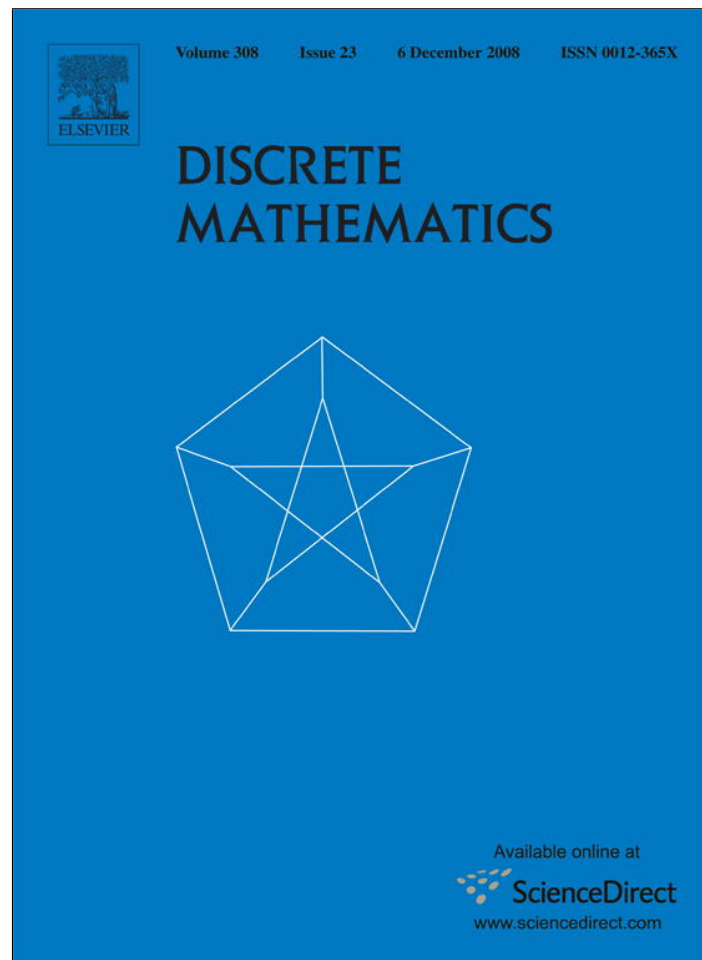


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# Distributing vertices along a Hamiltonian cycle in Dirac graphs

Gábor N. Sárközy<sup>a,b</sup>, Stanley Selkow<sup>a,\*</sup>

<sup>a</sup> *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA*

<sup>b</sup> *Computer and Automation Research Institute, Hungarian Academy of Sciences, P.O. Box 63, Budapest H-1518, Hungary*

Received 26 January 2005; received in revised form 24 April 2007; accepted 26 October 2007

Available online 20 February 2008

---

## Abstract

A graph  $G$  on  $n$  vertices is called a Dirac graph if it has a minimum degree of at least  $n/2$ . The distance  $\text{dist}_G(u, v)$  is defined as the number of edges in a shortest path of  $G$  joining  $u$  and  $v$ . In this paper we show that in a Dirac graph  $G$ , for every small enough subset  $S$  of the vertices, we can distribute the vertices of  $S$  along a Hamiltonian cycle  $C$  of  $G$  in such a way that all but two pairs of subsequent vertices of  $S$  have prescribed distances (apart from a difference of at most 1) along  $C$ . More precisely we show the following. There are  $\omega, n_0 > 0$  such that if  $G$  is a Dirac graph on  $n \geq n_0$  vertices,  $d$  is an arbitrary integer with  $3 \leq d \leq \omega n/2$  and  $S$  is an arbitrary subset of the vertices of  $G$  with  $2 \leq |S| = k \leq \omega n/d$ , then for every sequence  $d_i$  of integers with  $3 \leq d_i \leq d$ ,  $1 \leq i \leq k-1$ , there is a Hamiltonian cycle  $C$  of  $G$  and an ordering of the vertices of  $S$ ,  $a_1, a_2, \dots, a_k$ , such that the vertices of  $S$  are visited in this order on  $C$  and we have

$$|\text{dist}_C(a_i, a_{i+1}) - d_i| \leq 1, \quad \text{for all but one } 1 \leq i \leq k-1.$$

© 2007 Elsevier B.V. All rights reserved.

*Keywords:* Graph theory; Hamiltonian cycle

---

## 1. Introduction

### 1.1. Notation and definitions

For basic graph concepts see the monograph of Bollobás [2].

$V(G)$  and  $E(G)$  denote the vertex set and the edge set of the graph  $G$ .  $(A, B, E)$  denotes a bipartite graph  $G = (V, E)$ , where  $V$  is the disjoint union of  $A$  and  $B$ , and  $E \subset A \times B$ . For a graph  $G$  and a subset  $U$  of its vertices,  $G|_U$  is the restriction to  $U$  of  $G$ .  $N(v)$  is the set of neighbors of  $v \in V$ . Hence the size of  $N(v)$  is  $|N(v)| = \deg(v) = \deg_G(v)$ , the degree of  $v$ .  $\delta(G)$  stands for the minimum degree, and  $\Delta(G)$  for the maximum degree in  $G$ .  $\nu(G)$  is the size of a maximum matching in  $G$ . The distance  $\text{dist}_G(u, v)$  is defined as the number of edges in a shortest path of  $G$  joining  $u$  and  $v$ . For  $S \subset V(G)$  we write  $N(S) = \bigcap_{v \in S} N(v)$ , the set of common neighbors.  $N(x, y, z, \dots)$  is shorthand for  $N(\{x, y, z, \dots\})$ . For a vertex  $v \in V$  and set  $U \subset V - \{v\}$ , we write

---

\* Corresponding author. Tel.: +1 508 831 5357.

*E-mail addresses:* [gsarkozy@cs.wpi.edu](mailto:gsarkozy@cs.wpi.edu) (G.N. Sárközy), [sms@wpi.edu](mailto:sms@wpi.edu) (S. Selkow).

$\deg(v, U)$  for the number of edges from  $v$  to  $U$ . When  $A, B$  are subsets of  $V(G)$ , we denote by  $e(A, B)$  the number of ordered pairs  $(a, b)$  such that  $a \in A, b \in B$  and  $(a, b) \in E(G)$ . For non-empty  $A$  and  $B$ ,

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

is the *density* of the graph between  $A$  and  $B$ . We write

$$d(A) = d(A, A) = 2|E(G|_A)|/|A|^2.$$

**Definition 1.** The bipartite graph  $G = (A, B, E)$  is  $\varepsilon$ -regular if

$$X \subset A, Y \subset B, |X| > \varepsilon|A|, |Y| > \varepsilon|B| \quad \text{imply} \quad |d(X, Y) - d(A, B)| < \varepsilon,$$

otherwise it is  $\varepsilon$ -irregular. Furthermore, it is  $(\varepsilon, d)$ -regular if it is  $\varepsilon$ -regular and  $d(A, B) \geq d$ . Finally,  $(A, B, E)$  is  $(\varepsilon, d)$ -super-regular if it is  $\varepsilon$ -regular and

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

### 1.2. Distributing vertices along a Hamiltonian cycle in Dirac graphs

Let  $G$  be a graph on  $n \geq 3$  vertices. A *Hamiltonian cycle (path)* of  $G$  is a cycle (path) containing every vertex of  $G$ . A *Hamiltonian graph* is a graph containing a Hamiltonian cycle. A graph in which every pair of vertices can be connected with a Hamiltonian path is *Hamiltonian-connected*. A classical result of Dirac [4] asserts that if  $\delta(G) \geq n/2$  (call these graphs *Dirac graphs*), then  $G$  is Hamiltonian. This result of Dirac has generated an incredible amount of research; it has been generalized and strengthened in numerous ways (see the excellent survey of Gould [8]).

In a recent, interesting strengthening of Dirac's Theorem, Kaneko and Yoshimoto [10] showed that in a Dirac graph given any sufficiently small subset  $S$  of vertices, a Hamiltonian cycle  $C$  can be constructed such that there is a uniform lower bound on the distances on  $C$  between successive pairs of vertices of  $S$ .

**Theorem 1.** *Let  $G$  be a graph of order  $n$  with  $\delta(G) \geq n/2$  and let  $d$  be a positive integer with  $d \leq n/4$ . Then for any vertex set  $S$  with at most  $n/2d$  vertices, there exists a Hamiltonian cycle  $C$  with  $\text{dist}_C(u, v) \geq d$  for every  $u$  and  $v$  in  $S$ .*

Note that this result is sharp; the bound on the cardinality of  $S$  cannot be increased.

Gould called for further studies on density conditions that allow the distribution of “small” subsets of vertices along a Hamiltonian cycle (see Problem 1 in [8]). In this paper we show that not only can we have a lower bound on the distances but actually almost all of the distances between successive pairs of vertices of  $S$  can be specified almost exactly. Note that the partitions of graphs into special subgraphs of given size have received attention (see [5,6,9] and [19]). One example is the celebrated El-Zahar conjecture (see [5]), where we partition the graphs into cycles of a given length (instead of paths). The conjecture states that a graph  $G$  of order  $n = n_1 + n_2 + \dots + n_k$  with  $\delta(G) \geq \sum_{i=1}^k \lceil n_i/2 \rceil$  contains a 2-factor  $C_{n_1} \cup \dots \cup C_{n_k}$ . The case  $k = 1$  follows again from Dirac's Theorem, and the case  $k = 2$  was proved by El-Zahar in [5]. Another example is a result of Enomoto and Matsunaga who showed that in a graph  $G$  of order  $n = n_1 + n_2 + \dots + n_k$  with  $n_i \geq 2, \delta(G) \geq 3k - 2$ , for any set of  $k$  vertices  $\{v_1, \dots, v_k\}$  we can find a partition  $V(G) = V_1 \cup \dots \cup V_k$  such that  $|V_i| = n_i, v_i \in V_i$  and  $G|_{V_i}$  has no isolated vertices. This fact has been used to derive the best known error bounds in certain branches of coding theory [3].

