A fast parallel algorithm for finding Hamiltonian cycles in dense 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

Abstract

Suppose $0 < \eta < 1$ is given. We call a graph, $G$, on $n$ vertices an $\eta$-Chvatal graph if its degree sequence $d_1 \leq d_2 \leq \ldots \leq d_n$ satisfies: for $k < n/2$, $d_k \leq \min\{k + \eta n, n/2\}$ implies $d_{n-k-\eta n} \geq n - k$. (Thus for $\eta = 0$ we get the well-known Chvatal graphs.) An $NC^4$-algorithm is presented which accepts as input an $\eta$-Chvatal graph and produces a Hamiltonian cycle in $G$ as an output. This is a significant improvement on the previous best $NC$-algorithm for the problem, which finds a Hamiltonian cycle only in Dirac graphs ($\delta(G) \geq n/2$ where $\delta(G)$ is the minimum degree in $G$).

1 Introduction

1.1 Notations and definitions

For basic graph concepts see the monograph of Bollobás [4].
$V(G)$ and $E(G)$ denote the vertex-set and the edge-set of the graph $G$. $(A, B, E)$ denotes

*Research supported in part by the National Science Foundation under Grant No. DMS-0456401.
a bipartite graph $G = (V, E)$, where $V = A \cup B$, and $E \subseteq A \times B$. For a graph $G$ and a subset $U$ of its vertices, $G\mid_U$ is the restriction to $U$ of $G$. If $A \subseteq V(G)$ and $B \subseteq V(G)$ then $N_A(B)$ denotes the set of the neighbors of vertices of $B$ in $A$. The size of $N_A(v)$ is called the degree of $v$ in $A$, $\deg_A(v)$. $\deg(v) = \deg_G(v)$. $\delta(G)$ stands for the minimum, and $\Delta(G)$ for the maximum degree in $G$. $P_l$ ($C_l$) denotes the path (cycle) of length $l$ (counting edges). When $A, B$ are 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 $B$. In particular, we write $\deg(v, U) = e(\{v\}, U)$ for the number of edges from $v$ to $U$. 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 $d(A) = d(A, A) = 2|E(G\mid_A)|/|A|^2$.

**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.

We will often say simply that “the pair $(A, B)$ is $\varepsilon$-regular” with the graph $G$ implicit.

**Definition 2.** $(A, B)$ is $(\varepsilon, \delta)$-super-regular if it is $\varepsilon$-regular and

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

As the model of computation we choose the weakest possible version of a PRAM, in which concurrent reads or writes of the same location are not allowed (EREW, see [10] for a discussion of the various PRAM models.) When researchers investigate the parallel complexity of a problem, the main question is whether a polylogarithmic running time is achievable on a PRAM containing a polynomial number of processors. If the answer is positive than the problem and the corresponding algorithm are said to belong to class $NC$ introduced in [22]. When the running time is $O((\log n)^i)$, the algorithm is in $NC^i$.

1.2 Fast parallel algorithms for finding subgraphs in dense 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.
In a series of papers we developed a method based on the Regularity Lemma and the Blow-up Lemma for finding certain spanning subgraphs in dense graphs (see [12]-[18], [26]). Typical examples are spanning trees (Bollobás-conjecture, see [12]), Hamiltonian cycles or powers of Hamiltonian cycles (Pósa-Seymour conjecture, see [16, 17]) or $H$-factors for a fixed graph $H$ (Alon-Yuster conjecture, see [18]). Since both the Regularity Lemma and the Blow-up Lemma have now fast parallel algorithmic implementations (see [2] and [15]), the above existential results have fast parallel algorithmic versions.

In this paper, by using the above method, we study the problem of finding a Hamiltonian cycle in a graph $G$. As this is one of the most famous $NP$-complete problems, to solve the problem for general graphs in $NC$ is hopeless. However, in some subclasses of graphs we have a chance. For instance, there are known classes of graphs where all the members are Hamiltonian. One class is the tournaments (see [27]). Another class is the Dirac graphs where call a graph $G = (V, E)$ a Dirac-graph, if $\delta(G) \geq \frac{n}{2}$. Dirac’s classical theorem ([3], [4], [7]) guarantees the existence of a Hamiltonian cycle in a Dirac graph. Goldberg proposed the problem at STOC’87, whether we can construct one such a cycle in $NC$. [6] answered the question affirmatively. In [6], they also posed the problem, whether their result can be extended to wider classes of graphs, known to be Hamiltonian (see [3], [4]). They indicated the difficulty of the problem, by showing that solving the Hamiltonian cycle problem for graphs with $\delta(G) \geq \alpha n$ (where $0 < \alpha < \frac{1}{2}$) is just as hard as the original Hamiltonian cycle problem; it is $NP$-complete. The algorithm we present is the first such extension, it brakes the $\frac{1}{2}$-density barrier for a class of graphs and it is close to being best possible. We call a graph $G = (V, E)$ a Chvatal graph if its degree sequence satisfies:

$$d_{n-k} \geq n-k.$$  \hspace{1cm} (1)

