Distributing vertices along a Hamiltonian cycle in Dirac graphs

Gábor N. Sárközy
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA 01609
gsarkozy@cs.wpi.edu
and
Computer and Automation Research Institute
Hungarian Academy of Sciences
Budapest, P.O. Box 63
Budapest, Hungary, H-1518

Stanley Selkow
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA 01609
sms@cs.wpi.edu

February 19, 2004

Abstract

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

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

1 Introduction

1.1 Notation and definitions

For basic graph concepts see the monograph of Bollobás [2].
$+$ will sometimes be used for disjoint union of sets. $V(G)$ 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 $V = A + B$, 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 neighbours 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, and $\Delta(G)$ for the maximum degree in G. $\nu(G)$ is the size of a maximum matching in G. The distance $dist_G(u, v)$ is defined as the number of edges in a shortest subpath of G joining u and v. For $A \subset V(G)$ we write $N(A) = \cap_{v \in A} N(v)$, the set of common neighbours. $N(x, y, z, ...)$ is shorthand for $N(\{x, y, z, ...\})$. For a vertex $v \in V$ and set $U \subset V - \{v\}$, we write $deg(v, U)$ for the number of edges from v to U. When A, B are disjoint subsets of $V(G)$, we denote by $e(A, B)$ the number of edges of G with one endpoint in A and the other in B. 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. In particular, we write $d(A) = d(A, A) = 2|E(G|_A)|/|A|^2$.

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

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

otherwise it is ε-irregular.

We will often say simply that "the pair (A, B) is ε-regular" with the graph G implicit. We will also need a stronger version.

Definition 2. *(A, B) is (ε, δ)* **super-regular** *if it is ε-regular and*

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

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 **Hamiltonian-connected** graph is a graph in which every pair of vertices can be connected with a Hamiltonian path. A classical result of Dirac [3] asserts that if $\delta(G) \geq n/2$, 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 [4]).

In a recent, interesting strengthening of Dirac's Theorem, Kaneko and Yoshimoto [5] showed that in a Dirac graph small subsets of vertices can be somewhat uniformly distributed along a Hamiltonian cycle.

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 A with at most $n/2d$ vertices, there exists a Hamiltonian cycle C with $dist_C(u, v) \geq d$ for every u and v in A.*

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

In [4] Gould called for further studies on density conditions that allow the distribution of "small" subsets of vertices along a Hamiltonian cycle. In this paper we show that with similar conditions we can not only achieve that the distance between two subsequent vertices of A along C is at least d, but actually we can prescribe the exact distances (apart from a difference of at most 1) between all but two pairs of subsequent vertices of A along C. More precisely we show the following.

Theorem 2. *There are $\kappa, n_0 > 0$ 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 $3 \leq d \leq \kappa n/2$ and A is an arbitrary subset of the vertices of G with $2 \leq |A| = k \leq \kappa n/d$, 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 A, $a_1, a_2, \ldots, a_k$, such that the vertices of A are visited in this order on C and we have*

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

We need the discrepancies by 1 between $dist_C(a_i, a_{i+1})$ and d_i because of parity reasons. Indeed, consider the complete bipartite graph between U and V, where $|U| = |V| = n/2$. Take $A \subset U$, then the distance between subsequent vertices of A 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 $|dist_C(a_i, a_{i+1}) - d_i| > 1$, consider the following construction. Take two complete graphs on U and V with $|U| = |V| = n/2$. Let $A = A' \cup A''$ with $A' \subset U$, $A'' \subset V$ and $|A'| = |A''| = |A|/2$, and add the complete bipartite graphs between A' and V, and between A'' and U. Clearly on any Hamiltonian cycle we will have two distances much greater than d.

We believe that our theorem remains true for greater $|A|$'s as well, but we were unable to prove a stronger statement.

2 The main tools

In the proof the following lemma of Szemerédi plays a central role.