Here our main result is the following.

**Theorem 2.** *There are  $\omega, n_0 > 0$  such that if  $G$  is a graph on  $n \geq n_0$  vertices with  $\delta(G) \geq n/2, d$  is an arbitrary integer with  $3 \leq d \leq \omega n/2$  and  $S$  is an arbitrary subset of the vertices of  $G$  with  $2 \leq |S| = k \leq \omega n/d$ , then for every sequence  $d_i$  of integers with  $3 \leq d_i \leq d, 1 \leq i \leq k - 1$ , there is a Hamiltonian cycle  $C$  of  $G$  and an ordering of the vertices of  $S, a_1, a_2, \dots, a_k$ , such that the vertices of  $S$  are visited in this order on  $C$  and we have*

$$|\text{dist}_C(a_i, a_{i+1}) - d_i| \leq 1, \quad \text{for all but one } 1 \leq i \leq k - 1.$$

It would be desirable to eliminate the two discrepancies by 1 from the theorem. However, this is impossible. We need the discrepancies by 1 between  $\text{dist}_C(a_i, a_{i+1})$  and  $d_i$  because of parity reasons. Indeed, consider the complete bipartite graph between  $U$  and  $V$ , where  $|U| = |V| = n/2$ . Take  $S \subset U$ , then the distance between subsequent vertices of  $S$  along a Hamiltonian cycle is even, and if we have an odd  $d_i$  we cannot obtain a distance with that  $d_i$ .

To see that we might need an exceptional  $i$  for which  $|\text{dist}_C(a_i, a_{i+1}) - d_i| > 1$ , consider the following construction. Take two complete graphs on  $U$  and  $V$  with  $|U| = |V| = n/2$ . Let  $S = S' \cup S''$  with  $S' \subset U$ ,  $S'' \subset V$  and  $|S'| = |S''| = |S|/2$ , and add the complete bipartite graphs between  $S'$  and  $V$ , and between  $S''$  and  $U$ . Clearly on any Hamiltonian cycle we will have two distances much greater than  $d$ .

Let us also remark that we need the  $d_i \geq 3$  requirement as in certain parts of our proof (see [Subcase 1.1](#) in Section 5). We use connecting paths of length at least 4 between two vertices of  $S$ .

Finally we believe that our theorem remains true for greater values  $|S|$ 's (perhaps proportional to  $n/d$  as in [Theorem 1](#)) and for greater values of  $d$  as well, but we were unable to prove a stronger statement.

## 2. The main tools

In the proof the Regularity Lemma [23] plays a central role. Here we will use the following variation of the lemma. For the proof, see [18].

**Lemma 1** (*Regularity Lemma – Degree form*). *For every  $\varepsilon > 0$  and every integer  $m_0$  there is an  $M_0 = M_0(\varepsilon, m_0)$  such that if  $G = (V, E)$  is any graph on at least  $M_0$  vertices and  $\delta \in [0, 1]$  is any real number, then there is a partition of the vertex set  $V$  into  $l + 1$  sets (so-called clusters)  $V_0, V_1, \dots, V_l$ , and there is a subgraph  $G' = (V, E')$  with the following properties:*

- $m_0 \leq l \leq M_0$ ,
- $|V_0| \leq \varepsilon|V|$ ,
- all clusters  $V_i, i \geq 1$ , are of the same size  $L$ ,
- $\deg_{G'}(v) > \deg_G(v) - (\delta + \varepsilon)|V|$  for all  $v \in V$ ,
- $G'|_{V_i} = \emptyset$  ( $V_i$  are independent in  $G'$ ),
- all pairs  $G'|_{V_i \times V_j}, 1 \leq i < j \leq l$ , are  $\varepsilon$ -regular, each with a density 0 or exceeding  $\delta$ .

This form can easily be obtained by applying the original Regularity Lemma (with a smaller value of  $\varepsilon$ ), adding to the exceptional set  $V_0$  all clusters incident to many irregular pairs, and then deleting all edges between any other clusters where the edges either do not form a regular pair or they do but with a density of at most  $\delta$ .

An application of the Regularity Lemma in graph theory is now often coupled with an application of the Blow-up Lemma (see [13] for the original, [14] for an algorithmic version and [20,21] for two alternative proofs). Here we use a very special case of the Blow-up Lemma. This asserts that if  $(A, B)$  is a super-regular pair with  $|A| = |B|$  and  $x \in A, y \in B$ , then there is a Hamiltonian path starting with  $x$  and ending with  $y$ . More precisely.

**Lemma 2.** *For every  $\delta > 0$  there are  $\varepsilon_{BL} = \varepsilon_{BL}(\delta), n_{BL} = n_{BL}(\delta) > 0$  such that if  $\varepsilon \leq \varepsilon_{BL}$  and  $n \geq n_{BL}$ ,  $G = (A, B)$  is an  $(\varepsilon, \delta)$ -super-regular pair with  $|A| = |B| = n$  and  $x \in A, y \in B$ , then there is a Hamiltonian path in  $G$  starting with  $x$  and ending with  $y$ .*

We will also use some well-known properties of regular pairs. They can be found in [18]. The first one basically says that every regular pair contains a “large” super-regular pair.

**Lemma 3** ([18, Fact 1.3]). *Let  $(A, B)$  be an  $(\varepsilon, \delta)$ -regular pair and  $B'$  be a subset of  $B$  of size at least  $\varepsilon|B|$ . Then there are at most  $\varepsilon|A|$  vertices  $v \in A$  with  $|N(v) \cap B'| < (\delta - \varepsilon)|B'|$ .*

The next property says that subgraphs of a regular pair are also regular.

**Lemma 4** (*Slicing Lemma, [18, Fact 1.5]*). *Let  $(A, B)$  be an  $(\varepsilon, \delta)$ -regular pair, and, for some  $\beta > \varepsilon$ , let  $A' \subset A, |A'| \geq \beta|A|, B' \subset B, |B'| \geq \beta|B|$ . Then  $(A', B')$  is an  $(\varepsilon', \delta')$ -regular pair with  $\varepsilon' = \max\{\varepsilon/\beta, 2\varepsilon\}$  and  $|\delta' - \delta| < \varepsilon$ .*

We will also use two simple Pósa-type lemmas on Hamiltonian-connectedness. The second one is the bipartite version of the first one.

**Lemma 5** (See [1], Chapter 10, Theorem 13). Let  $G$  be a graph on  $n \geq 3$  vertices with degrees  $d_1 \leq d_2 \leq \dots \leq d_n$  such that for every  $2 \leq k \leq \frac{n}{2}$  we have  $d_{k-1} > k$ . Then  $G$  is Hamiltonian-connected.

**Lemma 6** (See [1], Chapter 10, Theorem 15). Let  $G = (A, B)$  be a bipartite graph with  $|A| = |B| = n \geq 2$  with degrees  $d_1 \leq d_2 \leq \dots \leq d_n$  from  $A$  and with degrees  $d'_1 \leq d'_2 \leq \dots \leq d'_n$  from  $B$ . Suppose that for every  $2 \leq j \leq \frac{n+1}{2}$  we have  $d_{j-1} > j$  and  $d'_{j-1} > j$ . Then  $G$  is Hamiltonian-connected.

Finally we will use the following two simple facts on the sizes of matchings in graphs.

**Lemma 7** (Erdős, Pósa [7], see also [2], Chapter 2, Theorem 4.2). Let  $G$  be a graph on  $n$  vertices. Then

$$\nu(G) \geq \min \left\{ \delta(G), \frac{n-1}{2} \right\}.$$

**Lemma 8.** Let  $G$  be a graph of order  $n$  and let  $S \subset V(G)$  with  $|S| \leq n/2$ . Then in  $G$  there is a matching of size at least

$$\delta(G) \frac{n - |S|}{2(\delta(G) + \Delta(G))} \geq \delta(G) \frac{n}{8\Delta(G)}$$

such that for each matching edge at least one of the endpoints is from  $V(G) \setminus S$ .

**Proof.** Let us take a maximal matching  $M$  with  $m$ -edges with the property that for each matching edge at least one of the endpoints is from  $V(G) \setminus S$ . Then for the number of edges  $E$  between  $V(M)$  and  $V(G) \setminus (V(M) \cup S)$ , we get the estimate

$$\delta(G)(n - 2m - |S|) \leq E \leq 2m\Delta(G).$$

From this we get

$$2m(\delta(G) + \Delta(G)) \geq \delta(G)(n - |S|),$$

which proves the lemma.  $\square$

### 3. Outline of the proof

In this paper we use the Regularity Lemma–Blow-up Lemma method again (see [11–17,22]). The method is usually applied to find certain spanning subgraphs in dense graphs. Typical examples are spanning trees (Bollobás conjecture, see [11]), Hamiltonian cycles or powers of Hamiltonian cycles (Pósa–Seymour conjecture, see [15,16]) or  $H$ -factors for a fixed graph  $H$  (Alon–Yuster conjecture, see [17]).