Chvatal proved (see e.g. [4]) that this generalization of the Dirac condition still guarantees the existence of a Hamiltonian cycle, and further this is the weakest possible such condition. More precisely, if the degree sequence does not satisfy (1), then we can construct a graph with a degree sequence majorizing this degree sequence and without a Hamiltonian cycle. Furthermore, Bondy and Chvatal ([5]) designed a sequential, polynomial time algorithm which finds a Hamiltonian cycle in Chvatal graphs, but the algorithm seems inherently sequential. The obvious question is whether there is an $NC$ algorithm for the same task. In this paper we come quite close to this ultimate goal. Let $0 < \eta < 1$ be fixed. We call a graph $G = (V, E)$ an $\eta$-Chvatal graph if its degree sequence satisfies:

$$d_{n-k} \geq n-k.$$  \hspace{1cm} (2)
Thus for \( \eta = 0 \) we get back the Chvatal condition. From the definition we can also see that this is a much wider class of graphs than the Dirac graphs. In this paper we show how to construct the Hamiltonian cycle in \( NC \) in \( \eta \)-Chvatal graphs.

**Theorem 1.** Let \( 0 < \eta < 1 \) be fixed. We can construct in \( NC^4 \) a Hamiltonian cycle in an \( \eta \)-Chvatal graph.

We note that there is also a randomized parallel algorithm for the problem [8]. However, Theorem 1 is the first deterministic algorithm in \( NC \) that goes beyond the Dirac condition.

We also note that an earlier, weaker version of this paper has appeared in [25] (see also [19] and [20]).

## 2 The main tools

In the proof the Regularity Lemma [28] plays a central role. Here we will use the following variation of the lemma. For a proof, see [2] and [20].

**Lemma 2 (Regularity Lemma – Algorithmic 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 an \( NC^1 \)-algorithm that finds a partition of the vertex-set \( V \) into \( l + 1 \) sets (so-called 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 \( v \in V \),
- \( G'\mid_{V_i} = \emptyset \) (\( V_i \) are independent in \( G' \)),
- all pairs \( G'\mid_{V_i \times V_j}, 1 \leq i < j \leq l \), are \( \varepsilon \)-regular, each with a density \( \theta \) 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 at most \( \delta \).
As we mentioned in the introduction, an application of the Regularity Lemma in graph theory is now often coupled with an application of the Blow-up Lemma (see [14] for the original, [15] for an algorithmic version and [23] and [24] for two alternate proofs). Here we use a very special case of the algorithmic 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 an \(NC^4\)-algorithm that finds a Hamiltonian path starting with \(x\) and ending with \(y\) (see [15]). More precisely.

**Lemma 3.** For every \(\delta > 0\) there are \(\varepsilon_0, n_0 > 0\) such that if \(\varepsilon \leq \varepsilon_0\) and \(n \geq n_0\), \(G = (A,B)\) is an \((\varepsilon, \delta)\) super-regular pair with \(|A| = |B| = n\) and \(x \in A, y \in B\), then there is an \(NC^4\)-algorithm that finds a Hamiltonian path in \(G\) starting with \(x\) and ending with \(y\).

We will also use the \(NC^4\) algorithm for the maximal independent set problem. Recall that a subset \(I\) of the vertices of a graph \(G\) is independent if there are no edges between any two vertices in \(I\). An independent set \(I\) is maximal if it is not a proper subset of any other independent set. Karp and Wigderson ([11]) were the first to give an \(NC^4\)-algorithm for this problem.

**Lemma 4.** It is possible to construct a maximal independent set in a graph in \(NC^4\).

Better algorithms were later described in [1], [9] and in [21]. We call this the MIS algorithm.

3 Outline of the proof

We will assume throughout the paper that \(n\) is sufficiently large (otherwise clearly we can find a Hamiltonian cycle in \(NC^4\)). We will use the following main parameters

\[0 < \varepsilon \ll \delta \ll \alpha \ll \eta \ll 1,\]

where \(\alpha\) depends on \(\eta\), \(\delta\) depends on \(\alpha\) and \(\eta\) and \(\varepsilon\) depends on \(\delta\), \(\alpha\) and \(\eta\), and \(a \ll b\) means that \(a\) is sufficiently small compared to \(b\). For simplicity we do not compute the actual dependencies, although it could be done.

Let us consider an \(\eta\)-Chvatal graph \(G\) of order \(n\). Then its degree sequence \(d_1 \leq d_2 \leq \ldots \leq d_n\) satisfies (2). Note that in particular (2) implies

\[\delta(G) = d_1 \geq \eta n.\]

We must show that we can find in \(NC^4\) a Hamiltonian cycle in \(G\). 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 1 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 the high level description of our algorithm is the following.

**Program Find-Hamiltonian-cycle**

*Given:* A $\eta$-Chvatal graph $G$ on $n$ vertices.

*Compute:* A Hamiltonian cycle of $G$.

- **Step 1:** We apply Lemma 2 for $G$, with $\varepsilon$ and $\delta$ as in (3). We get a partition of $V(G') = \bigcup_{0 \leq i \leq l} V_i$. 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 $\varepsilon$-regular pair in $G'$ with density exceeding $\delta$.