Lemma 1 (Regularity Lemma [15]). *For every positive ε and positive integer m there are positive integers M and n_1 with the following property: for every graph G with $n \geq n_1$ vertices there is a partition of the vertex set into $l+1$ classes (clusters)*

$$V = V_0 + V_1 + V_2 + ... + V_l$$

such that

- $m \leq l \leq M$
- $|V_1| = |V_2| = ... = |V_l|$
- $|V_0| < \varepsilon n$
- *at most εl^2 of the pairs $\{V_i, V_j\}$ are ε-irregular.*

We will use the following simple consequence of Lemma 1.

Lemma 2 (Degree form). *For every $\varepsilon > 0$ there is an $M = M(\varepsilon)$ such that if $G = (V, E)$ is any graph and $\delta \in [0, 1]$ is any real number, then there is a partition of the vertex-set V into $l+1$ clusters $V_0, V_1, ..., V_l$, and there is a subgraph $G' = (V, E')$ with the following properties:*

- $l \leq M$,
- $|V_0| \leq \varepsilon |V|$,
- *all clusters V_i, $i \geq 1$, are of the same size $L \leq \lceil \varepsilon |V| \rceil$.*
- $deg_{G'}(v) > deg_G(v) - (\delta + \varepsilon)|V|$ *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 ε-regular, each with a density either 0 or exceeding δ.*

The other main tool 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. This is a very special case of the Blow-up Lemma ([8], [9]). More precisely.

Lemma 3. *For every $\delta > 0$ there are $\varepsilon_0, n_2 > 0$ such that if $\varepsilon \leq \varepsilon_0$ and $n \geq n_2$, $G = (A, B)$ is an (ε, δ) 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 two simple Pósa-type lemmas on Hamiltonian-connectedness. The second one is the bipartite version of the first one.

Lemma 4 (see [1]). *Let G be a graph on $n \geq 3$ vertices with degrees $d_1 \leq d_2 \leq \ldots \leq d_n$ such that for every $2 \leq k \leq \frac{n}{2}$ we have $d_{k-1} > k$. Then G is Hamiltonian-connected.*

Lemma 5 (see [1]). *Let $G = (A, B)$ be a bipartite graph with $|A| = |B| = n \geq 2$ with degrees $d_1 \leq d_2 \leq \ldots \leq d_n$ from A and with degrees $d'_1 \leq d'_2 \leq \ldots \leq d'_n$ from B. Suppose that for every $2 \leq j \leq \frac{n+1}{2}$ we have $d_{j-1} > j$ and that for every $2 \leq k \leq \frac{n+1}{2}$ we have $d'_{k-1} > k$. Then G is Hamiltonian-connected.*

Finally we will use the following simple fact.

Lemma 6 (Erdős, Pósa, see [2]). *Let G be a graph on n vertices. Then*

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

In case we have a good upper bound on the maximum degree of G, we can strengthen this lemma in the following way.

Lemma 7. *In a graph G of order n*

$$\nu(G) \geq \delta(G) \frac{n}{2(\delta(G) + \Delta(G))} \geq \delta(G) \frac{n}{4\Delta(G)}.$$

In fact, let us take a maximal matching M with m edges. Then for the number of edges E between M and $V(G) \setminus M$ we get $\delta(G)(n - 2m) \leq E \leq 2m\Delta(G)$, which proves the lemma.

3 Outline of the proof

In this paper we use the Regularity Lemma-Blow-up Lemma method again (see [6]-[12], [14]). The method is usually applied to find certain spanning subgraphs in dense graphs. Typical examples are spanning trees (Bollobás-conjecture, see [6]), Hamiltonian cycles or powers of Hamiltonian cycles (Pósa-Seymour conjecture, see [10, 11]) or H-factors for a fixed graph H (Alon-Yuster conjecture, see [12]).