Let us consider a graph  $G$  of order  $n$  with

$$\delta(G) \geq \frac{n}{2}. \tag{1}$$

We will assume throughout the paper that  $n$  is sufficiently large. We will use the following main parameters

$$0 < \omega \ll \varepsilon \ll \delta \ll \alpha \ll 1, \tag{2}$$

where  $a \ll b$  means that  $a$  is sufficiently small compared to  $b$ . For simplicity, we do not compute the actual dependencies, although it could be done.

Let  $d$  be an arbitrary integer with  $3 \leq d \leq \omega n/2$  and let  $S$  be an arbitrary subset of the vertices of  $G$  with

$$2 \leq |S| = k \leq \omega n/d. \tag{3}$$

Consider an arbitrary sequence  $\underline{d} = \{d_i | 3 \leq d_i \leq d, 1 \leq i \leq k-1\}$ . A cycle  $C$  in  $G$  (or a path  $P$ ) is called an  $(S, \underline{d})$ -cycle (or an  $(S, \underline{d})$ -path) if there is an ordering of the vertices of  $S$ ,  $a_1, a_2, \dots, a_k$ , such that the vertices of  $S$  are visited in this order on  $C$  (on  $P$ ) such that

$$|\text{dist}_C(a_i, a_{i+1}) - d_i| \leq 1, \quad 1 \leq i \leq k-1.$$

We must show that there is a Hamiltonian cycle that is almost an  $(S, \underline{d})$ -cycle, namely we can have

$$|\text{dist}_C(a_i, a_{i+1}) - d_i| > 1$$

for only one  $1 \leq i \leq k - 1$ .

First in the next section, in the non-extremal part of the proof, we show this assuming that the following extremal condition does not hold for our graph  $G$ . We show later in Section 5 that [Theorem 2](#) is true in the extremal case as well.

**Extremal Condition (EC) with parameter  $\alpha$ :** *There exist (not necessarily disjoint)  $A, B \subset V(G)$  such that*

- $|A|, |B| \geq \left(\frac{1}{2} - \alpha\right)n$ , and
- $d(A, B) < \alpha$ .

In the non-extremal case, when  $G$  does not satisfy the EC with parameter  $\alpha$ , we apply the Regularity Lemma ([Lemma 1](#)) for  $G$ , with  $\varepsilon$  and  $\delta$  as in (2). We get a partition of  $V(G') = \cup_{0 \leq i \leq l} V_i$ . We define the following *reduced graph*  $G_r$ : The vertices of  $G_r$  are  $p_1, \dots, p_l$ , and we have an edge between vertices  $p_i$  and  $p_j$  if the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular in  $G'$  with density exceeding  $\delta$ . Thus we have a one-to-one correspondence  $f : p_i \rightarrow V_i$  between the vertices of  $G_r$  and the clusters of the partition. This function  $f$  allows us to move from  $G_r$  to  $G'$  (or  $G$ ). Since in  $G'$ ,  $\delta(G') > \left(\frac{1}{2} - \varepsilon - (\delta + \varepsilon)\right)n = \left(\frac{1}{2} - \delta - 2\varepsilon\right)n$ , an easy calculation shows that in  $G_r$  we have

$$\delta(G_r) \geq \left(\frac{1}{2} - 2\delta\right)l. \tag{4}$$

Indeed, because the neighbors of  $u \in V_i$  in  $G'$  can only be in  $V_0$  and in the clusters which are neighbors of  $p_i$  in  $G_r$ , then for a  $V_i$ ,  $1 \leq i \leq l$  we have:

$$\left(\frac{1}{2} - \delta - 2\varepsilon\right)nL \leq \sum_{u \in V_i} \deg(u) \leq \varepsilon nL + \deg(p_i)L^2.$$

From this using  $\varepsilon \leq \delta/3$  we get inequality (4):

$$\deg_{G_r}(p_i) \geq \left(\frac{1}{2} - \delta - 3\varepsilon\right) \frac{n}{L} \geq \left(\frac{1}{2} - 2\delta\right)l.$$

Applying [Lemma 7](#) we can find a matching  $M$  in  $G_r$  of size at least  $\left(\frac{1}{2} - 2\delta\right)l$ . Put  $|M| = m$ . Let us put the vertices of the clusters not covered by  $M$  into the exceptional set  $V_0$ . For simplicity  $V_0$  still denotes the resulting set. Then

$$|V_0| \leq 4\delta lL + \varepsilon n \leq 5\delta n. \tag{5}$$

Denote the  $i$ th pair in  $f(M)$  by  $(V_1^i, V_2^i)$  for  $1 \leq i \leq m$ .

The rest of the non-extremal case is organized as follows. In [Section 4.1](#) first we find an  $(S, \underline{d})$ -path  $P$  where actually  $\text{dist}(a_i, a_{i+1}) = d_i$  for all  $1 \leq i \leq k - 1$ . Then in [Section 4.2](#) we find short connecting paths  $P_i$  between the consecutive edges in the matching  $f(M)$  (for  $i = m$  the next edge is  $i = 1$ ). The first connecting path  $P_1$  between  $(V_1^1, V_2^1)$  and  $(V_1^2, V_2^2)$  will also contain  $P$ , each of the others has length exactly 3. In [Section 4.3](#) we will take care of the exceptional vertices and make some adjustments by extending some of the connecting paths so that the distribution of the remaining vertices inside each edge in  $f(M)$  is perfect, i.e., there are the same number of vertices left in both clusters of the edge. Finally applying [Lemma 2](#) we close the Hamiltonian cycle in each edge thus giving a Hamiltonian  $(S, \underline{d})$ -cycle where  $\text{dist}(a_i, a_{i+1}) = d_i$  for all  $1 \leq i \leq k - 1$  (thus in the non-extremal case both discrepancies by 1 in [Theorem 2](#) are eliminated). Note that the material of [Sections 4.2](#) and [4.3](#) is fairly standard by now and is independent of the problem of prescribed distances. Similar arguments have appeared in other works (see e.g. [[16,17](#)] and [[22](#)]). For the sake of completeness we present the full proof here, but the readers familiar with this technique may skip these sections.

#### 4. The non-extremal case

Throughout this section we assume that the extremal condition with parameter  $\alpha$  does not hold for  $G$ . We apply the Regularity Lemma (Lemma 1) for  $G$  with  $\varepsilon$  and  $\delta$  given in (2) and  $m_0 = 1/\varepsilon$ , define the reduced graph  $G_r$ , and find the matching  $M$  in  $G_r$  as described above in the outline.

##### 4.1. Finding an $(S, \underline{d})$ -path

We are going to use the following fact repeatedly.

**Fact 1.** For any distinct  $x, y \in V(G)$  there are at least  $\delta n$  internally disjoint paths of length 3 in  $G$  connecting  $x$  and  $y$ .

**Proof.** Indeed, using (1) we may choose  $A \subset N_G(x)$  with  $|A| = \lfloor \frac{n}{2} \rfloor$  and  $B \subset N_G(y)$  with  $|B| = \lfloor \frac{n}{2} \rfloor$ . The fact that the EC with parameter  $\alpha$  does not hold for  $G$  implies  $d(A, B) \geq \alpha$ . From this it follows that we have at least  $\frac{\alpha}{2}|A|$  vertices  $v$  in  $A$ , for which we have  $\deg(v, B) \geq \frac{\alpha}{2}|B|$ . Indeed, otherwise we would have

$$d(A, B) = \frac{e(A, B)}{|A||B|} < \frac{\frac{\alpha}{2}|A||B| + \frac{\alpha}{2}|A||B|}{|A||B|} = \alpha,$$

a contradiction. Then using  $\delta \leq \alpha/20$  we can select greedily a matching of size at least

$$\frac{\alpha}{5}|B| = \frac{\alpha}{5} \lfloor \frac{n}{2} \rfloor \geq \frac{\alpha}{20}n \geq \delta n,$$

such that each edge has one endpoint in  $A$  and one endpoint in  $B$ , these endpoints are vertex disjoint from  $\{x, y\}$ , and from this Fact 1 follows. Indeed, take a vertex  $v \in A$  with  $\deg(v, B) \geq \frac{\alpha}{2}|B|$ , and select one of these edges to  $B$  as the first matching edge. Remove this edge, and apply this repeatedly in the leftover, namely take a vertex  $v' \in A$  with  $\deg(v', B) \geq \frac{\alpha}{2}|B|$ , and select one of the remaining edges to  $B$  as a matching edge. As long as the matching that we have so far covers fewer than  $\frac{\alpha}{2}|B|$  vertices, we can select the next matching edge. We remove the at most 2 matching edges that have a non-empty intersection with  $\{x, y\}$ , and we get the desired matching of size at least  $\frac{\alpha}{5}|B|$ .  $\square$

We construct an  $(S, \underline{d})$ -path  $P = Q_1 \cup \dots \cup Q_{k-1}$  in the following way. Let  $a_1, \dots, a_k$  be the vertices of  $S$  in an arbitrary order (so note that here actually we can prescribe the order of the vertices of  $S$  as well). First we construct a path  $Q_1$  of length  $d_1$  connecting  $a_1$  and  $a_2$ . Using the minimum degree condition (1), we construct greedily a path  $Q'_1$  starting from  $a_1$  that has length  $d_1 - 3$  (note that  $d_1 \geq 3$ ). Denote the other end point of  $Q'_1$  by  $a'_1$ . Applying Fact 1, we connect  $a'_1$  and  $a_2$  by a path  $Q''_1$  of length 3 that is internally disjoint from  $Q'_1$ . Then  $Q_1 = Q'_1 \cup Q''_1$  is a path connecting  $a_1$  and  $a_2$  with length  $d_1$ .