- **Step 2:** Find a perfect matching $M$ in $G_r$ (we will show that one must exist). Put $|M| = m = \lfloor \frac{l}{2} \rfloor$. Denote the $i$-th pair in $M$ by $(V_i^1, V_i^2)$ for $1 \leq i \leq m$.

- **Step 3:** Put the cluster of $G_r$ that is not covered by $M$ (in case $l$ is odd) and some additional exceptional vertices (to achieve super-regularity) into $V_0$, denote the resulting set still by $V_0$ for simplicity.

- **Step 4:** Redistribute the vertices in $V_0$ among the clusters in $M$ in such a way, that we preserve super-regularity, and we add only a “few” vertices to each cluster.

- **Step 5:** Find short connecting paths $P_i$ between the consecutive edges in the matching $M$ (for $i = m$ the next edge is $i = 1$). These paths will be parts of the final Hamiltonian cycle.

- **Step 6:** Make some adjustments to achieve that we have the same number of vertices left in $V_i^1$ and in $V_i^2$ for each $1 \leq i \leq m$.

- **Step 7:** Apply Lemma 3 in each $(V_i^1, V_i^2), 1 \leq i \leq m$ to close the Hamiltonian cycle.

In the next section, in the non-extremal case, we will discuss the above steps one-by-one. Finally in Section 5 we show that Theorem 1 is true in the extremal case as well.
4 The non-extremal case

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

4.1 Step 1

Using the fact that $\deg G'(v) > \deg G(v) - (\delta + \varepsilon)n$, we will show that $G_r$ satisfies a similar degree condition as the original graph $G$. In fact, let us denote the degree sequence of $G_r$ by $d_{r1}, d_{r2}, \ldots, d_{rl}$. We will show that $d_{r1} \geq (\eta - 2\delta)l$ and that $k < \frac{l}{2}, d_{rk} \leq \min\{k + (\eta - 2\delta)l, \left(\frac{1}{2} - 2\delta\right)l\}$ implies $d_{r(l-k-(\eta-\varepsilon))l} \geq l - k - 2\delta l.$ (6)

We know that in $G'$ the neighbors of $u \in V_i$ can only be in $V_0$ and in the clusters which are neighbors of $V_i$ in $G_r$. Then $d_{r1} \geq (\eta - 2\delta)l$ is immediate from (4) and (5). For the second half of the statement let us assume that for a $1 \leq k < \frac{l}{2}$ we have $d_1^r \leq \ldots \leq d_k^r \leq \min\{k + (\eta - 2\delta)l, \left(\frac{1}{2} - 2\delta\right)l\}.$ (7)

We must show that for this $k$ (6) holds. (5) and (7) imply that we have at least $kL$ vertices $u \in V(G)$ for which $\deg_G(u) < \deg_{G'}(u) + (\delta + \varepsilon)n \leq \varepsilon n + d_k^rL + (\delta + \varepsilon)n \leq d_k^rL + 2\delta n \leq \min\{kL + \eta n, \frac{n}{2}\}.$

Hence in $G$ $d_{kL} \leq \min\{kL + \eta n, \frac{n}{2}\}.$

But then (2) implies that $d_{n-kL-nm} \geq n - kL.$ In this case using (5) there are at least $kL + \eta n$ vertices $v \in V(G)$ for which $\deg_{G'}(v) > \deg_G(v) - (\delta + \varepsilon)n \geq n - kL - (\delta + \varepsilon)n.$ (8)

This and $|V_0| \leq \varepsilon n$ imply that there are at least $k + (\eta - \varepsilon)l$ clusters $V_i$ which contain at least one vertex satisfying (8). But then for these clusters $V_i$ we have $\deg_{G'}(V_i) \geq l - k - 2\delta l,$ and thus proving (6).
4.2 Step 2

We find a maximum matching $M$ in $G_r$ (here we take advantage of the fact that $l$ is a constant, so we do not have to worry about the running time). We will prove that $M$ is a perfect matching. Assume indirectly that it is not, and consider two clusters $V_i$ and $V_j$ from the independent set $V(G_r) \setminus V(M)$. We will show that there is an alternating path $P$ with respect to $M$ connecting $V_i$ and $V_j$. But then we can increase the size of $M$ by one, a contradiction, by exchanging the matching edges on $P$ with the non-matching edges on $P$.

The existence of the alternating path $P$ will follow from the following fact (this fact will be used in Steps 5 and 6 as well).

**Fact 5.** If $V_i, V_j \in V(G_r)$ then there are at least $\delta^2 l$ internally disjoint alternating paths (with respect to $M$) of length at most $1/\delta$ connecting $V_i$ and $V_j$ in $G_r$, where the first and last edges on the paths are non-matching edges.