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

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

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 $4 \leq d \leq \kappa n/2$ and let A be an arbitrary subset of the vertices of G with

$$2 \leq |A| = k \leq \kappa n/d, \tag{3}$$

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 $(A, \underline{d})$-cycle (or an $(A, \underline{d})$-path) if there is an ordering of the vertices of A, $a_1, a_2, \ldots, a_k$, such that the vertices of A are visited in this order on C (on P) and we have

$$|dist_C(a_i, a_{i+1}) - d_i| \leq 1, \ 1 \leq i \leq k-1.$$

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

$$|dist_C(a_i, a_{i+1}) - d_i| > 1$$

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): *There exist (not necessarily disjoint) $A, B \subset V(G)$ such that*

- $|A| = |B| = \lfloor \frac{n}{2} \rfloor$, *and*
- $d(A, B) < \alpha$.

In the non-extremal case we apply Lemma 2 for G, with ε and δ 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 the clusters V_i, $1 \leq i \leq l$, and we have an edge between two clusters if they form an ε-regular pair in G' with density exceeding δ. Since in G', $\delta(G') > (\frac{1}{2} - (\delta + \varepsilon))n$, an easy calculation shows that in G_r we have

$$\delta(G_r) \geq \left(\frac{1}{2} - 3\delta\right) l. \tag{4}$$

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 V_i in G_r, then for a $V_i, 1 \le i \le l$ we have:

$$\left(\frac{1}{2} - (\delta + \varepsilon)\right) nL \le \sum_{u \in V_i} deg_{G'}(u) \le \varepsilon nL + deg_{G_r}(V_i) L^2.$$

From this we get inequality (4):

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

Applying Lemma 6 we can find a matching M in G_r of size at least $\left(\frac{1}{2} - 3\delta\right) l$. 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

$$|V_0| \le 6\delta l L + \varepsilon n \le 7\delta n. \tag{5}$$

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

The rest of the non-extremal case is organized as follows. In Section 4.1 first we find an $(A, \underline{d})$-path P. Then in Section 4.2 we find short connecting paths P_i between the consecutive edges in the matching 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, the others have 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 M is perfect, i.e. there are the same number of vertices left in both clusters of the edge. Finally applying Lemma 3 we close the Hamiltonian cycle in each edge and thus giving a Hamiltonian $(A, \underline{d})$-cycle.

4 The non-extremal case

Throughout this section we assume that the extremal case EC does not hold.

4.1 Finding an $(A, \underline{d})$-path

We are going to use the following fact several times.

Fact 1. *If $x, y \in V(G)$ then there are at least δn internally disjoint paths of length 3 connecting x and y.*

Indeed, if we choose $A \subset N_G(x)$ with $|A| = \lfloor \frac{n}{2} \rfloor$ and $B \subset N_G(y)$ with $|B| = \lfloor \frac{n}{2} \rfloor$, then the fact that EC does not hold implies $d(A, B) \geq \alpha$ and Fact 1 follows.

We construct an $(A, \underline{d})$-path $P = Q_1 \cup \ldots \cup Q_k$ in the following way. Let $a_1, \ldots, a_k$ be the vertices of A in an arbitrary order (so note that here actually we can prescribe the order of the vertices of A as well). First we construct a path Q_1 of length d_1 connecting a_1 and a_2. For this purpose first we construct greedily a path Q'_1 starting from a_1 that has length $d_1 - 3$ ((1) makes this possible). Denote the other endpoint 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 endpoint 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 $(A, \underline{d})$-path P. (1), (2), (3) and Fact 1 imply that we never get stuck since

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

Observe that here in the non-extremal case there is no discrepancy between $dist(a_i, a_{i+1})$ and d_i for all $1 \leq i \leq k-1$, and furthermore we can specify the order of the vertices of A as well.

4.2 Connecting paths

For the first connecting path P_1 between (V_1^1, V_2^1) and (V_1^2, V_2^2), first we connect a typical vertex u of V_2^1 (more precisely a vertex u with $deg(u, V_1^1) \geq (\delta - \varepsilon)|V_1^1|$, most vertices in V_2^1 satisfy this) and a_1 with a path of length 3, and then we connect a_k and a typical vertex w of V_1^2 (so $deg(w, V_2^2) \geq (\delta - \varepsilon)|V_2^2|$) 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. 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. Thus P_1 has length at most $\kappa n + 6$, all other P_i-s have length 3.