We iterate this procedure. For the construction of  $Q_2$ , first we greedily construct a path  $Q'_2$  starting from  $a_2$  that is internally disjoint from  $Q_1$  and has length  $d_2 - 3$ . Denote the other end point of  $Q'_2$  by  $a'_2$ . Applying Fact 1, we connect  $a'_2$  and  $a_3$  by a path  $Q''_2$  of length 3 that is internally disjoint from  $Q_1 \cup Q'_2$ . Then  $Q_2 = Q'_2 \cup Q''_2$  is a path connecting  $a_2$  and  $a_3$  with length  $d_2$ .

By iterating this procedure we get an  $(S, \underline{d})$ -path  $P$ . (1)–(3) and Fact 1 imply that we never get stuck since

$$|V(P)| = 1 + \sum_{i=1}^{k-1} d_i \leq 1 + (k-1)d \leq \omega n \ll \delta n. \tag{6}$$

Observe that here in the non-extremal case there is no discrepancy between  $\text{dist}(a_i, a_{i+1})$  and  $d_i$  for all  $1 \leq i \leq k-1$ , and furthermore we can construct an  $(S, \underline{d})$ -path for any ordering of the vertices of  $S$ .

##### 4.2. Connecting paths

The first connecting path  $P_1$  between  $(V_1^1, V_2^1)$  and  $(V_1^2, V_2^2)$  will include as a subpath the  $(S, \underline{d})$ -path  $P$ . To construct this  $P_1$  first by using Fact 1 we connect a typical vertex  $u$  of  $V_2^1$  (more precisely a vertex  $u$  with  $\deg(u, V_1^1) \geq (\delta - \varepsilon)L$ , most vertices in  $V_2^1$  satisfy this by Lemma 3) and  $a_1$  with a path of length 3. Then we connect  $a_k$  and a typical vertex  $w$  of  $V_1^2$  (so  $\deg(w, V_2^2) \geq (\delta - \varepsilon)L$ ) with a path of length 3. To construct the second

connecting path  $P_2$  between  $(V_1^2, V_2^2)$  and  $(V_1^3, V_2^3)$  we just connect a typical vertex of  $V_2^2$  and a typical vertex  $V_1^3$  with a path of length 3 that is vertex disjoint from  $P_1$ . Continuing in this fashion, finally we connect a typical vertex of  $V_2^m$  with a typical vertex of  $V_1^1$  with a path of length 3 that is vertex disjoint from all the other connecting paths. Thus  $P_1$  has length at most  $\omega n + 6$ , all other  $P_i$ 's have length 3. Note that we can always find these connecting paths that are vertex disjoint from the connecting paths constructed so far. Indeed, the total number of vertices in the union of these paths is at most

$$2\omega n + 4l \leq \frac{\varepsilon}{2} \frac{n}{2M_0} + 4M_0 \leq \frac{\varepsilon}{2} \frac{n}{2M_0} + \frac{\varepsilon}{2} \frac{n}{2M_0} = \varepsilon \frac{n}{2M_0} \leq \varepsilon \frac{n}{2l} \leq \varepsilon L,$$

using  $\omega \leq \frac{\varepsilon}{8M_0}$  and  $n \geq \frac{16M_0^2}{\varepsilon}$ . Then we can find endpoints for the next connecting path that are vertex disjoint from the connecting paths constructed so far since from every cluster (of size  $L$ ) we used up only at most  $\varepsilon L$  vertices, so most of the typical vertices from a cluster are still available. Furthermore, when applying Fact 1 to connect the endpoints, since  $\varepsilon L \leq \varepsilon n \leq \frac{\delta}{2}n$  we still have  $\frac{\delta}{2}n$  internally disjoint paths of length 3 connecting the endpoints that are vertex disjoint from the connecting paths constructed so far.

We remove the internal vertices of these connecting paths from the clusters, but for simplicity we keep the notation for the resulting clusters. These connecting paths will be subpaths of the final Hamiltonian cycle. If the number of remaining vertices (in the clusters and in  $V_0$ ) is odd, then we take another typical vertex  $w$  of  $V_1^2$  and we extend  $P_1$  by a path of length 3 that ends with  $w$ . This way we decreased the number of vertices by 3, so we may always assume that the number of remaining vertices is even. Note that by removing vertices we might have created discrepancies between the sizes of the clusters in an edge of  $f(M)$ , this will be adjusted later at the end of the non-extremal case.

#### 4.3. Adjustments and the handling of the exceptional vertices

Let us note again that the material of this section is fairly routine in this kind of proofs. For the sake of completeness we present the full proof, but the reader familiar with this technique may skip this section.

We already have an exceptional set  $V_0$  of vertices in  $G$ . We add some more vertices to  $V_0$  to achieve super-regularity. From  $V_1^i$  (and similarly from  $V_2^i$ ) we remove all vertices  $u$  for which  $\deg(u, V_2^i) < (\delta - \varepsilon)L$ .  $\varepsilon$ -regularity and Lemma 3 guarantee that at most  $\varepsilon L$  such vertices exist in each cluster  $V_1^i$ .

Thus using (5) and  $\varepsilon \leq \delta$ , we still have

$$|V_0| \leq 5\delta n + \varepsilon n \leq 6\delta n. \tag{7}$$

Since we are looking for a Hamiltonian cycle, we have to include the vertices of  $V_0$  on the Hamiltonian cycle as well. We are going to extend some of the connecting paths  $P_i$ , so now they are going to contain the vertices of  $V_0$ . Let us consider the first vertex (in an arbitrary ordering of the vertices in  $V_0$ )  $v$  in  $V_0$ . We find a pair  $(V_1^i, V_2^i)$  such that either

$$\deg(v, V_1^i) \geq \delta L, \tag{8}$$

in which case we say that  $v$  and  $V_1^i$  are *friendly*, or

$$\deg(v, V_2^i) \geq \delta L, \tag{9}$$

in which case we say that  $v$  and  $V_2^i$  are friendly. In case (8) holds we assign  $v$  to the cluster  $V_2^i$ , and in case (9) holds we assign  $v$  to the cluster  $V_1^i$ . In case (8) holds we extend  $P_{i-1}$  (for  $i = 1, P_m$ ) inside the pair  $(V_1^i, V_2^i)$  by a path of length 3, and in case (9) holds we extend  $P_i$  inside the pair  $(V_1^i, V_2^i)$  by a path of length 3, so that now in both cases the paths end with  $v$ . Indeed, in case (8) holds (it is similar for (9)) consider the endpoint  $w$  of  $P_{i-1}$  in  $V_1^i$ . Choosing  $X = N(w) \cap V_2^i$  and  $Y = N(v) \cap V_1^i$ , by (8), the fact that  $w$  was typical and  $\varepsilon \leq \delta/3$  we can apply the regularity condition for  $X$  and  $Y$ , so in particular we have  $d(X, Y) \geq \delta - \varepsilon$ . Then we can take an arbitrary edge  $(v_1, v_2)$  between  $X$  and  $Y$  and then  $(w, v_1, v_2, v)$  gives us the desired extension of  $P_{i-1}$ .

To finish the procedure for  $v$ , in case (8) holds we add one more vertex  $v'$  to  $P_{i-1}$  after  $v$  such that  $(v, v') \in E(G)$  and  $v'$  is a typical vertex of  $V_1^i$ , so  $\deg(v', V_2^i) \geq (\delta - \varepsilon)L$ . In case (9) holds we add one more vertex  $v'$  to  $P_i$  before  $v$  such that  $(v, v') \in E(G)$ ,  $v'$  is a typical vertex of  $V_2^i$ . Thus now  $v$  is included as an internal vertex on the extended connecting path  $P_{i-1}$  or  $P_i$ .

After handling  $v$ , we repeat the same procedure for the other vertices in  $V_0$ . However, we have to pay attention to several technical details. First, of course in repeating this procedure we always consider the remaining vertices in each cluster; the internal vertices on the extended connecting paths are always removed. For simplicity we keep the notation. Note that the number of remaining vertices is always even during the whole process.

Second, we make sure that we never assign too many vertices of  $V_0$  to any cluster, and thus we never use up too many vertices from any cluster in the matching. First we claim that each  $v \in V_0$  is friendly with at least  $l/4$  clusters in the matching. Indeed, assume for a contradiction that there were only  $c < l/4$  friendly clusters for a  $v \in V_0$ . Then, since  $v$  has fewer than  $\delta L$  neighbors in clusters that are not friendly with  $v$ , using (7) and  $\varepsilon < \delta < \frac{1}{56}$  we have

$$\deg_G(v) < cL + (2m - c)\delta L + |V_0| \leq \frac{l}{4}L + \delta lL + 6\delta n \leq \left(\frac{1}{4} + 7\delta\right)n \leq \frac{3}{8}n < \frac{n}{2},$$

which is a contradiction to (1). We assign the vertices  $v \in V_0$  as evenly as possible to the pairs (in the matching) of the friendly clusters. Since each vertex  $v \in V_0$  has at least  $l/4$  friendly clusters, each cluster gets assigned at most  $\frac{4|V_0|}{l}$  vertices  $v \in V_0$ . However, as this is proportional to  $\delta L$ , this creates an additional problem, namely as we keep removing vertices we might lose the super-regularity property inside the matching edges, in the worst case it would be possible that we used up all the  $\delta L$  neighbors of a vertex in the other set. Note, that we never lose  $\varepsilon$ -regularity, the Slicing Lemma (Lemma 4 with  $\beta = 1/2$ ) implies that as long as we still have at least half of the vertices remaining in both clusters, the remaining pair is still  $(2\varepsilon, \delta/2)$ -regular.

