FINDING HAMILTON CYCLES IN ROBUSTLY EXPANDING DIGRAPHS

DEMETRES CHRISTOFIDES, PETER KEEVASH, DANIELA KÜHN AND DERYK OSTHUS

Abstract. We provide an NC algorithm for finding Hamilton cycles in directed graphs with a certain robust expansion property. This property captures several known criteria for the existence of Hamilton cycles in terms of the degree sequence and thus we provide algorithmic proofs of (i) an 'oriented' analogue of Dirac's theorem and (ii) an approximate version (for directed graphs) of Chvátal's theorem. Moreover, our main result is used as a tool in a recent paper by Kühn and Osthus, which shows that regular directed graphs of linear degree satisfying the above robust expansion property have a Hamilton decomposition, which in turn has applications to TSP tour domination.

1. Introduction

In this paper we study the problem of finding Hamilton cycles in directed graphs efficiently. The decision problem is one of the most famous NP-complete problems so we will restrict our attention to some specific classes of directed graphs which are known to be Hamiltonian and provide fast parallel algorithms for finding Hamilton cycles in such graphs. These algorithms immediately translate into sequential algorithms with polynomial running time. Our model of computation will be the EREW PRAM, in which concurrent reading or writing is not allowed. We say that a problem belongs to the class NC if it can be solved in polylogarithmic time on a PRAM containing a polynomial number of processors. If the algorithm has running time $O((\log n)^i)$, then we say that it belongs to the class NC^i. For a discussion of the various PRAM models, we refer the reader to [12].

By Dirac's theorem [9], one class of undirected graphs which are known to be Hamiltonian is the class of graphs with minimum degree at least $\frac{n}{2}$, where n is the order of the graph. Although Dirac's proof was not formulated in algorithmic terms, it can be easily turned into a polynomial time algorithm for finding a Hamilton cycle in such graphs. Goldberg raised the question of whether the problem of finding such a cycle belongs to NC. This question was answered affirmatively by Dahlhaus, Hajnal and Karpinski [8] who designed an NC^4 algorithm for this problem.

Following Dirac's theorem, there was a series of results by various authors giving even weaker conditions which still guarantee Hamiltonicity. Finally, Chvátal [6] showed that if the degree sequence $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$ of a graph G satisfies $d_k \geqslant k+1$ or $d_{n-k} \geqslant n-k$ whenever $k < \frac{n}{2}$, then G is Hamiltonian. Chvátal's condition is best possible in the sense that for every degree sequence $d_1 \leqslant \cdots \leqslant d_n$ not satisfying this condition, there is a non-Hamiltonian graph on n vertices whose degree sequence dominates $d_1 \leqslant \cdots \leqslant d_n$. Chvátal's original proof was not algorithmic. A sequential polynomial time algorithm for finding Hamilton cycles in such graphs was found later by Bondy and Chvátal [5]. No NC-algorithm for finding Hamilton cycles in such graphs is known yet. Recently however, Sárközy [25] proved the following approximate result.

Theorem 1. *Let $0 < \eta < 1$ be fixed and let G be a graph of order n whose degree sequence satisfies*

$$d_k > \min\{k+\eta n, n/2\} \text{ or } d_{n-k-\eta n} \geqslant n-k$$

whenever $k < \frac{n}{2}$. Then there is an NC^4 algorithm for finding a Hamilton cycle in G.

Let us now turn our attention to directed graphs (digraphs). The digraphs considered in this paper do not have loops and we allow at most 2 edges between any pair of vertices, at most one in each direction. When referring to paths and cycles in digraphs we

D. Christofides and D. Osthus were supported by the EPSRC, grant no. EP/E02162X/1.

always mean that these are directed without mentioning this explicitly. For an analogue of Dirac's theorem for digraphs it is natural to consider the *minimum semi-degree* $\delta^0(G)$ of a digraph G, which is the minimum of its minimum out-degree $\delta^+(G)$ and its minimum in-degree $\delta^-(G)$. The corresponding analogue is a theorem of Ghouila-Houri [10] which states that every digraph G on n vertices with minimum semi-degree at least $\frac{n}{2}$ contains a Hamilton cycle. Thomassen [27] asked for an analogue for oriented graphs (these are digraphs without 2-cycles). One could expect that for such graphs, a much weaker degree condition suffices. Indeed Häggkvist [13] pointed out that a minimum semi-degree of $\frac{3n-4}{8}$ is necessary and conjectured that it is also sufficient to guarantee a Hamilton cycle in any oriented graph of order n. The following approximate version of this conjecture was proved by Kelly, Kühn and Osthus [15].

Theorem 2. *For every $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that for every oriented graph G of order $n \geqslant n_0$ the following hold:*

(i) *If $\delta(G) + \delta^+(G) + \delta^-(G) \geqslant \left(\frac{3}{2} + \alpha\right) n$, then G contains a Hamilton cycle;*
(ii) *if $d^+(x) + d^-(y) \geqslant \left(\frac{3}{4} + \alpha\right) n$ whenever $xy \notin E(G)$, then G contains a Hamilton cycle.*

In particular, if $\delta^0(G) \geqslant \left(\frac{3}{8} + \alpha\right) n$, then G contains a Hamilton cycle. (Here, $\delta(G)$ denotes the minimum number of edges incident to a vertex of G.)