**Proof of Fact 5:** First we will show the following expansion property. For all $X \subset V(G_r)$, $1 \leq |X| < l/2$ we have

$$|N_{G_r}(X)| \geq \min \left\{|X| + \frac{\eta}{4} l, \left(\frac{1}{2} - 2\delta\right) l\right\}.$$  \hfill (9)

If $1 \leq |X| \leq \frac{\eta}{2} l$, then take an arbitrary cluster $V \in X$, and using (3) we have

$$|N_{G_r}(X)| \geq \text{deg}_{G_r}(V) \geq d_r^{\ast} \geq (\eta - 2\delta) l \geq \frac{\eta}{2} l + \frac{\eta}{4} l \geq |X| + \frac{\eta}{4} l,$$

proving (9) in this case. Thus we may assume $\frac{\eta}{2} l < |X| < \frac{l}{2}$. Denote $k = |X| - \frac{\eta}{2} l$. Then we have $1 \leq k < l/2$. We have two cases:

**Case 1:** $d_r^{\ast} \geq \min \left\{k + (\eta - 2\delta) l, \left(\frac{1}{2} - 2\delta\right) l\right\}$.

In this case (9) is obvious.

**Case 2:** $d_r^{\ast} < \min \left\{k + (\eta - 2\delta) l, \left(\frac{1}{2} - 2\delta\right) l\right\}$.

From (6), we get

$$d_r^{\ast} \geq l - k - 2\delta l.$$

Thus in this case the clusters with the

$$k + (\eta - \varepsilon) l \geq |X| + \frac{\eta}{4} l$$

largest degrees have at least $\frac{\eta}{4} l$ neighbors in $X$, and therefore they are in $N_{G_r}(X)$, proving (9) again.
In order to prove Fact 5 first let \( V_i, V_j \in V(G_r) \setminus V(M) \). We will define a sequence of sets \( N_1, N_2, \ldots, (N'_1, N'_2, \ldots) \) in \( V(G_r) \) such that the clusters in \( N_i \) (\( N'_i \)) are reachable from \( V_i \) (\( V_j \)) by an alternating path of length \( i \) where the first edge is a non-matching edge. Let \( N_1 = N_{G_r}(V_i) \), and \( N_2 \) is the set of neighbors in \( M \) of the vertices in \( N_1 \). Similarly, in general if \( N_{2i} \) is already defined, then \( N_{2i+1} = N_{G_r}(N_{2i}) \), and \( N_{2(i+1)} \) is the set of neighbors in \( M \) of the vertices in \( N_{2i+1} \). Here we used the fact that the clusters of \( N_{2i+1} \) are always matched in \( M \), since otherwise we could get a bigger matching, thus a contradiction, just as above.

Then the expansion property (9) implies that with \( N = N_{2\lceil \frac{\eta}{2} \rceil} \) we have

\[
|N| \geq \left( \frac{1}{2} - 2\delta \right) l. \tag{10}
\]

Similarly, the sequence \( N'_1, N'_2, \ldots \) can be defined and with \( N' = N'_{2\lceil \frac{\eta}{2} \rceil} \) we have

\[
|N'| \geq \left( \frac{1}{2} - 2\delta \right) l. \tag{11}
\]

Then (3), (10), (11) and fact that here EC does not hold clearly imply that

\[
d(N, N') \gg \delta
\]

and thus we have “many” edges between \( N \) and \( N' \). This gives one alternating path \( P \) of length at most \( 10/\eta \ll 1/\delta \) between \( V_i \) and \( V_j \) in \( G_r \). We remove the internal vertices of \( P \) from \( G_r \) and repeat the above procedure. It is not hard to see that the above procedure goes through again and by iterating the above procedure \( \delta^2 l \) times we get Fact 5. Indeed, the total number of internal vertices on these paths is only at most \( \delta^2 l \frac{1}{4} = \delta l \), and subtracting this much does not change the above procedure. This shows that \( M \) is a perfect matching, and then we get Fact 5 for every \( V_i, V_j \in V(G_r) \).

### 4.3 Step 3

We already have an exceptional set \( V_0 \) of vertices in \( G \). We add the cluster of \( G_r \) that is not covered by \( M \) (in case \( l \) is odd) and some additional exceptional vertices (to achieve super-regularity) into \( V_0 \), denote the resulting set still by \( V_0 \) for simplicity. From \( V'_1 \) (and similarly from \( V'_2 \)) in parallel we remove all vertices \( u \) for which \( \text{deg}(u, V'_2) < (\delta - \varepsilon)|V'_2| \). \( \varepsilon \)-regularity guarantees that at most \( \varepsilon|V'_1| \leq \varepsilon L \) such vertices exist in each cluster \( V'_1 \). Thus we still have

\[
|V_0| \leq 3\varepsilon n. \tag{12}
\]
4.4 Step 4

We will redistribute the vertices in $V_0$ in blocks of size $\lfloor \varepsilon L \rfloor$; in particular the number of blocks is a constant. Let us take the first block of $\lfloor \varepsilon L \rfloor$ vertices in $V_0$. For each vertex $w$ in this block in parallel we find a pair $(V_1^1, V_2^1)$ such that either

$$\deg(w, V_1^1) \geq \delta|V_1^1|,$$

or

$$\deg(w, V_2^1) \geq \delta|V_2^1|.$$  \hspace{1cm} (13)

or

$$\deg(w, V_2^1) \geq \delta|V_2^1|.$$  \hspace{1cm} (14)

