+1)(w-u).$$

By Lemma 3.6, $K_{|A_1'|,|\tilde{B}_2|}$ and $K_{|B_1'|,|\tilde{A}_2|}$ each contains an H-factor. Next we find H-factors of G_1 and G_2 as above.

Part II: Assume m is odd. In this case we use an idea used in the proof of Lemma 16 in [13]: we will use $hcf_c(H)=1$ to remove a small number of copies of H such that the remaining vertices of G form a balanced, bipartite graph of size $2n' = m'h$ where n' is divisible by H. Then, we apply the proof of *Part I* to this graph, and complete our tiling.

Because m is odd and $mh = 2n$ is even, then h must be even. Moreover, since $hcf_c(H) = 1$, there exists a component $C_i[U_i, W_i]$ of H with an odd number of vertices. Since c_i is odd, $w_i - u_i$ is odd. Now, let c_1 be the 2-coloring of H with color classes U and W of sizes u and w, respectively (then $U_i \subset U$ and $W_i \subset W$). We obtain another coloring c_2 of H by swapping the colors of U_i and W_i from c_1. Suppose that c_2 has color classes U' and W' such that $|U'| = u' \leq w' = |W'|$. Since h is even, u and w have the same parity, and u' and w' have the same parity. Additionally, since $w_i - u_i$ is odd, the parities of u, w and u', w' are different.

Let $k_1=(h/2)-u'$ and $k_2=(h/2)-u$ (so $k_1,k_2\geq 0$). From $G[A_1,B_2]$, remove k_1 copies of H with u vertices in A_1 and w vertices in B_2, and remove k_2 copies of H with w' vertices in A_1 and u' vertices in B_2. This is possible because $G[A_1,B_2]$ is almost complete. Denote the sets of the remaining vertices in X and Y by X' and Y', respectively.

We claim that $|X'|=|Y'|$. Since $|X|=|Y|$, it suffices to show that $|X|-|X'|=|Y|-|Y'|$. In fact, since $|X|-|X'|=k_1u+k_2w'$ and $|Y|-|Y'|=k_1w+k_2u'$, by the definitions of k_1 and k_2,

$$\begin{aligned} & k_1u+k_2w'=k_1w+k_2u' \\ \Leftrightarrow & \left(\frac{u'+w'}{2}-u'\right)u+\left(\frac{u+w}{2}-u\right)w'=\left(\frac{u'+w'}{2}-u'\right)w+\left(\frac{u+w}{2}-u\right)u' \\ \Leftrightarrow & \frac{w'-u'}{2}u+\frac{w-u}{2}w'=\frac{w'-u'}{2}w+\frac{w-u}{2}u', \end{aligned}$$

which is equivalent to the identity $\frac{1}{2}(w'-u')(w-u)=\frac{1}{2}(w'-u')(w-u)$.

Let $n'=|X'|=|Y'|$. We have $n-n'=(k_1+k_2)(h/2)=(h-u-u')(h/2)$. Since h is even and $u+u'$ is odd, we have $n-n'\equiv(h/2) \bmod h$. Since $n=mh/2\equiv(h/2) \bmod h$, we derive that n' is divisible by h. In the new graph $G'=G[X',Y']$, we have

$$\delta(G')\geq\delta(G)-(h-u-u')\frac{h}{2}\geq\frac{u}{h}n+c_1(H)-(w-u)\frac{h}{2}$$

by using $u'\geq u$. By (8), $c_1(H)=\beta(w-u)^2+\zeta h^2+(w-u)(h/2)+w$. We thus have $\delta(G')\geq(u/h)n+\beta(w-u)^2+\zeta h^2+w$. Hence, (15) and (16) hold and we may apply the proof of Part I to G' obtaining an H-factor. ■

5. PROOF OF THEOREM 1.5

Let H be a bipartite graph on h vertices with $u=\sigma(H)$ and $w=h-\sigma(H)$. Let G be a balanced bipartite graph on $2n$ vertices with $\delta(G)\geq(u/h)n$. We assume $u<w$; otherwise, we can obtain the desired H-tiling as follows. Add $3h$ new vertices to each side of G and join them with all the existing vertices on the opposite side. The new graph G' has $\delta(G')\geq(n/2)+3h=(n+3h)/2+3h/2$. By Theorem 1.2, G' contains an H-factor $\mathcal{H}$, which gives rise to an H-tiling of G that misses at most $6h(h-1)$ vertices because at most $6h$ copies of H in $\mathcal{H}$ may contain the vertices of $G'-G$, and each copy of H may contain at most $h-1$ vertices of G.