We remove the vertices on these connecting paths from the clusters, but for simplicity we keep the notation for the resulting clusters. These connecting paths will be parts 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. So we may always assume that the number of remaining vertices is even.

4.3 Adjustments and the handling of the exceptional vertices

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 $deg(u, V_2^i) < (\delta - \varepsilon)|V_2^i|$. ε-regularity guarantees that at most $\varepsilon|V_1^i| \leq \varepsilon L$ such vertices exist in each cluster V_1^i.

Thus using (5), we still have

$$|V_0| \leq 7\delta n + 2\varepsilon n \leq 9\delta n.$$

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) w in V_0. We find a pair (V_1^i, V_2^i) such that either

$$deg(w, V_1^i) \geq \delta|V_1^i|, \tag{6}$$

or

$$deg(w, V_2^i) \geq \delta|V_2^i|. \tag{7}$$

We assign w to the pair (V_1^i, V_2^i). We extend P_{i-1} (for $i = 1$, P_m) in (V_1^i, V_2^i) by a path of length 3 in case (6) holds, and by a path of length 4 in case (7) holds, so that now the path ends with w. To finish the procedure for w, in case (6) holds we add one more vertex w' to P_{i-1} after w such that $(w, w') \in E(G)$ and w' is a typical vertex of V_1^i, so $deg(w', V_2^i) \geq (\delta - \varepsilon)|V_2^i|$. In case (7) holds we add two more vertices w', w'' to P_{i-1} after w such that $(w, w'), (w', w'') \in E(G)$, w' is a typical vertex of V_2^i and w'' is a typical vertex of V_1^i.

After handling w, 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 free vertices in each cluster; the vertices on the connecting paths are always removed. Second, we make sure that we never assign too many vertices of V_0 to one pair (V_1^i, V_2^i). It is not hard to see (using (1) and $\delta \ll 1$) that we can guarantee that we always assign at most $\sqrt{\delta}|V_1^i|$ vertices of V_0 to a pair (V_1^i, V_2^i). Finally, since we are removing vertices from a pair (V_1^i, V_2^i), we might violate the super-regularity. Note that we never violate the ε-regularity. Therefore, we do the following. After handling (say) $\lfloor \delta^2 n \rfloor$ vertices from V_0, we update V_0 as follows. In a pair (V_1^i, V_2^i) we remove all vertices u from V_1^i (and similarly from V_2^i) for which $deg(u, V_2^i) < (\delta - \varepsilon)|V_2^i|$ (again, we consider only the remaining vertices). Again, we added at most $2\varepsilon n$ vertices to V_0. In V_0 we handle these vertices first and then we move on to the other vertices in V_0.

After we are done with all the vertices in V_0, we might have a small discrepancy ($\leq 2\sqrt{\delta}|V_1^i|$) among the remaining vertices in V_1^i and in V_2^i in a pair. Therefore, we have to make some adjustments. Let us take a pair (V_1^i, V_2^i) with a discrepancy ≥ 2 (if one such pair exists), say $|V_1^i| \geq |V_2^i| + 2$ (only remaining vertices are considered). Using the fact that EC does not hold we find an **alternating path** (with respect to M) in G_r of length 6 starting with V_1^i and ending with V_2^i. Let us denote this path by

$$V_1^i, V_2^{i_1}, V_1^{i_1}, V_1^{i_2}, V_2^{i_2}, V_1^i, V_2^i$$

(the construction is similar if the clusters in $(V_1^{i_1}, V_2^{i_1})$ or in $(V_1^{i_2}, V_2^{i_2})$ are visited in different order). We remove a typical vertex from V_1^i and we add it to $V_1^{i_1}$, then we remove a typical vertex from $V_1^{i_1}$ and we add it to $V_2^{i_2}$, finally we remove a typical vertex from $V_2^{i_2}$ and we add it to V_2^i. When we add a new vertex to a pair (V_1^j, V_2^j), we extend the connecting path P_{j-1} by a path of length 4 in the pair so that it now includes the new vertex.