(3) and (4) imply that for every vertex $w$ there is a pair $(V_1^1, V_2^1)$ for which either (13) or (14) holds. In case (13) holds we assign $w$ to $V_2^1$, and in case (14) holds we assign $w$ to $V_1^1$. After a block is finished, since during the whole process the number of forbidden pairs is at most $4 \sqrt{3}L$ vertices have been assigned to it from all the blocks so far, and in the next block we will not consider this pair in (13) and in (14). Then using (4) and (12) we can redistribute all the vertices of $V_0$ among the pairs, since during the whole process the number of forbidden pairs is at most $4 \sqrt{3}L$.

4.5 Step 5

First using Fact 5 we can find $m$ connecting paths $P_i^r$ in $G_r$ from $V_j^0$ to $V_j^{i+1}$ for every $1 \leq i \leq m$ (for $i = m$ we go from $V_j^m$ back to $V_j^1$). Note that these paths in $G_r$ may not be internally vertex disjoint. Note also that Fact 5 actually gives alternating paths, but now we just look at these as ordinary paths. From these paths $P_i^r$ in $G_r$ we can construct vertex disjoint connecting paths $P_i$ in $G$ connecting a typical vertex $v_0^1$ of $V_j^0$ to a typical vertex $v_j^{i+1}$ of $V_j^{i+1}$. More precisely we construct $P_1$ with the following simple greedy strategy. Denote $P_1^r = (p_1, \ldots, p_t), 2 \leq t \leq 1/\delta$, where according to the definition $p_1 = V_1^1$ and $p_t = V_1^2$. Let the first vertex $u_1 (= v_0^1)$ of $P_1$ be a vertex $u_1 \in V_1^1$ for which $\deg_G(u_1, p_2) \geq (\delta - \varepsilon)L$ and $\deg_G(u_1, V_1^1) \geq (\delta - \varepsilon)L$. By $\varepsilon$-regularity most of the vertices satisfy this in $V_1^1$. The second vertex $u_2$ of $P_1$ is a vertex $u_2 \in p_2 \cap N_G(u_1)$ for which $\deg_G(u_2, p_3) \geq (\delta - \varepsilon)L$. Again by $\varepsilon$-regularity most vertices satisfy this in $p_2 \cap N_G(u_1)$. The third vertex $u_3$ of $P_1$ is a vertex $u_3 \in p_3 \cap N_G(u_2)$ for which $\deg_G(u_3, p_4) \geq (\delta - \varepsilon)L$. We continue in this fashion, finally the last vertex $u_t (= v_j^1)$ of $P_1$ is a vertex $u_t \in p_t \cap N_G(u_{t-1})$ for which $\deg_G(u_t, V_j^1) \geq (\delta - \varepsilon)L$.

Then we move on to the next connecting path $P_2$. Here we follow the same greedy procedure, we pick the next vertex from the next cluster in $P_2^r$. However, if the cluster has occurred already on the path $P_i^r$ (or on any other connecting paths later in the procedure), then we just have to make sure that we pick a vertex that has not been used so far. Since
the total number of vertices on the connecting paths will be a constant, this is feasible. (Furthermore, this also implies that the running time of this step is a constant as well.)

We continue in this fashion and construct the vertex disjoint connecting paths $P_i$ in $G$, $1 \leq i \leq m$. These will be parts of the final Hamiltonian cycle in $G$. We remove the internal vertices of these paths from $G$. In case the number of remaining vertices is odd, since $P_i$ does not have to be an alternating path, we can clearly make it one cluster longer. Thus we may always assume that the number of remaining vertices is even.

4.6 Step 6

At this point we might have a small discrepancy ($\leq 2\sqrt{\varepsilon}|V_i|$) among the remaining vertices in $V_i$ and in $V_{i'}$ in a pair. Therefore, we have to make some adjustments. Let us take a pair $(V_i, V_{i'})$ with a discrepancy $d \geq 2$ (if one such pair exists), say $|V_i| = |V_{i'}| + d$ (only remaining vertices are considered). Using Fact 5 we find an alternating path (with respect to $M$) in $G'$ starting with $V_i$ and ending with $V_{i'}$. Let us denote this path by

$$V_i, V_{i'1}, V_{i2}, V_{i'1}, \ldots, V_{i't}, V_{i'1}$$

where $t \leq 1/\delta$. (15)

(Here for simplicity we assumed that on this path all the pairs are visited in the order $V_i, V_{i'}$, otherwise it is similar). In parallel we remove $d$ typical vertices from $V_i$ and we add them to $V_{i'1}$, then we remove $d$ typical vertices from $V_{i'1}$ and we add them to $V_{i'2}$, etc., finally we remove $d$ typical vertices from $V_{i't}$ and we add them to $V_{i'}$.