Therefore, we do the following periodic *super-regularity updating* procedure inside the pairs. After removing  $\lfloor \frac{\delta}{8}L \rfloor$  vertices from a pair  $(V_1^i, V_2^i)$ , we do the following. In the pair  $(V_1^i, V_2^i)$  (that is still  $(2\varepsilon, \delta/2)$ -regular) we find all vertices  $u$  from  $V_1^i$  (and similarly from  $V_2^i$ ) for which  $\deg(u, V_2^i) < (\frac{\delta}{2} - 2\varepsilon)|V_2^i|$  (where we consider only the remaining vertices). Consider one such vertex  $u$ . Similarly as in the way of handling  $v \in V_0$  using  $\varepsilon$ -regularity we extend the connecting path  $P_{i-1}$  or  $P_i$  by a path of length 4 inside the pair (using two vertices from both clusters of the pair so we do not change the difference between the sizes of the clusters in the pair; this fact will be important later) so that it now includes  $u$  as an internal vertex (here  $u$  plays the role of  $v \in V_0$  in the above). By iterating this procedure we can eliminate all of these exceptional  $u$  vertices. Then between two updates in a pair  $(V_1^i, V_2^i)$ , for the degrees of vertices  $u \in V_1^i$  (and similarly in  $V_2^i$ ) we always have

$$\deg(u, V_2^i) \geq \left(\frac{\delta}{2} - 2\varepsilon\right)|V_2^i| - \frac{\delta}{8}L \geq \left(\frac{\delta}{2} - 2\varepsilon\right)\frac{L}{2} - \frac{\delta}{8}L = \left(\frac{\delta}{8} - \varepsilon\right)L \geq \frac{\delta}{16}L,$$

and thus we maintain a super-regularity condition. Furthermore, Lemma 3 implies that we find at most  $2\varepsilon L$  exceptional vertices in one cluster in one update. Thus during the whole process the total number of vertices that we use up from a cluster with this super-regularity updating procedure is at most  $\frac{64\varepsilon}{\delta}L \leq \delta L$  using  $\varepsilon \leq \frac{\delta^2}{64}$ .

Returning to the  $V_0$ -vertices, using (7), each cluster gets assigned at most  $\frac{4|V_0|}{l} \leq 24\delta n/l \leq 25\delta L$  vertices from  $V_0$  during the whole process. Note that in order to handle an assigned  $V_0$ -vertex we have to use at most 2 additional vertices from both clusters of the pair where the vertex was assigned. Thus, we use up at most  $100\delta L$  vertices from each cluster for handling the vertices in  $V_0$  and an additional at most  $\delta L$  vertices in any other way (super-regularity updating procedure, connecting paths and the exceptional vertices removed in the beginning), so altogether we used up at most  $101\delta L$  vertices from each cluster.

After we are done with this, in the remainder of each pair  $(V_1^i, V_2^i)$  we have  $|V_1^i|, |V_2^i| \geq (1 - 101\delta)L (\geq L/2)$  (using  $\delta \leq 1/202$ ) and the pair is still  $(2\varepsilon, \delta/16)$ -super-regular. At this point we might have a small difference ( $\leq 101\delta L$ ) between the number of remaining vertices in  $V_1^i$  and in  $V_2^i$  in a pair. Therefore, we have to make some adjustments. For this purpose we will need some facts about  $G_r$ . First we will show that  $G_r$  satisfies structural properties similar to that of  $G$ .

**Fact 2.** *EC with parameter  $\alpha/2$  does not hold for  $G_r$ .*

Indeed, otherwise suppose for a contradiction that there are  $A, B \subset V(G_r)$  such that  $|A|, |B| \geq \left(\frac{1}{2} - \frac{\alpha}{2}\right)l$  and  $d_{G_r}(A, B) < \frac{\alpha}{2}$ . We will show that in this case the EC with parameter  $\alpha$  would hold for  $G$  as well, a contradiction.

Consider  $f(A)$  and  $f(B)$ . We have  $f(A), f(B) \subset V(G)$  with

$$|f(A)|, |f(B)| \geq \left(\frac{1}{2} - \frac{\alpha}{2}\right) (1 - \varepsilon)n \geq \left(\frac{1}{2} - \alpha\right)n,$$

giving the first condition in the definition of EC with parameter  $\alpha$ . For the second condition in the definition, concerning the number of edges in  $G$  between  $f(A)$  and  $f(B)$  we get the following upper bound.

$$\begin{aligned} |E(G|_{f(A) \times f(B)})| &< \frac{\alpha}{2}|f(A)||f(B)| + (\delta + \varepsilon)|f(A)|n \\ &\leq \frac{\alpha}{2}|f(A)||f(B)| + 6\delta|f(A)||f(B)| < \alpha|f(A)||f(B)| \end{aligned}$$

(using  $\delta < \alpha/12$ ). Here the first term comes from the edges in  $G'$  between  $f(A)$  and  $f(B)$  (they must come from  $G_r$ -edges), and the second term comes from the edges in  $G \setminus G'$  between  $f(A)$  and  $f(B)$ . Thus indeed EC with parameter  $\alpha$  would hold for  $G$ , a contradiction, proving Fact 2.

The next fact will be similar to Fact 1.

**Fact 3.** For any (not necessarily distinct)  $p, q \in V(G_r)$  there are at least  $\frac{\alpha}{90}l$  internally disjoint alternating (with respect to edges in  $M$ ) paths (cycles if  $p = q$ ) of length 5 connecting  $p$  and  $q$ , where the  $M$ -edges are the 2nd and 4th edges along the paths.

Indeed, consider the sets  $N_{G_r}(p) \cap V(M)$  and  $N_{G_r}(q) \cap V(M)$ . Let us denote by  $A$  the pairs (in  $M$ ) of the clusters in the first set and by  $B$  the pairs of the clusters in the second set. From (4) and  $m = |M| \geq \left(\frac{1}{2} - 2\delta\right)l$  we have  $|A|, |B| \geq \left(\frac{1}{2} - 6\delta\right)l$ . Using  $\delta < \alpha/12$  and Fact 2 we know that  $d_{G_r}(A, B) \geq \alpha/2$ . Then, as in the proof of Fact 1 we can select a matching  $M'$  of size at least  $\frac{\alpha}{10}|B| \geq \frac{\alpha}{30}l$  from  $A$  to  $B$ . By throwing away some edges from  $M'$ , we can find a matching  $M''$  of size at least  $\frac{\alpha}{90}l$  from  $A$  to  $B$  such that for any edge  $e \in M$  we have at most one edge of  $M''$  that is incident to  $e$ . Then the statement of Fact 3 follows. Note that we have

$$\frac{\alpha}{90}l \geq \frac{\alpha}{90}m_0 = \frac{\alpha}{90\varepsilon},$$

so using (2) there are quite many paths guaranteed by Fact 3.

With these preparations, let us take a pair  $(V_1^i, V_2^i)$  with a difference  $\geq 2$  (if one such pair exists), say  $|V_1^i| \geq |V_2^i| + 2$  (only the remaining vertices are considered). Using Fact 3 with  $p = q = f^{-1}(V_1^i)$  we can find an alternating path in  $G_r$  of length 5 starting and ending with  $f^{-1}(V_1^i)$ . Let us denote this path by

$$f^{-1}(V_1^i), p_1, p'_1, p_2, p'_2, f^{-1}(V_1^i)$$

where  $(p_1, p'_1)$  and  $(p_2, p'_2)$  are edges in matching  $M$  (and thus they correspond to super-regular pairs), and the other 3 edges of the path are edges in  $G_r$  (and thus they correspond to regular pairs). We remove a typical vertex  $u_1$  from  $V_1^i$  (a vertex for which  $\deg(u_1, f(p_1)) \geq \left(\frac{\delta}{2} - 2\varepsilon\right)|f(p_1)|$ , most of the remaining vertices satisfy this in  $V_1^i$ ) and add it to  $f(p'_1)$  (and thus we preserve the super-regularity of  $(f(p_1), f(p'_1))$ ). We remove a typical vertex  $u_2$  from  $f(p'_1)$  (a vertex for which  $\deg(u_2, f(p_2)) \geq \left(\frac{\delta}{2} - 2\varepsilon\right)|f(p_2)|$ , most of the remaining vertices satisfy this in  $f(p'_1)$ ) and add it to  $f(p'_2)$  (and thus we preserve the super-regularity of  $(f(p_2), f(p'_2))$ ). Finally we remove a typical vertex  $u_3$  from  $f(p'_2)$  (a vertex for which  $\deg(u_3, V_1^i) \geq \left(\frac{\delta}{2} - 2\varepsilon\right)|V_1^i|$ , most remaining vertices satisfy this in  $f(p'_2)$ ) and add it to  $V_2^i$  (and thus we preserve the super-regularity of  $(V_1^i, V_2^i)$ ). Furthermore, similarly as above in the super-regularity updating when we add a new vertex to a pair  $(V_1^j, V_2^j)$ , using  $\varepsilon$ -regularity we extend the connecting path  $P_{j-1}$  or  $P_j$  by a path of length 4 inside the pair (using two vertices from both clusters of the pair so that we do not change the difference between the sizes of the clusters in the pair) so that it now includes the new vertex as an internal vertex. Thus the overall effect of these changes is that the difference  $|V_1^i| - |V_2^i|$  decreases by 2, but the other differences  $|V_1^j| - |V_2^j|$  do not change for  $1 \leq j \leq m, j \neq i$ .