Now we are one step closer to the perfect distribution, and by iterating this procedure we can assure that the discrepancy in every pair is ≤ 1. We consider only those pairs for which the discrepancy is exactly 1, so in particular the number of remaining vertices in one such a pair is odd. From the construction 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 6 between V_1^i and V_2^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 left. In this case Lemma 3 closes the Hamiltonian cycle in every pair.

5 The extremal case

First we assume that we have the following special case.

Case 1: There is a partition $V(G) = A_1 \cup A_2$ with $|A_1| = \lfloor \frac{n}{2} \rfloor$ and $d(A_1) < \alpha^{1/3}$.

Note that in this case from (1) we also have $d(A_1, A_2) > 1 - \alpha^{1/3}$. 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.

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

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

Note that (1) implies that if $v \in A_i$ is exceptional, then

$$deg(v, A_i) \geq \alpha^{1/6} |A_i|.$$

But then since $d(A_1, A_2) > 1 - \alpha^{1/3}$, we get that the number of exceptional vertices in A_i is at most $\alpha^{\frac{1}{6}}|A_i|$. We remove the exceptional vertices from each set and add them to A_2 if they have more neighbors in A_1, and add them to A_1 if they have more neighbors in A_2. 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 $0 \leq r \leq 2\alpha^{1/6}|A_2|$. In $G|_{A_1 \times A_2}$ apart from at most $2\alpha^{1/6}|A_2|$ exceptional vertices all the degrees are at least $(1 - 3\alpha^{1/6})|A_2|$, and the degrees of the exceptional vertices are at least $|A_2|/3$.

Our goal is to achieve $r = 0$. If there is a vertex $x \in A_2$ for which

$$deg(x, A_2) \geq \alpha^{1/7} |A_2|, \tag{8}$$

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 (8) or $|A_1| = |A_2|$. Assume that we have the first case. Since we have $\Delta(G|_{A_2}) < \alpha^{1/7}|A_2|$, (1) and Lemma 7 imply that $G|_{A_2}$ has an r-matching M denoted by $\{u_1, v_1\}, \ldots, \{u_r, v_r\}$. Furthermore, for every edge in M we can guarantee that at least one of the endpoints (say u_i) is not in A. This matching M 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 are still much more than the number of these exceptional vertices. These degree conditions and (2) imply the following fact (similar to Fact 1).

Fact 2. *If $x, y \in A_1$ then in $G|_{A_i \times A_2}$ there are at least δn internally disjoint paths of length 4 connecting x and y. If $x, y \in A_2$ then in $G|_{A_i \times A_2}$ there are at least δn internally disjoint paths of length 2 connecting x and y. If $x \in A_i$, $y \in A_{i'}$ then in $G|_{A_i \times A_2}$ there are at least δn internally disjoint paths of length 3 connecting x and y.*

Let A be an arbitrary subset of the vertices of G satisfying (3). In this case we construct the desired Hamiltonian cycle in the following way. First by using Fact 2 and a similar procedure as in Section 4.1 we find in $G|_{A_1 \times A_2}$ an $(A, \underline{d})$-path

$$P = P(a_1, a_k) = Q_1 \cup \ldots \cup Q_k$$

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 $dist_C(a_i, a_{i+1}) = d_i + 1$. 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, say $a_1 = v_i$, then we start Q_1 with the edge $\{v_i, u_i\}$ (note that $u_i \notin A$). 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 endpoint of Q_1' by a_1'. Applying Fact 2, 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 $Q_2, \ldots, Q_k$ similarly and thus we get $P = Q_1 \cup \ldots \cup Q_k$. Say the remaining edges of M which are not traversed by P are