Now we are closer to the perfect distribution by one more pair, and by iterating this procedure we can assure that the discrepancy in every pair is at most 1. Furthermore, similarly as in Step 4, after handling each such pair, we declare a pair forbidden if at least $\sqrt{\varepsilon}L$ vertices have been added to it or removed from it during the whole process in Step 6 so far. Then we will not consider a forbidden pair in the next iteration. (3) and Fact 5 imply that we can always find the alternating path in (15), since during the whole process in Step 6 the number of forbidden pairs is at most $\sqrt{\varepsilon}L \ll \delta^2 L$.

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 $(V_i, V_{i'})$ and $(V_{i'}, V_{i''})$ is one such pair with $|V_i| = |V_{i'}| + 1$ and $|V_{i'}| = |V_{i''}| + 1$ (otherwise similar), then similar to the construction above, we find an alternating path in $G'$ between $V_i$ and $V_{i''}$, and we move one typical vertex of $V_i$ through the intermediate clusters to $V_{i''}$. 11
4.7 Step 7

Thus we may assume that the distribution is perfect, in every pair \((V_i^1, V_j^2)\) we have the same number of vertices left. Furthermore, each pair \((V_i^1, V_j^2)\) is super-regular with somewhat weaker parameters (say \((\sqrt{\varepsilon}, \delta/2)\)-super-regular). In this case Lemma 3 closes the Hamiltonian cycle in every pair.

We note that here in the non-extremal case we can also prove the following. For every pair of vertices \(u, v \in V(G)\), we can find in \(NC^4\) a Hamiltonian path in \(G\) connecting \(u\) and \(v\). Indeed, the only difference in the above is that instead of the connecting path \(P_m\), we will have one connecting path \(P_1^m\) connecting \(v^m_2\) and \(v\) and another one \(P_2^m\) connecting \(u\) and \(v^m_1\); all the other details above are the same. This fact will be used later in the extremal case.

5 The extremal case

First we treat two special cases and then we handle the general extremal case.

**Case 1:** Assume that we have a partition \(V(G) = A_1 \cup A_2\) with \(|A_1| = \lceil \frac{n}{2} \rceil\) and
\[
d(A_1, A_2) < \sqrt[4]{\alpha}.
\]
Thus the bipartite graph between \(A_1\) and \(A_2\) is very sparse.

First we claim, that in this case in \(G\)
\[
d_k > \min \left\{ k + \eta n, \frac{n}{2} \right\}
\]
always holds for \(k < \frac{n}{2}\). Suppose (17) is not true, thus for some \(k < \frac{n}{2}\)
\[
d_k \leq \min \left\{ k + \eta n, \frac{n}{2} \right\}.
\]
(2) then gives
\[
d_{n-k-\eta m} \geq n - k.
\]
But from (19), \(k \leq (1 - \eta)\frac{n}{2}\) follows. Otherwise
\[
n - k - \eta m \leq n - (1 - \eta)\frac{n}{2} - \eta m = (1 - \eta)\frac{n}{2} < k \quad \text{and thus}
\]
\[
\frac{n}{2} < n - k \leq d_{n-k-\eta m} \leq d_k
\]
a contradiction with (18). Thus we can assume that $k \leq (1 - \eta)\frac{n}{2}$ and that (19) holds. Then for at least $k + \eta n \geq \eta n$ vertices $v$

$$deg_G(v) \geq n - k \geq n - (1 - \eta)\frac{n}{2} = \frac{n}{2} + \eta \frac{n}{2}.$$ 

Thus in either $A_1$ or $A_2$ (suppose in $A_1$) we have $\geq \eta n$ from these vertices. But then for these vertices $v$

$$deg_{A_2}(v) \geq \eta|A_2|,$$

a contradiction with (16) from (3). Thus we can assume that (17) holds.

We define **exceptional** vertices $v \in A_i, i \in \{1, 2\}$, as

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

Note that from the density condition (16), the number of exceptional vertices in $A_i$ is at most $\sqrt{\alpha}|A_i|$. In parallel we remove the exceptional vertices from each set and then we add each extra vertex to the set where it has more neighbors. We still denote the sets by $A_1$ and $A_2$. Thus in $G|_{A_i}, i \in \{1, 2\}$, it is certainly true that apart from at most $3\sqrt{\alpha}|A_i|$ exceptional vertices for all the vertices $v \in A_i$ we have

$$deg_{G|_{A_i}}(v) \geq deg_G(v) - 3\sqrt{\alpha}|A_i|,$$  

(20)

and for the exceptional vertices using (4) we have

$$deg_{G|_{A_i}}(v) \geq \eta|A_i|/2.$$  

(21)

(3), (17), (20) and (21) imply that if we denote the degree sequence of $G|_{A_i}$ by $d_1^i, d_2^i, \ldots, d_{|A_i|}^i$, then we have for all $1 \leq k \leq |A_i|$ the following.

$$d_k^i > \min \left\{ k + \frac{\eta}{4}|A_i|, (1 - \eta)|A_i| \right\}.$$  

(22)

Thus in particular $G|_{A_i}, i \in \{1, 2\}$ are $\eta/4$-Chvatal graphs, and furthermore (22) clearly implies that the extremal case EC cannot hold for them. (17) implies that we can find two independent edges (bridges) $e_1 = (u_1, v_1)$ and $e_2 = (u_2, v_2)$ between $A_1$ and $A_2$, where $u_1, u_2 \in A_1, v_1, v_2 \in A_2$. Running the non-extremal version of our algorithm twice we can find a Hamiltonian path in $G|_{A_1}$ connecting $u_1$ and $u_2$ (see the last remark at the end of Step 7) and a Hamiltonian path in $G|_{A_2}$ connecting $v_1$ and $v_2$. This gives us the desired Hamiltonian cycle in $G$. 