Part 1 of the following lemma is a replacement of Corollary 3.7 when $hcf_{\chi,c}(H)\neq 1$; Part 2 is needed for the extremal case.

Lemma 5.1.

1. *Let $G[X,Y]$ be a complete bipartite graph with $(u/w)\leq|X|/|Y|\leq 1$. Then G has a $K_{u,w}$-tiling that leaves out $l(X)$ vertices in X and $l(Y)$ vertices in Y such that $l(X)+l(Y)\leq h+(w-u)-2$. In this $K_{u,w}$-tiling, at least $m/2-h$ copies of $K_{u,w}$ have their w-vertex sides in Y, where $m=\lfloor(|X|+|Y|)/h\rfloor$.*
2. *Let $m>c$ be positive integers. Then $G[X,Y]=K_{mu-c,mw+c}$ contains a $K_{u,w}$-tiling that covers all but at most $(c+u-1)(h/u)$ vertices.*

Proof. *Part* 1: Let $r \equiv |X|+|Y| \bmod h$ (then $0 \le r \le h-1$). We may write $|X| = mu+t$ and $|Y| = mw-t+r$. Since $|X|/|Y| \ge u/w$, we have $|X| \ge (|X|+|Y|)(u/h) \ge mu$, which implies that $t \ge 0$. We next write $t = q(w-u)+p$ for some integers q and $0 \le p \le w-u-1$. We now have two cases.

First, if $p \le r$, then we may tile G with m copies of $K_{u,w}$ where $m-q$ copies have their w-vertex sides placed in Y, and q copies have their w-vertex sides placed in X. This tiling covers $(m-q)w+qu = mw-t+p = |Y|-(r-p)$ vertices of Y and $(m-q)u+qw = mu+t-p = |X|-p$ vertices of X. Let $l(X) = p$ and $l(Y) = r-p$. We have $l(X)+l(Y) = r \le h-1$.

Otherwise, $p > r$. In that case, tile G with $m-q-1$ copies of $K_{u,w}$ with their w-vertex sides placed in Y, and q copies of $K_{u,w}$ with their w-vertex sides placed in X. This tiling covers $(m-q-1)w+qu = mw-(t-p)-w = |Y|+p-(r+w)$ vertices of Y and $(m-q-1)u+qw = mu+t-(p+u) = |X|-(p+u)$ vertices of X. Let $l(X) = p+u$ and $l(Y) = r+w-p$. We have $l(X)+l(Y) = r+h \le h+w-u-2$ since $r < p \le w-u-1$.

In both cases, our H-tiling contains at least $m-q-1$ copies of $K_{u,w}$ with their w-vertex sides in Y. Since $|X| \le |Y|$, we have $mu+t \le mw-t+r$, or $2t \le m(w-u)+r$. With $t = q(w-u)+p$, this gives $m \ge 2q+(2p-r)/(w-u)$. By using $r \le h-1$, we have $m-q-1 \ge (m/2)-(h-1)/(2(w-u))-1 \ge m/2-h$.

Part 2: Write $c = pu+q$ for integers p, q such that $0 \le q < u$. If $q = 0$, then $|X| = mu-pu$ and $G \supset K_{(m-p)u,(m-p)w}$, which consists of $m-p$ copies of $K_{u,w}$. It leaves $c+pw = ph = ch/u$ vertices in Y uncovered. Otherwise $q \ge 1$ and $G \supset K_{(m-p-1)w,(m-p-1)u}$, which consists of $m-p-1$ copies of $K_{u,w}$. It leaves $u-q$ vertices in X and $c+(p+1)w$ vertices in Y uncovered. The total number of uncovered vertices is

$$u-q+c+(p+1)w = u+pu+(p+1)w = h+ph = h\left(\frac{c-q}{u}+1\right) \le (c+u-1)\frac{h}{u}.$$

Proof of Theorem 1.5. First note what is different here from Theorem 1.4: (i). we do not assume that H is in Class 1; (ii) the $\delta(G)$ condition has no extra constant $c_1(H)$; (iii) at most $c_2(H)$ vertices may be left outside the desired H-tiling. Below we closely follow the proof of Theorem 1.4 but focus on the impact of these differences.