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

Then we connect the endpoint a_k of P and u_{i_1} by a path Q_1 of length 2 or 3, connect v_{i_1} and u_{i_2} by a path Q_2 of length 2, etc. Finally connect $v_{i_{r'-1}}$ and $u_{i_{r'}}$ by a path $Q_{r'}$ of length 2. Consider the following path.

$$P' = (P, Q_1, \{u_{i_1}, v_{i_1}\}, Q_2, \{u_{i_2}, v_{i_2}\}, \ldots, Q_{r'}, \{u_{i_{r'}}, v_{i_{r'}}\}).$$

In case $a_1 \in A_2$, add one more vertex from A_1 to the end of the path. Remove P' from $G|_{A_1 \times A_2}$ apart from the endvertices a_1 and $v_{i_{r'}}$. From (2), (3) and the degree conditions we get that the resulting graph satisfies the conditions of Lemma 5 and thus it is Hamiltonian-connected. This closes the desired Hamiltonian cycle. For this purpose we could also use Lemma 3, the remaining bipartite graph is super-regular with the appropriate choice of parameters, but here the much simpler Lemma 5 also suffices. Note also that here we have no exceptional i, so we have

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

Case 2: Assume next that we have a partition $V(G) = A_1 \cup A_2$ with $|A_1| = \lfloor \frac{n}{2} \rfloor$ 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

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

Note that again 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 add them to the set where they have more neighbors. We still denote the sets by A_1 and A_2. Thus in $G|_{A_i}, i \in \{1, 2\}$, apart from at most $2\alpha^{1/6}|A_i|$ exceptional vertices all the degrees are at least $(1 - 2\alpha^{1/6})|A_i|$, 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 and 2).

Fact 3. *If $x, y \in A_i$ then in $G|_{A_i}$ there are at least δn internally disjoint paths of length 3 connecting x and y. Furthermore, if at least one of the vertices x and y is non-exceptional then there are at least δn internally disjoint paths of length 2 connecting x and y.*

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

$$A' = A \cap A_1, A'' = A \cap A_2, k' = |A'|, k'' = |A''|,$$

$$\underline{d}' = \{d_i \mid 1 \leq i \leq k' - 1\} \text{ and } \underline{d}'' = \{d_i \mid k' + 1 \leq i \leq k - 1\}.$$

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 endpoints (say u_i) is non-exceptional and it is not in A. 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 $u_1 \in A_1 \setminus A'$ and $u_2 \in A_2 \setminus A''$. Here we construct the desired Hamiltonian cycle in the following way. First by using Fact 3 and a similar procedure as in Section 4.1 we find in $G|_{A_1}$ an $(A', \underline{d}')$-path $P' = P'(a_1, v_2)$ with endpoints $a_1 \in A$ and v_2 (if $v_2 \in A'$ then this is just the last vertex $v_2 = a_{k'}$ from A 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 $(A'', \underline{d}'')$-path $P'' = P''(a_{k'+1}, v_1)$ with endpoints $a_{k'+1} \in A$ and v_1. Then in $G|_{A_1}$ we remove the path P' apart from the endvertex a_1. From (2), (3) and the degree conditions we get that the resulting graph satisfies the conditions of Lemma 4 and thus it is Hamiltonian-connected. Take a Hamiltonian path $P_1 = P_1(u_1, a_1)$ with endpoints u_1 and a_1. Similarly in $G|_{A_2}$ we remove the path P'' apart from the endvertex $a_{k'+1}$ and we find a Hamiltonian path $P_2 = P_2(u_2, a_{k'+1})$ with endpoints 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