13
**Case 2:** Assume next that there is a partition $V(G) = A_1 \cup A_2$ with $|A_1| = \lfloor \frac{n}{2} \rfloor$ and $d(A_1) < \sqrt[4]{\alpha}$. Thus the graph $G|_{A_1}$ is very sparse.

A vertex $v \in A_1$ is called exceptional if it has a relatively large neighborhood in $A_1$, more precisely if we have

$$\deg_G(v, A_1) \geq \sqrt[4]{\alpha}|A_1|.$$ 

From $d(A_1) < \sqrt[4]{\alpha}$ we get that the number of exceptional vertices in $A_1$ is at most $\sqrt[4]{\alpha}|A_1|$.

In parallel for each exceptional vertex $v$, we add $v$ to $A_2$ if it has more neighbors in $A_1$ than in $A_2$. We still denote the resulting sets by $A_1$ and $A_2$. Thus in $G|_{A_1}$ it is certainly true that apart from at most $\sqrt[4]{\alpha}|A_1|$ exceptional vertices for all the vertices $v \in A_1$ we have

$$\deg_G(v, A_2) \geq \deg_G(v) - \sqrt[4]{\alpha}|A_1|,$$ 

(23)

and for the exceptional vertices we have

$$\deg_G(v, A_2) \geq \eta|A_2|/2.$$ 

(24)

Let $k = |A_1| - \lfloor \frac{n}{2} \rfloor$. Similarly as above in the proof of (17) for this $k$ we have

$$d_k > \min \left\{ k + \eta n, \frac{n}{2} \right\}.$$ 

(25)

From (23) and (25) it follows that there are at least $\frac{n}{2}$ vertices $v \in A_1$ for which

$$\deg_G(v, A_2) \geq (1 - 2\sqrt[4]{\alpha})|A_2|.$$ 

This in turn implies that we can have at most $4\sqrt[4]{\alpha}|A_2|$ exceptional vertices $v \in A_2$ for which

$$\deg_G(v, A_1) \geq \frac{\eta}{4}|A_1|$$ 

(26)

does not hold. In parallel we remove each of these exceptional vertices from $A_2$ and add them to $A_1$. We still denote the resulting sets by $A_1$ and $A_2$. We have $d = ||A_1| - |A_2|| \leq 5\sqrt[4]{\alpha}|A_2|$.

The rest of the proof in this case will be a bipartite adaptation of the proof in the non-extremal case. Therefore we are not going into details, we just point out the major differences. First, since the “heart” of the non-extremal case was the expansion property (9), here we need a bipartite version of this as well. First (2), (3), (23) and (24) imply that similarly as in (9) for all $X \subset A_1$, $1 \leq |X| \leq |A_1|$ we have

$$|N_{A_2}(X)| \geq \min \left\{ |X| + \frac{\eta}{4}|A_2|, (1 - \eta)|A_2| \right\}.$$ 

(27)
But then (3), (26) and (27) imply in turn that we have a similar expansion property from the other direction as well; more precisely for all \( Y \subset A_2, 1 \leq |Y| \leq |A_2| \) we have

\[
|N_{A_1}(Y)| \geq \min \left\{ |Y| + \frac{\eta}{3}|A_1|, (1 - \eta)|A_1| \right\}.
\]  

(28)

Indeed, for small \(|Y|\) we get this from (26), and for larger \(|Y|\) we get this from (27) by choosing \( X = A_1 \setminus N_{A_1}(Y) \) since then \( N(X) \subset A_2 \setminus Y \).