Non-extremal case: We assume G is not in the extremal case, which is defined exactly as in Theorem 1.4. First note that Theorem 4.1 has no $c_1(H)$ in the minimum degree condition, and Lemma 4.3 does not assume that H is in Class 1. We thus apply Lemma 4.3 to get a decomposition of G into super-regular cluster pairs $(P_1, Q_1), \ldots, (P_k, Q_k)$, and exceptional sets X_0, Y_0. We can not apply Lemma 4.4 directly because it assumes that H is in Class 1. If we follow the proof of Lemma 4.4, we can apply Claim 4.7 to get rid of the exceptional sets but we cannot use Claim 4.8 because we do not have $hcf_c(H) = 1$. Actually, even if h divides $|P_i|+|Q_i|$, we cannot use Corollary 3.7 to obtain an H-factor on $P_i \cup Q_i$ because we do not have $hcf_{\chi,c}(H) = 1$. Instead we can only apply Lemma 5.1 to obtain an H-tiling that omits at most $h+(w-u)-2$ vertices of $P_i \cup Q_i$. If we apply Lemma 5.1 to each (P_i, Q_i), then we obtain an H-tiling of G that omits at most $2hk$ vertices, where $k \le 2pM(\varepsilon)$ is a large constant depending on the large constant $M(\varepsilon)$ defined in the Regularity Lemma.

In order to reduce the number of uncovered vertices to a constant $O(h^2)$, we use the connection among $P_i, Q_i, i = 1, \ldots, k$, to gather all uncovered vertices in a few cluster pairs. This approach can be found in [18]. To facilitate our calculation, we need all P_i (and thus all Q_i) to have the same size. Let us go back to the moment right after we decompose the clusters of R. The second terms of the pairs in (7) are all possible sizes for P_i. It is easy to see that $N_0/(q(q^2-p^2))$ divides all of them. We then divide each P_i to subclusters of size $N_1 := N_0/(q(q^2-p^2))$, accordingly divide its partner Q_i to subclusters of size $N_2 := (q/p)N_1$, and match the resulting subclusters from P_i and those from Q_i arbitrarily. Let us still denote new cluster pairs by (P_i, Q_i), and use k for the number of the new cluster pairs. Let k_1 be the number of (P_i, Q_i) with $P_i \subset X$. We have $k_1 = k/2$ because there are the same number of vertices of G contained in the clusters of X and in the clusters of Y (note that $|X_0| = |Y_0|$). We call P_i and Q_i the *partners* of each other. To distinguish them, we call $P_1, \ldots, P_k$ *small* clusters and $Q_1, \ldots, Q_k$ *large* clusters.

Now let R' be the bipartite graph on $\{P_i, Q_i : i = 1, \ldots, k\}$ such that two clusters C, C' are adjacent if $d_{G'}(C, C') > 0$. Consider a vertex $C \in V(R')$. Since each cluster, P_i or Q_i, has at most N_2 vertices, by the same calculation as in (5), we derive that $\delta_{R'}(C) \geq (u/h - 2\gamma)n/N_2$. Since $N_1(k/2) + N_2(k/2) = \sum_{C \subset X} |C| \leq n$ and $N_1 > N_2(u/w)$, we obtain that $(n/N_2) > (1 + u/w)(k/2) = hk/(2w)$. Consequently $\delta_{R'}(C) \geq (u/(2w) - (h/w)\gamma)k$.

We next define a directed graph D_X whose vertices are all the current clusters in X, namely, $P_1, \ldots, P_{k/2}$, $Q_{k/2+1}, \ldots, Q_k$, and direct an edge from a cluster C to another C' if and only if $d(C, C'') > 0$, where C'' is the cluster in Y matched to C'. Then the minimum out-degree $\delta(D_X) = \delta_{R'}(C) \geq (u/(2w) - (h/w)\gamma)k$. Define the *sink* of D_X as a subset $S \subseteq V(D_X)$ such that for every vertex $v \in V(D_X)$, there is a vertex $s \in S$ and a directed path from v to s. A simple fact on digraphs (e.g. Lemma 6.7 in [18]) states that every digraph D contains a sink of size at most $|D|/\delta(D)$. Then D_X has a sink S_X of size at most

$$\frac{k}{\delta(D_X)} \leq \frac{k}{\left(\dfrac{u}{2w} - \dfrac{h}{w}\gamma\right)k} = \frac{2w}{u - 2h\gamma}.$$

Since $\gamma \ll 1$, this implies $|S_X| \leq 2w/u$. We similarly define the digraph D_Y on all the clusters of Y and obtain a sink S_Y of size at most $2w/u$. Let M_S be the set of all cluster pairs that contain at least one member of $S_X \cup S_Y$. Then $|M_S| \leq 4w/u$.