Now we are one step closer to the perfect distribution, and by iterating this procedure we can assure that the difference in every pair is at most 1. However, similarly as above we have to make sure that we never use up too many vertices from each cluster in this part of the procedure. Note that altogether we use up at most  $10^4\delta n$  vertices in this

part of the procedure. We declare a cluster *forbidden* if we used up  $\alpha L$  vertices from that cluster. Then from Fact 3 it follows that we can always find an alternating path that does not contain any forbidden clusters assuming  $\delta \leq \alpha^2/10^6$ . Furthermore, as above we perform periodically the super-regularity update inside each pair.

Thus we may assume that the difference in every pair  $(V_1^i, V_2^i)$  is at most 1. We consider only those pairs for which the difference is exactly 1, so in particular the number of remaining vertices in one such pair is odd. Since we have an even number of vertices left, it follows that we have an even number of such pairs. We pair up these pairs arbitrarily. If  $(V_1^i, V_2^i)$  and  $(V_1^j, V_2^j)$  is one such pair with  $|V_1^i| = |V_2^i| + 1$  and  $|V_1^j| = |V_2^j| + 1$  (otherwise similar), then similar to the construction above, we find an alternating path in  $G_r$  of length 5 between  $V_1^i$  and  $V_1^j$ , and we move a typical vertex of  $V_1^i$  through the intermediate clusters to  $V_2^j$ .

Thus we may assume that the distribution is perfect, in every pair  $(V_1^i, V_2^i)$  we have the same number of vertices ( $\geq (1 - 2\alpha)L \geq L/2$ ) left in both clusters and all the pairs are still  $(2\epsilon, \delta/16)$ -super-regular. Then using (2) all the conditions of Lemma 2 are satisfied and then Lemma 2 closes the Hamiltonian cycle in every pair.

### 5. The extremal case

For the extremal case, first we will deal with two special cases. In Case 1,  $G$  contains an almost complete bipartite graph. In Case 2,  $G$  contains the union of two almost complete graphs. Finally we will show that the extremal case reduces to one of these two cases.

**Case 1.** Assume that there is a partition  $V(G) = A_1 \cup A_2$  with  $(\frac{1}{2} - \alpha)n \leq |A_1| \leq \frac{n}{2}$  and  $d(A_1) < \alpha^{1/3}$ .

Note that in this case from (1) we also have  $d(A_1, A_2) > 1 - 2\alpha^{1/3}$ . Thus, roughly speaking in this case we have very few edges in  $G|_{A_1}$ , and we have an almost complete bipartite graph between  $A_1$  and  $A_2$ .

In this case a vertex  $v \in A_i, i \in \{1, 2\}$ , is called *exceptional* if it is not connected to most of the vertices in the other set, more precisely if we have

$$\deg(v, A_{i'}) \leq (1 - 2\alpha^{1/6})|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Note that since  $d(A_1, A_2) > 1 - 2\alpha^{1/3}$ , the number of exceptional vertices in  $A_i$  is at most  $\alpha^{1/6}|A_i|$ . We remove the exceptional vertices from each set and we redistribute them in such a way that  $e(A_1, A_2)$  is maximized. We still denote the resulting sets by  $A_1$  and  $A_2$ . Assume that  $|A_1| \leq |A_2|$ , so  $|A_2| - |A_1| = r$ , where  $0 \leq r \leq 3\alpha^{1/6}|A_2|$ . It is easy to see that in  $G|_{A_1 \times A_2}$  we certainly have the following degree conditions. Apart from at most  $3\alpha^{1/6}|A_i|$  exceptional vertices for all vertices  $v \in A_i, i \in \{1, 2\}$  we have

$$\deg(v, A_{i'}) \geq (1 - 4\alpha^{1/6})|A_{i'}|, \quad \{i, i'\} = \{1, 2\},$$

and for the exceptional vertices  $v \in A_i, i \in \{1, 2\}$  we have

$$\deg(v, A_{i'}) \geq \frac{|A_{i'}|}{3}, \quad \{i, i'\} = \{1, 2\}.$$

Thus note that in  $G|_{A_1 \times A_2}$  the degrees of the exceptional vertices are certainly much more than the number of these exceptional vertices, so the degree conditions of Lemma 6 are satisfied with much room to spare. However,  $|A_1|$  may not be equal to  $|A_2|$ .

Our goal is to achieve  $|A_1| = |A_2|$  (if this is not true already). Thus we do the following. If  $|A_1| < |A_2|$  and there is a vertex  $x \in A_2$  for which

$$\deg(x, A_2) \geq \alpha^{1/7}|A_2|, \tag{10}$$

then we remove  $x$  from  $A_2$  and add it to  $A_1$ . We iterate this procedure until either there are no more vertices in  $A_2$  satisfying (10) or  $|A_1| = |A_2|$ .

**Subcase 1.1.** We have  $|A_1| < |A_2|$ , but there are no more vertices in  $A_2$  satisfying (10). Since we have  $\delta(G|_{A_2}) \geq \frac{r}{2}$  (using (1)) and  $\Delta(G|_{A_2}) < \alpha^{1/7}|A_2|$  (since there are no more vertices in  $A_2$  satisfying (10)), applying Lemma 8 for  $G|_{A_2}$  and using (2) we get that  $G|_{A_2}$  has an  $r$ -matching  $M_r$  denoted by  $\{u_1, v_1\}, \dots, \{u_r, v_r\}$  such that for every edge

in  $M_r$  at least one of the end points (say  $u_i$ ) is not in  $S$ . This matching  $M_r$  will be used to balance the discrepancy between  $|A_1|$  and  $|A_2|$ .

Note that in  $G|_{A_1 \times A_2}$  the degrees of the exceptional vertices (and now we have exceptional vertices only in  $A_1$ ) are still much more than the number of these exceptional vertices since for  $\alpha \ll 1$  we have  $\alpha^{1/6} \ll \alpha^{1/7}$ . These degree conditions and (2) imply the following fact (similar to Fact 1).

**Fact 4.** For any distinct  $x, y \in A_2$  there are at least  $n/4$  internally disjoint paths of length 2 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ . For any distinct  $x, y \in A_1$  there are at least  $\frac{\alpha^{1/7}}{8}n$  internally disjoint paths of length 4 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ . Finally for any  $x$  and  $y$  that are in different sets, say  $x \in A_1$  and  $y \in A_2$ , there are at least  $\frac{\alpha^{1/7}}{8}n$  internally disjoint paths of length 3 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ .

**Proof.** Indeed for the first statement note that from  $\Delta(G|_{A_2}) < \alpha^{1/7}|A_2|$  we get  $\deg(x, A_1), \deg(y, A_1) \geq (1 - 2\alpha^{1/7})|A_1|$  and thus the number of common neighbors of  $x$  and  $y$  in  $A_1$  is at least  $(1 - 4\alpha^{1/7})|A_1| \geq n/4$  and from this the first statement follows. For the second statement consider two disjoint equal-size subsets from the two neighborhoods  $X \subset N(x, A_2)$  and  $Y \subset N(y, A_2)$  with  $|X| = |Y| \geq \frac{\alpha^{1/7}}{4}|A_2| \geq \frac{\alpha^{1/7}}{8}n$  (using  $\deg(x, A_2), \deg(y, A_2) \geq \frac{\alpha^{1/7}}{2}|A_2|$ ). Pairing up the vertices between  $X$  and  $Y$  arbitrarily and applying the first statement for each pair we get the second statement. Finally for the last statement consider a subset of the neighborhood  $X \subset N(x, A_2)$  with  $|X| \geq \frac{\alpha^{1/7}}{8}n$ . Applying the first statement for each of the pairs  $(y, u)$  where  $u \in X$  we get the last statement.  $\square$

Let  $S$  be an arbitrary subset of the vertices of  $G$ , satisfying (3), to be distributed along the Hamiltonian cycle at approximately the specified distances. Let us take an arbitrary ordering  $a_1, a_2, \dots, a_k$  of the vertices in  $S$ . In this subcase we construct the desired Hamiltonian cycle in the following way. First, by using Fact 4 repeatedly and a similar procedure as in Section 4.1 we find in  $G|_{A_1 \times A_2}$  an  $(S, \underline{d})$ -path

$$P = P(a_1, a_k) = Q_1 \cup \dots \cup Q_{k-1}$$

connecting the vertices  $a_1$  and  $a_k$ . The only difference from Section 4.1 is that here because of parity reasons we might have  $\text{dist}_C(a_i, a_{i+1}) = d_i + 1$ . Indeed, first we construct a path  $Q_1$  of length  $d_1$  or  $d_1 + 1$  connecting  $a_1$  and  $a_2$ . If  $a_1$  is covered by an edge of  $M_r$ , say  $a_1 = v_i$ , then we start  $Q_1$  with the edge  $\{v_i, u_i\}$  (note that  $u_i \notin S$ ). If  $d_1 = 3$ , then to get  $Q_1$  we connect  $u_i$  and  $a_2$  in  $G|_{A_1 \times A_2}$  by a path of length 2 in case  $a_2 \in A_2$ , and by a path of length 3 in case  $a_2 \in A_1$ . If  $d_1 > 3$ , then we greedily construct a path  $Q'_1$  that has length  $d_1 - 3$ , starts with the edge  $\{v_i, u_i\}$  and continues in  $G|_{A_1 \times A_2}$ . Denote the other end point of  $Q'_1$  by  $a'_1$ . Applying Fact 4, we connect  $a'_1$  and  $a_2$  by a path  $Q''_1$  of length 3 in case they are in different sets, and by a path of length 4 in case they are in the same set. Then  $Q_1 = Q'_1 \cup Q''_1$  is a path connecting  $a_1$  and  $a_2$  with length  $d_1$  or  $d_1 + 1$ .