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

However, $dist_C(a_{k'}, a_{k'+1})$ 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 $v_1, v_2 \in A''$, since otherwise we are back to 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 $(A', \underline{d}')$-path $P' = P'(a_1, a_{k'})$ with endpoints 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 endvertex a_1 and we find a Hamiltonian path $P_1 = P_1(u_1, a_1)$ with endpoints u_1 and a_1. Define

$$A''' = A'' \setminus \{a_k\} \text{ and } \underline{d}''' = \{d_i \mid k'+1 \le i \le k-2\} = \underline{d}'' \setminus \{d_{k-1}\}.$$

We find in $G|_{A_2}$ an $(A''', \underline{d}''')$-path $P'' = P''(a_{k'+1}, a_{k-1})$ with endpoints $a_{k'+1}$ and a_{k-1}. We remove P'' from $G|_{A_2}$ apart from the endvertex a_{k-1} and we find a Hamiltonian path $P_2 = P_2(a_{k-1}, v_1)$ with endpoints 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

$$dist_C(a_i, a_{i+1}) = d_i \text{ for all } 1 \le i \le k-2,$$

but $dist_C(a_{k-1}, a_k)$ could be very different from d_{k-1}.

Case 3: Assume finally that the extremal case EC holds, so we have $A, B \subset V(G)$, $|A| = |B| = \lfloor \frac{n}{2} \rfloor$ and $d(A, B) < \alpha$. 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.

- $\sqrt{\alpha} n \le |A \cap B| < (1 - \sqrt{\alpha}) \frac{n}{2}$. This case is not possible under the given conditions. In fact, otherwise we would get

$$|A \cap B| \frac{n}{2} \le \sum_{u \in A \cap B} deg_G(u) = \sum_{u \in A \cap B} deg_G(u, A \cup B)$$
$$+ \sum_{u \in A \cap B} deg_G(u, V(G) \setminus (A \cup B)) \le$$
$$\le 2\alpha n^2 + |A \cap B| \left(|A \cap B| + 1\right),$$

  or

$$|A \cap B| \left(\frac{n}{2} - |A \cap B| - 1 \right) \le 2\alpha n^2,$$

  a contradiction under the given conditions.

- $|A \cap B| \ge (1 - \sqrt{\alpha}) \frac{n}{2}$. 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] G. A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* 2 (1952), 68-81.

[4] R.J. Gould, Advances on the Hamiltonian problem - a survey, *Graphs and Combinatorics* 19 (2003), pp. 7-52.

[5] A. Kaneko, Yoshimoto, On a Hamiltonian cycle in which specified vertices are uniformly distributed, *Journal of Combinatorial Theory, Ser. B* 81 (2001), pp. 100-109.

[6] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of a packing conjecture of Bollobás, *Combinatorics, Probability and Computing* 4 (1995), pp. 241-255.

[7] J. Komlós, G. N. Sárközy and E. Szemerédi, On the square of a Hamiltonian cycle in dense graphs, *Random Structures and Algorithms* 9 (1996), pp. 193-211.

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

[9] J. Komlós, G. N. Sárközy and E. Szemerédi, An algorithmic version of the Blow-up Lemma, *Random Structures and Algorithms* 12 (1998), pp. 297-312.

[10] J. Komlós, G. N. Sárközy and E. Szemerédi, On the Pósa-Seymour conjecture, *Journal of Graph Theory* 29 (1998), pp. 167-176.

[11] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Annals of Combinatorics*, 2 (1998), pp. 43-60.

[12] J. Komlós, G. N. Sárközy, E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Mathematics* 235 (2001), pp. 255-269.

[13] L. Ng, M. Schultz, k-ordered Hamiltonian graphs, *Journal of Graph Theory* 1 (1997), pp. 45-47.

[14] G. N. Sárközy, S. Selkow, E. Szemerédi, On the number of Hamiltonian cycles in Dirac graphs, *Discrete Mathematics* 265 (2003), pp. 237-250.

[15] E. Szemerédi, Regular partitions of graphs, Colloques Internationaux C.N.R.S. Nº 260 - *Problèmes Combinatoires et Théorie des Graphes*, Orsay (1976), 399-401.