After this detour, we go back to the proofs of Lemmas 4.3 and 4.4: we obtain the super-regularity of all (P_i, Q_i) as in the proof Lemma 4.3 and then eliminate the exceptional sets X_0, Y_0 by Claim 4.7. Note that in these steps we only remove a small number of vertices from each cluster and thus do not change the adjacency in R', D_X, D_Y. Now all (P_i, Q_i) are super-regular and ratios $|P_i|/|Q_i|$ are slightly larger than u/w. Let $l(P_i)$ and $l(Q_i)$ be the numbers of leftover vertices in P_i and Q_i when we apply Lemma 5.1 to $K_{|P_i|,|Q_i|}$ (then $l(P_i) + l(Q_i) \leq h + w - u - 2$). Since $|P_i| + |Q_i|$ is sufficiently large, by Lemma 5.1, the values of $l(P_i), l(Q_i)$ do not change after we remove cu vertices from P_i and cw vertices from Q_i for any fixed integer c.

Before actually tiling (P_i, Q_i), we remove $l(C)$ vertices from each C not included in M_S as follows. Assume that $C \subset X$ and $l(C) = l_0$. By the definition of S_X, there is a directed path $C_0 C_1 \ldots C_t$ from $C_0 := C$ to some $C_t \in S_X$ in D_X. Let C'_j denote the partner of C_j for $1 \leq j \leq t$. For $0 \leq j < t$, we find l_0 disjoint copies of $K_{u,w}$, each of which consists one vertex of C_j and $w + u - 1$ vertices from $C_{j+1} \cup C'_{j+1}$ such that C_{j+1} loses $u - 1$ vertices if it is small or loses $w - 1$ vertices if it is large. At the end, C_0 loses l_0 vertices, C_t loses $l_0(u-1)$ vertices (if it is small) or $l_0(w-1)$ (if it is large) while any of the clusters $C_1, \ldots, C_{t-1}, C'_1, \ldots, C'_t$ loses $l_0 u$ vertices (if it is small) or $l_0 w$ vertices (if it is large). As a result, $l(C)$ becomes zero while $l(C_1), l(C_1'), \ldots, l(C_{t-1}), l(C_{t-1}')$ stay the same. We apply this procedure to every cluster C not included in M_S such that $l(C)=0$ at the end. Note that each cluster loses constant many (at most $4khw$) vertices even if it is contained in all the directed paths because there are at most $2k$ paths, and each path uses at most $(2h)w$ vertices from a single cluster. Hence, the resulting cluster pairs are still super-regular and satisfy $u/w \le |P_i|/|Q_i| \le 1$. Now we apply Lemma 5.1 and the Blow-up Lemma to each (P_i, Q_i) and obtain an perfect H-tiling unless $(P_i, Q_i) \in M_S$. Since each cluster pair in M_S contains at most $h+w-u-2$ uncovered vertices, we obtain an H-tiling of G that misses at most $|M_S|(h+w-u-2) \le (4w/u)(h+w-u-2) < 8h^2$ vertices.

Extremal case: Following the proof of Theorem 4.9, we first define A_i, B_i for $i = 0,1,2$. Claim 4.11 still holds because it only needs $\delta(G) \ge (u/h)n$. Then we do not need to separate the cases on the parity of m. Define $G_1 = G[A_1, B_2 \cup B_0]$ and $G_2 = G[B_1, A_2 \cup A_0]$ as well. Assume that $v(G_1) \equiv r \bmod h$ and $v(G_2) \equiv h-r \bmod h$ for some $0 \le r < h$. We remove arbitrary h vertices from A_1, $h-r$ vertices from B_1, and r vertices from B_2 and ignore them permanently. Denote the sets of the remaining vertices by $X', Y', A_1', A_2', B_1', B_2'$. Then $|X'| = |Y'| = n-h$. Let $\tilde{A}_2 = A_2' \cup A_0$ and $\tilde{B}_2 = B_2' \cup B_0$. Update G_1, G_2 as $G[A_1', \tilde{B}_2]$ and $G[B_1', \tilde{A}_2]$, respectively. Since both $v(G_1)$ and $v(G_2)$ are divisible by h, we have

$$|A_1'| = m_1 w + s, \quad |B_1'| = m_2 w + t, \quad |\tilde{A}_2| = m_2 u - t, \quad |\tilde{B}_2| = m_1 u - s \tag{17}$$

for some integers m_1, m_2, s, t. Without loss of generality, assume that $m_1 \ge m_2$ and consequently $t \ge s$. Let $c_0 = h+w-1$. We separate the cases when $t \le c_0$ and $t > c_0$.