We iterate this procedure; we construct  $Q_2, \dots, Q_{k-1}$  similarly and thus we get  $P = Q_1 \cup \dots \cup Q_{k-1}$ . Say, the remaining edges of  $M_r$  which are not traversed by  $P$  are

$$\{u_{i_1}, v_{i_1}\}, \dots, \{u_{i_{r'}}, v_{i_{r'}}\} \quad \text{for } 0 \leq r' \leq r.$$

Then we connect the end point  $a_k$  of  $P$  and  $u_{i_1}$  by a path  $R_1$  of length 2 or 3, connect  $v_{i_1}$  and  $u_{i_2}$  by a path  $R_2$  of length 2, etc. Finally connect  $v_{i_{r'-1}}$  and  $u_{i_{r'}}$  by a path  $R_{r'}$  of length 2. Consider the following path.

$$P' = (P, R_1, \{u_{i_1}, v_{i_1}\}, R_2, \{u_{i_2}, v_{i_2}\}, \dots, R_{r'}, \{u_{i_{r'}}, v_{i_{r'}}\}).$$

Note that we never get stuck in the construction of this path; namely when applying Fact 4 we can always choose paths that are internally disjoint from the path that has been constructed so far, since using (2) we have

$$|V(P')| \leq 1 + \sum_{i=1}^{k-1} (d_i + 1) + 4r \leq 2\omega n + 12\alpha^{1/6}n \ll \alpha^{1/7}n.$$

In case  $a_1 \in A_2$ , add one more vertex from  $A_1$  to the end of the path  $P'$ . Remove  $P'$  from  $G|_{A_1 \times A_2}$  apart from the end vertices  $a_1$  and  $v_{i_{r'}}$ . From (2) and (3) and the degree conditions we get that the resulting graph still satisfies the conditions of Lemma 6 and thus it is Hamiltonian-connected. This closes the desired Hamiltonian cycle and finishes Case 1. For this purpose we could also use Lemma 2 because the remaining bipartite graph is super-regular with the

appropriate choice of parameters, but here the much simpler Lemma 6 also suffices. Note also that here we have no exceptional  $i$ , so we have

$$|\text{dist}_G(a_i, a_{i+1}) - d_i| \leq 1 \quad \text{for all } 1 \leq i \leq k - 1,$$

and here this is true again for any ordering of the vertices in  $S$ .

**Subcase 1.2.** We have  $|A_1| = |A_2|$ . The proof is similar to the proof of Subcase 1.1. Corresponding to Fact 4 here we have the following.

**Fact 5.** For any distinct non-exceptional  $x, y \in A_i, i \in \{1, 2\}$  there are at least  $n/4$  internally disjoint paths of length 2 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ . For any distinct (possibly exceptional)  $x, y \in A_i, i \in \{1, 2\}$  there are at least  $\frac{\alpha^{1/7}}{8}n$  internally disjoint paths of length 4 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ . For any  $x$  and  $y$  that are in different sets, say  $x \in A_1$  and  $y \in A_2$ , there are at least  $\frac{\alpha^{1/7}}{8}n$  internally disjoint paths of length 3 in  $G|_{A_1 \times A_2}$  connecting  $x$  and  $y$ .

**Proof.** Indeed for the first statement note that since  $x, y \in A_i$  are non-exceptional we have

$$\deg(x, A_{i'}) \geq (1 - 4\alpha^{1/6})|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Then the number of common neighbors of  $x$  and  $y$  in  $A_{i'}$  is at least  $(1 - 8\alpha^{1/6})|A_{i'}| \geq n/4$  and from this the first statement follows. For the second statement consider two disjoint equal-size subsets of non-exceptional vertices from the two neighborhoods  $X \subset N(x, A_{i'})$  and  $Y \subset N(y, A_{i'})$  with  $|X| = |Y| \geq \frac{\alpha^{1/7}}{8}n$ . Pairing up the vertices between  $X$  and  $Y$  arbitrarily and applying the first statement for each pair we get the second statement. Finally for the last statement consider a subset of non-exceptional vertices of the neighborhood  $X \subset N(x, A_2)$  with  $|X| \geq \frac{\alpha^{1/7}}{8}n$ . Applying the first statement for each of the pairs  $(y, u)$  where  $u \in X$  we get the last statement.  $\square$

The remaining portion of Subcase 1.2 is similar to Subcase 1.1. By using Fact 5 repeatedly we find in  $G|_{A_1 \times A_2}$  an  $(S, \underline{d})$ -path connecting the vertices  $a_1$  and  $a_k$ . Here the situation is even simpler as we do not have to worry about the matching edges. We remove this path and apply Lemma 6 in the leftover.

**Case 2.** Assume next that we have a partition  $V(G) = A_1 \cup A_2$  with  $(\frac{1}{2} - \alpha)n \leq |A_1| \leq \frac{n}{2}$  and  $d(A_1, A_2) < \alpha^{1/3}$ . Thus roughly speaking,  $G|_{A_1}$  and  $G|_{A_2}$  are almost complete and the bipartite graph between  $A_1$  and  $A_2$  is sparse.

Again we define exceptional vertices  $v \in A_i, i \in \{1, 2\}$ , as

$$\deg(v, A_{i'}) \geq \alpha^{1/6}|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Note that from the density condition  $d(A_1, A_2) < \alpha^{1/3}$ , the number of exceptional vertices in  $A_i$  is at most  $\alpha^{1/6}|A_i|$ . We remove the exceptional vertices from each set and we redistribute them in such a way that  $e(A_1, A_2)$  is minimized. We still denote the sets by  $A_1$  and  $A_2$ . It is easy to see that in  $G|_{A_i}, i \in \{1, 2\}$ , apart from at most  $3\alpha^{1/6}|A_i|$  exceptional vertices all the degrees are at least  $(1 - 3\alpha^{1/6})|A_i|$ , and the degrees of the exceptional vertices are at least  $|A_i|/3$ . These degree conditions and (2) imply the following fact (similar to Facts 1, 4 and 5).

**Fact 6.** For any distinct  $x, y \in A_i, i \in \{1, 2\}$ , where at most one of the vertices is exceptional, there are at least  $n/8$  internally disjoint paths of length 2 in  $G|_{A_i}$  connecting  $x$  and  $y$ . For any distinct  $x, y \in A_i, i \in \{1, 2\}$  there are at least  $n/8$  internally disjoint paths of length 3 in  $G|_{A_i}$  connecting  $x$  and  $y$ .

Assume that  $|A_1| \leq |A_2|$ . Let  $S$  be an arbitrary subset of the vertices of  $G$  satisfying (3). Put

$$\begin{aligned} S' &= S \cap A_1, \quad S'' = S \cap A_2, \quad k' = |S'|, \quad k'' = |S''|, \\ \underline{d}' &= \{d_i | 1 \leq i \leq k' - 1\} \quad \text{and} \quad \underline{d}'' = \{d_i | k' + 1 \leq i \leq k - 1\}. \end{aligned}$$

We show that we can find two vertex disjoint edges (called bridges)  $\{u_1, v_1\}, \{u_2, v_2\}$  in  $G|_{A_1 \times A_2}$  such that for both of these bridges at least one of the end points (say  $u_i$ ) is non-exceptional and it is not in  $S$ . This is trivial if  $|A_1| < |A_2|$ , since then for every  $u \in A_1$  we have  $\deg(u, A_2) \geq 2$ . Thus we may assume that  $|A_1| = |A_2|$ . But then for every  $u \in A_1$  we have  $\deg(u, A_2) \geq 1$  and for every  $v \in A_2$  we have  $\deg(v, A_1) \geq 1$ , and thus again we can pick the two bridges.

We distinguish two subcases.