Before starting the bipartite version of the non-extremal case, we need one more technical step; namely we would like to achieve that \( d = |A_1| - |A_2| = 0 \). Without loss of generality assume \( |A_2| > |A_1| \). If there is a vertex \( v \in A_2 \) for which \( deg_{A_2}(v) \geq \eta^2|A_2| \) then we put \( v \) in \( A_1 \), and thus reducing \(|A_2| - |A_1|\). Therefore we may assume that there is no such \( v \in A_2 \) and let us denote \( 0 < d = |A_2| - |A_1| \leq 5\sqrt[4]{\alpha}|A_2| \). By (2), we know that in \( A_2 \) there are still at least \( \eta n \) vertices \( u \in A_2 \) for which \( deg_{A_2}(u) \geq \frac{d}{2} \), namely these are vertices with \( deg_{E}(u) \geq \frac{d}{2} \). Denote the set of edges leaving these vertices in \( A_2 \) by \( E \). Running MIS on the linegraph defined on \( E \) and using the maximum degree condition we can find in \( NC^4 d \) independent edges from \( E \), denoted by \( e_1 = (u_1, v_1), e_2 = (u_2, v_2), \ldots, e_d = (u_d, v_d) \). Next we will find in \( NC^4 d \) short vertex disjoint connecting paths in \( G|_{A_1 \times A_2} \) between \( v_i \) and \( u_{i+1} \) for \( 1 \leq i \leq d - 1 \). Similarly to the non-extremal case the expansion property (28) implies that for each \( 1 \leq i \leq d - 1 \) there are many internally disjoint connecting paths between \( v_i \) and \( u_{i+1} \). Then running MIS on the appropriately defined auxiliary graph (the vertices are the connecting paths, and we put an edge between two connecting paths if they share a common vertex) we can select \( d - 1 \) vertex disjoint connecting paths connecting \( v_i \) and \( u_{i+1} \) from \( 1 \leq i \leq d - 1 \). We add one more arbitrary edge of \( G|_{A_1 \times A_2} \) to the path from \( v_d \); denote its other endvertex by \( u_{d+1} \in A_1 \). This way we get a path \( P \) connecting \( u_1 \) and \( u_{d+1} \) that contains the \( d \) edges \( e_1, \ldots, e_d \). We remove the internal vertices of this path \( P \) from \( G \).

Now we just have to find in \( NC^4 d \) a Hamiltonian path in the leftover (where now we have the same number of vertices on the two sides) connecting \( u_1 \) and \( u_{d+1} \). This together with \( P \) gives us the desired Hamiltonian cycle in \( G \) in this case.

Thus now we may assume \( |A_1| = |A_2| \), (27) and (28), and we have to find a Hamiltonian cycle in \( G|_{A_1 \times A_2} \). We follow a bipartite adaptation of the non-extremal case. Note that (27) and (28) clearly imply that the extremal case cannot hold here. We apply the bipartite version of the Regularity Lemma to get two partitions

\[
A_1 = V_0^1 + V_1^1 + V_2^1 + \ldots + V_l^1,
\]

\[
A_2 = V_0^2 + V_1^2 + V_2^2 + \ldots + V_l^2.
\]

The reduced graph \( G_r \) is a bipartite graph as well between \( A_1^r \) and \( A_2^r \) satisfying similar conditions to (27) and (28) (with somewhat weaker parameters). We take a maximum
matching \( M \) again, and we show that it is a perfect matching. For this we show that the expansion conditions imply similarly to Fact 5 that if we have a cluster \( V_i \in A_1^r \) and a cluster \( V_j \in A_2^r \) then there are many short internally disjoint alternating paths connecting \( V_i \) and \( V_j \). Note that we might not have these paths between \( V_i \) and \( V_j \) belonging to the same partite set, but fortunately we never need this, as \(|A_1|\) and \(|A_2|\) are already balanced. From this bipartite version of Fact 5, it follows again that \( M \) is a perfect matching and that we can perform all the other steps of the non-extremal case. All details can implemented again in \( NC^4 \) and are omitted here.

**Extremal Case:** 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.

- \(|A \cap B| \leq \lceil \sqrt[3]{\alpha n} \rceil\). The statement follows from Case 1. Indeed, let \( A_1 = A, A_2 = V(G) \setminus A_1 \), then clearly \( d(A_1, A_2) < \sqrt[3]{\alpha} \) if \( \alpha \ll 1 \) holds.

- \( \lfloor \sqrt[3]{\alpha n} \rfloor < |A \cap B| < (1 - \sqrt[3]{\alpha}) \frac{n}{2} \). This case is not possible under the given conditions. In fact, otherwise denote \( k = |A \cap B| - \lfloor \sqrt[3]{\alpha n} \rfloor \). Then \( 1 \leq k < n/2 \). We have two subcases:
  - **Subcase 1:** \( d_k \geq \min \{ k + \eta n, \frac{n}{2} \} \).
    In this case we have
    \[
    \sum_{u \in A \cap B} \deg_G(u, A \cup B) \geq \sqrt[3]{\alpha n} \min \left\{ \eta n - \lfloor \sqrt[3]{\alpha n} \rfloor, \frac{n}{2} - |A \cap B| \right\} \geq \alpha^{2/3} n^2,
    \]
    a contradiction with \( d(A, B) < \alpha \).
  - **Subcase 2:** \( d_k < \min \{ k + \eta n, \frac{n}{2} \} \).
    From (2), we get
    \[
    d_{n-k-\eta n} \geq n - k.
    \]
    Thus in this case the vertices with the \( k + \eta n \) largest degrees have at least \( \sqrt[3]{\alpha n} \) neighbors in \( A \cap B \). From these vertices at least \( \eta n/2 \) vertices are in \( A \cup B \). Thus again
    \[
    \sum_{u \in A \cap B} \deg_G(u, A \cup B) \geq \frac{\eta}{2} \sqrt[3]{\alpha n^2},
    \]
    a contradiction with \( d(A, B) < \alpha \).

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

This finishes the extremal case and the proof of Theorem 1.
References