First assume that $t \le c_0$ (so $s \le t \le c_0$). As in the proof of Theorem 4.9, we remove $|A_0| + |B_0|$ copies of $K_{w,u}$ from G_1 and G_2, each of which contains a vertex from $A_0 \cup B_0$, such that (17) holds for (slightly) smaller values of m_1 and m_2. If $t < 0$, then

$$\frac{u}{w} \le \frac{|\tilde{B}_2|}{|A_1'|} < 1 \quad \text{and} \quad \frac{u}{w} \le \frac{|\tilde{A}_2|}{|B_1'|} < 1.$$

By Lemma 5.1, Part 1, $K_{|\tilde{B}_2|, |A_1'|}$ and $K_{|\tilde{A}_2|, |B_1'|}$ each contains an H-tiling that misses at most $h+(w-u)-2$ vertices. Consequently, by Lemma 4.10, G_1 and G_2 contain the same H-tilings. The number of uncovered vertices in this case is at most $2h+2(h+w-u-2)$. If $0 \le t \le c_0$, then by Lemma 5.1, Part 2, and Lemma 4.10, each of G_1 and G_2 contains an H-tiling that misses at most $(c_0+u-1)h/u = (h+w-1+u-1)h/u \le (2h-2)h$ vertices. The total number of uncovered vertices in this case is at most $2h+2(2h-2)h$.

Now assume that $t > c_0$. After removing $c_1(H)$ and replacing $r\zeta h$ by h in (14), we obtain that $\delta(B_1', A_1') \ge t-h$. Applying Lemma 4.13, we find $2(t-h-w+1) = 2(t-c_0)$ vertex-disjoint w-stars with $t-c_0$ of them centered at A_1' and other $t-c_0$ of them centered at B_1'. After moving the centers of these stars to $\tilde{A}_2$ and $\tilde{B}_2$, we have

$$|A_1'| = m_1 w + s - t + c_0, \quad |\tilde{B}_2| = m_1 u - s + t - c_0; \quad |B_1'| = m_2 w + c_0, \quad |\tilde{A}_2| = m_2 u - c_0.$$

After getting rid of $A_0 \cup B_0$ as before, we apply Lemma 4.10 together with Lemma 5.1, Part 2, to obtain H-tilings in G_1 and G_2, each of which misses at most $(c_0+u-1)h/u \le (2h-2)h$ vertices (note that $s-t+c_0 \le c_0$). The total number of uncovered vertices in this case is at most $2h+2(2h-2)h$.

In summary, the number of uncovered vertices in the extremal case is at most $2h+2(2h-2)h<4h^2$. ■

6. CONCLUDING REMARKS

As mentioned in Section 1, Theorem 1.4 implies an approximate version of Theorem 1.1 for bipartite H, in which the constant C is replaced by $o(n)$. Furthermore, if the following conjecture of Bollobás and Scott [2] is true, Theorem 1.4 implies Theorem 1.1 with $C \le 8h^3$ for bipartite H.

Conjecture 6.1 (Bollobás and Scott [2]). *Every graph G of even order contains a balanced bipartite spanning subgraph B such that for every vertex v in G,* $\deg_B(v) \ge (\deg_G(v)/2) - \frac{1}{2}$.

In fact, to derive Theorem 1.1 for bipartite H, it suffices to have a weaker version of Conjecture 6.1: every graph G of even order contains a balanced, bipartite spanning subgraph B such that $\delta(B) \ge (\delta(G)/2) - c$, where c is some absolute constant.

After seeing the similarity between Theorem 1.1 and Theorems 1.4, it is reasonable to expect such a result for r-partite tiling. In an r-partite graph G, we define the *pairwise minimum degree* $\bar{\delta}(G)$ as the minimum degree from a vertex in one partition set to any other partition set.