**Subcase 2.1.**  $u_1$  and  $u_2$  are in different sets, say  $u_1 \in A_1 \setminus S'$  and  $u_2 \in A_2 \setminus S''$ . Here we construct the desired Hamiltonian cycle in the following way. First by using Fact 6 and a similar procedure as in Section 4.1 we find in  $G|_{A_1}$  an  $(S', \underline{d}')$ -path  $P' = P'(a_1, v_2)$  with end points  $a_1 \in S$  and  $v_2$  (if  $v_2 \in S'$  then this is just the last vertex  $v_2 = a_{k'}$  from  $S$  on the path, otherwise we connect the last vertex  $a_{k'}$  and  $v_2$  by a path of length 3). Similarly we find in  $G|_{A_2}$  an  $(S'', \underline{d}'')$ -path  $P'' = P''(a_{k'+1}, v_1)$  with end points  $a_{k'+1} \in S$  and  $v_1$ . Then in  $G|_{A_1}$  we remove the path  $P'$  apart from the end vertex  $a_1$ . From (2) and (3) and the degree conditions we get that the resulting graph satisfies the conditions of Lemma 5 and thus it is Hamiltonian-connected. Take a Hamiltonian path  $P_1 = P_1(u_1, a_1)$  with end points  $u_1$  and  $a_1$ . Similarly in  $G|_{A_2}$  we remove the path  $P''$  apart from the end vertex  $a_{k'+1}$  and we find a Hamiltonian path  $P_2 = P_2(u_2, a_{k'+1})$  with end points  $u_2$  and  $a_{k'+1}$ . Then in this case the desired Hamiltonian cycle  $C$  is the following.

$$C = (P', \{v_2, u_2\}, P_2, P'', \{v_1, u_1\}, P_1).$$

Note that here actually in  $C$  we have

$$\text{dist}_C(a_i, a_{i+1}) = d_i \quad \text{for all } 1 \leq i \leq k' - 1 \text{ and } k' + 1 \leq i \leq k - 1.$$

However,  $\text{dist}_C(a_{k'}, a_{k'+1})$  could be very different from  $d_{k'}$ .

**Subcase 2.2.**  $u_1$  and  $u_2$  are in the same set (say  $A_1$ ). Here we do the following. We may assume that  $v_1, v_2 \in S''$ , since otherwise we are back at Subcase 2.1. We denote  $v_2$  by  $a_{k'+1}$  and  $v_1$  by  $a_k$ . First we find in  $G|_{A_1}$  again an  $(S', \underline{d}')$ -path  $P' = P'(a_1, a_{k'})$  with end points  $a_1$  and  $a_{k'}$ . We connect  $a_{k'}$  and  $u_2$  with a path  $Q = Q(a_{k'}, u_2)$  of length  $d_{k'} - 1$  that is internally disjoint from  $P'$  and  $u_1$ . The degree conditions guarantee that this is possible (even if  $d_{k'} = 3$ , since  $u_2$  is non-exceptional). Then we remove  $P'$  and  $Q$  from  $G|_{A_1}$  apart from the end vertex  $a_1$  and we find a Hamiltonian path  $P_1 = P_1(u_1, a_1)$  with end points  $u_1$  and  $a_1$ . Define

$$S''' = S'' \setminus \{a_k\} \quad \text{and} \quad \underline{d}''' = \{d_i | k' + 1 \leq i \leq k - 2\} = \underline{d}'' \setminus \{d_{k-1}\}.$$

We find in  $G|_{A_2}$  an  $(S''', \underline{d}''')$ -path  $P'' = P''(a_{k'+1}, a_{k-1})$  with end points  $a_{k'+1}$  and  $a_{k-1}$ . We remove  $P''$  from  $G|_{A_2}$  apart from the end vertex  $a_{k-1}$  and we find a Hamiltonian path  $P_2 = P_2(a_{k-1}, v_1)$  with end points  $a_{k-1}$  and  $v_1 = a_k$ . Then in this case the Hamiltonian cycle  $C$  is the following.

$$C = (P', Q, \{u_2, v_2\}, P'', P_2, \{v_1, u_1\}, P_1).$$

Note that here actually in  $C$  we have

$$\text{dist}_C(a_i, a_{i+1}) = d_i \quad \text{for all } 1 \leq i \leq k - 2,$$

but  $\text{dist}_C(a_{k-1}, a_k)$  could be very different from  $d_{k-1}$ . This finishes Case 2.

Assume finally that the extremal condition holds with parameter  $\alpha$ , so we have  $A, B \subset V(G)$ ,  $|A|, |B| \geq (\frac{1}{2} - \alpha)n$  and  $d(A, B) < \alpha$ . We may also assume  $|A|, |B| \leq n/2$ . We have three possibilities.

- $|A \cap B| < \sqrt{\alpha n}$ . The statement follows from Case 2. Indeed, let  $A_1 = A$ ,  $A_2 = V(G) \setminus A_1$ , then clearly  $d(A_1, A_2) < \alpha^{1/3}$  if  $\alpha \ll 1$  holds.
- $\sqrt{\alpha n} \leq |A \cap B| < (1 - \sqrt{\alpha})\frac{n}{2}$ . This case is not possible under the given conditions. In fact, otherwise we would get

$$\begin{aligned} |A \cap B| \frac{n}{2} &\leq \sum_{u \in A \cap B} \deg_G(u) = \sum_{u \in A \cap B} \deg_G(u, A \cup B) + \sum_{u \in A \cap B} \deg_G(u, V(G) \setminus (A \cup B)) \\ &\leq 2\alpha n^2 + |A \cap B| (|A \cap B| + 1), \end{aligned}$$

or

$$|A \cap B| \left( \frac{n}{2} - |A \cap B| - 1 \right) \leq 2\alpha n^2. \tag{11}$$

Here in the given range for  $|A \cap B|$  the left side is always greater than

$$(1 - \sqrt{\alpha}) \frac{n}{2} \left( \sqrt{\alpha} \frac{n}{2} - 1 \right) \geq \sqrt{\alpha} \frac{n^2}{8} \gg 2\alpha n^2,$$

(using  $\alpha \ll 1$ ) a contradiction to (11).

- $|A \cap B| \geq (1 - \sqrt{\alpha}) \frac{n}{2}$ . The statement follows from [Case 1](#) by choosing  $A_1 = A$ ,  $A_2 = V(G) \setminus A_1$ , and then  $d(A_1) < \alpha^{1/3}$ .

This finishes the extremal case and the proof of [Theorem 2](#).

## References

- [1] C. Berge, *Graphs and Hypergraphs*, North-Holland, New York, 1973.
- [2] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.
- [3] I. Csiszár, J. Körner, Graph decomposition: a new key to coding theorems, *IEEE Trans. Inform. Theory* 27 (1981) 5–12.
- [4] G.A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* 2 (1952) 68–81.
- [5] M.H. El-Zahar, On circuits in graphs, *Discrete Mathematics* 50 (1984) 227–230.
- [6] H. Enomoto, S. Matsunaga, Graph decompositions without isolated vertices III, *Journal of Graph Theory* 24 (1997) 155–164.
- [7] P. Erdős, L. Pósa, On the maximal number of disjoint circuits of a graph, *Publ. Math. Debrecen* 9 (1962) 3–12.
- [8] R.J. Gould, Advances on the Hamiltonian problem — a survey, *Graphs and Combinatorics* 19 (2003) 7–52.
- [9] E. Győri, On division of graphs to connected subgraphs, in: *Combinatorics (Proc. 5th Hungarian Colloq., Keszthely, 1976) I*, North-Holland, 1978, pp. 485–494.
- [10] A. Kaneko, Yoshimoto, On a Hamiltonian cycle in which specified vertices are uniformly distributed, *Journal of Combinatorial Theory, Ser. B* 81 (2001) 100–109.
- [11] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of a packing conjecture of Bollobás, *Combinatorics, Probability and Computing* 4 (1995) 241–255.
- [12] J. Komlós, G.N. Sárközy, E. Szemerédi, On the square of a Hamiltonian cycle in dense graphs, *Random Structures and Algorithms* 9 (1996) 193–211.
- [13] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, *Combinatorica* 17 (1997) 109–123.
- [14] J. Komlós, G.N. Sárközy, E. Szemerédi, An algorithmic version of the Blow-up Lemma, *Random Structures and Algorithms* 12 (1998) 297–312.
- [15] J. Komlós, G.N. Sárközy, E. Szemerédi, On the Pósa-Seymour conjecture, *Journal of Graph Theory* 29 (1998) 167–176.
- [16] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Annals of Combinatorics* 2 (1998) 43–60.
- [17] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Mathematics* 235 (2001) 255–269.
- [18] J. Komlós, M. Simonovits, Szemerédi's Regularity Lemma and its applications in Graph Theory, in: D. Miklós, V.T. Sós, T. Szönyi (Eds.), *Combinatorics, Paul Erdős is Eighty*, in: *Bolyai Society Mathematical Studies*, vol. 2, Budapest, 1996, pp. 295–352.
- [19] L. Lovász, Subgraphs with prescribed valencies, *Journal of Combinatorial Theory* 8 (1970) 391–416.
- [20] V. Rödl, A. Ruciński, Perfect matchings in  $\epsilon$ -regular graphs and the Blow-up Lemma, *Combinatorica* 19 (1999) 437–452.
- [21] V. Rödl, A. Ruciński, A. Taraz, Hypergraph packing and graph embedding, *Combinatorics, Probability and Computing* 8 (1999) 363–376.
- [22] G.N. Sárközy, S. Selkow, E. Szemerédi, On the number of Hamiltonian cycles in Dirac graphs, *Discrete Mathematics* 265 (2003) 237–250.
- [23] E. Szemerédi, Regular partitions of graphs, in: *Problèmes Combinatoires et Théorie des Graphes (Colloques Internationaux C.N.R.S. No. 260)* Orsay, 1976, pp. 399–401.