Conjecture 6.2. *Let H be a graph with order h and chromatic number r. There exist integers C and m_0 such that for all $m \ge m_0$, if G is a balanced r-partite graph with $n = mh$ vertices in each partition set such that*

$$\bar{\delta}(G) \ge \begin{cases} (1-1/\chi_{cr}(H))n + C & \text{if } H \text{ is in Class } 1, \\ (1-1/\chi(H))n + C & \text{otherwise,} \end{cases}$$

then G contains an H-factor.

At present Conjecture 6.2 is out of reach as it has not been confirmed for $H = K_r$ with $r > 4$. In other words, we do not have the multipartite version of the Hajnal–Szemerédi theorem. This problem was studied by Fischer [5], who obtained an almost perfect tiling for the case of K_3 and K_4. Magyar and Martin [15] proved Conjecture 6.2 for K_3 with $C = 1$; Martin and Szemerédi [16] proved Conjecture 6.2 for K_4 with $C = 0$. Furthermore, Martin and Zhao [17] proved Conjecture 6.2 for all complete tripartite graphs $K_{s,s,s}$. Given the success on the tiling of K_3 and K_4, it may not be very hard to prove Conjecture 6.2 for all 3-chromatic or 4-chromatic H.

ACKNOWLEDGMENTS

The authors thank referee for their valuable comments that improved the presentation.

REFERENCES

[1] N. Alon and R. Yuster, *H*-factors in dense graphs, J Combin Theory Ser B 66 (1996), 269–282.
[2] B. Bollobás and A. D. Scott, Problems and results on judicious partitions, Rand Struc Alg 21 (2002), 414–430.
[3] A. Czygrinow and L. DeBiasio, A note on bipartite graph tiling, submitted.
[4] G. Dirac, Some theorem on abstract graphs, Proc London Math Soc 2 (1952), 69–81.
[5] E. Fischer, Variants of the Hajnal–Szemerédi theorem, J Graph Theory 31(4) (1999), 275–282.
[6] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdös, Combinatorial Theory and Its Applications II (Proceedings of the Colloquium, Balatonfüred, 1969), North-Holland, Amsterdam, 1970, pp. 601–623.
[7] P. Hall, On representatives of subsets, J London Math Soc 10 (1935), 26–30.
[8] J. Hladký and M. Schacht, Note on bipartite graph tilings, SIAM J Discrete Math 24(2) (2010), 357–362.
[9] J. Komlós, Tiling Turán theorems, Combinatorica 20 (2000), 203–218.
[10] J. Komlós, G. N. Sárközy, and E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1997), 109–123.
[11] J. Komlós, G. N. Sárközy, and E. Szemerédi, Proof of the Alon–Yuster conjecture, Combinatorics (Praque, 1998), Discrete Math 235 (2001), 255–269.
[12] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, Combinatorics, Paul Erdős is Eighty, 2 (Keszthely, 1993), Bolyai Society Mathematical Studies 2, The Jnos Bolyai Mathematical Society, Budapest, 1996, pp. 295–352.
[13] D. Kühn and D. Osthus, The minimum degree threshold for perfect graph packings, Combinatorica 29 (2009), 65–107.
[14] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, Surveys in Combinatorics, Cambridge University Press, Cambridge, 2009, pp. 137–167.
[15] Cs. Magyar and R. Martin, Tripartite version of the Corrádi–Hajnal theorem, Discrete Math 254(1–3) (2002), 289–308.
[16] R. Martin and E. Szemerédi, Quadripartite version of the Hajnal–Szemerédi theorem, Discrete Math 308(19) (2008), 4337–4360.
[17] R. Martin and Y. Zhao, Tiling tripartite graphs with 3-colorable graphs, Electron J Combin 16(1) (2009), 16, Research Paper 109.
[18] A. Shokoufandeh and Y. Zhao, Proof of a conjecture of Komlós, Random Structures Algorithms 23(2) (2003), 180–205.
[19] E. Szemerédi, Regular partitions of graphs, Problèmes Combinatoires et Théorie des Graphes (Colloques Internationaux CNRS, University of Orsay, Orsay, 1976), Colloques Internationaux CNRS 260, Colloques Internationaux, Paris, 1978, pp. 399–401.
[20] Y. Zhao, Bipartite graph tiling, SIAM J Disc Math 23(2) (2009), 888–900.