Finally, the conjecture of Häggkvist was proved for all large enough oriented graphs by Keevash, Kühn and Osthus [14].

What about an analogue of Chvátal's theorem for digraphs? No such analogue has yet been proved. For a digraph G, let us write $d_1^+ \leqslant \cdots \leqslant d_n^+$ for its out-degree sequence and $d_1^- \leqslant \cdots \leqslant d_n^-$ for its in-degree sequence. The following conjecture of Nash-Williams [24] would provide an analogue of Chvátal's theorem for digraphs.

Conjecture 3. *Let G be a strongly connected digraph of order n and suppose that for all $k < \frac{n}{2}$*

(i) *$d_k^+ \geqslant k + 1$ or $d_{n-k}^- \geqslant n - k$;*
(ii) *$d_k^- \geqslant k + 1$ or $d_{n-k}^+ \geqslant n - k$.*

Then G contains a Hamilton cycle.

Recently, the following approximate version of Conjecture 3 for large digraphs was proved by Kühn, Osthus and Treglown [22].

Theorem 4. *For every $\eta > 0$ there exists an integer $n_0 = n_0(\eta)$ such that the following holds. Suppose G is a digraph on $n \geqslant n_0$ vertices such that for all $k < \frac{n}{2}$*

(i) *$d_k^+ \geqslant k + \eta n$ or $d_{n-k-\eta n}^- \geqslant n - k$;*
(ii) *$d_k^- \geqslant k + \eta n$ or $d_{n-k-\eta n}^+ \geqslant n - k$.*

Then G contains a Hamilton cycle.

It is natural to ask whether the Hamilton cycles guaranteed in Theorems 2 and 4 can be found efficiently. The main tools used to prove the above results were a version of Szemerédi's Regularity Lemma for digraphs [3] and the Blow-up Lemma [17]. Although both of them have algorithmic versions, (see [1] for the undirected version of the Regularity Lemma and [18] for the Blow-up Lemma) the authors needed to use a version of the Blow-up Lemma due to Csaba [7] which is not yet known to be algorithmic. Using a different approach, in this paper we give algorithmic versions of Theorems 2 and 4. In particular we avoid the use of Csaba's version of the Blow-up Lemma. More generally, our main result will work for all digraphs which have certain expansion properties. To state our result we first need some definitions.

Given $0 < \nu \leqslant \tau \leqslant \frac{1}{2}$, we call a digraph G a (ν, τ)-*outexpander* if for every $S \subseteq V(G)$ with $\tau|G| \leqslant |S| \leqslant (1-\tau)|G|$ we have $|N^+(S)| \geqslant |S| + \nu|G|$. Here, $N^+(S)$ denotes the set of all outneighbours of vertices of S. Although all digraphs we consider in this paper are outexpanders, this notion of expansion is not strong enough in order to be inherited by the reduced graph after we apply the Regularity Lemma. (Consider for example two disjoint cliques of equal size, joined by a matching.) For this reason, we will instead use the notion of robust outexpansion (introduced in [22] for similar reasons). Given a digraph G and $S \subseteq V(G)$, the ν-*robust out-neighbourhood* of S is the set

$$RN^+_{\nu,G}(S) = \{x \in V(G) : |N^-(x) \cap S| \geqslant \nu|G|\}.$$

We will usually drop the subscript G if it is clear to which digraph we are referring to. We call G a *robust* (ν, τ)-*outexpander* if $|RN^+_\nu(S)| \geqslant |S| + \nu|G|$ for every $S \subseteq V(G)$ with $\tau|G| \leqslant |S| \leqslant (1-\tau)|G|$. Thus a robust (ν, τ)-outexpander is also a (ν, τ)-outexpander.

We can now state our main theorem which implies algorithmic versions of Theorems 2 and 4. Here, and later on, we write $0 < a_k \ll \ldots \ll a_1 \leqslant 1$ to mean that there are increasing functions $f_2, \ldots, f_k$ such that, given $0 < a_1 \leqslant 1$, whenever we choose positive reals $a_2 \leqslant f_2(a_1), \ldots, a_k \leqslant f_k(a_{k-1})$, all calculations needed in the proofs of our statements are valid.

Theorem 5. *Let n_0 be an integer and let ν, τ, β be constants such that $0 < 1/n_0 \ll \nu \leqslant \tau \ll \beta \ll 1$. Let G be a digraph on $n \geqslant n_0$ vertices with $\delta^0(G) \geqslant \beta n$ and suppose G is a robust (ν, τ)-outexpander. Then G contains a Hamilton cycle. Moreover, there is an NC^5 algorithm for finding such a Hamilton cycle. In particular, there is a sequential polynomial time algorithm for finding such a Hamilton cycle.*

A non-algorithmic version of Theorem 5 was already proved in [22]. To see that the digraphs considered in Theorem 4 are robust outexpanders, we refer the reader to Lemma 11 of [22]. The fact that the graphs in Theorem 2(i) are robust outexpanders is proved in Lemma 12.1 of [21]. Lemma 6.2 of [15] shows that the oriented graphs considered in Theorem 2(ii) are outexpanders. A similar proof shows that they are in fact robust outexpanders.

Theorem 5 is also used as a tool in a paper by Kühn and Osthus [21], which shows that if a digraph G is a robust outexpander whose minimum semi-degree is linear in n, then G has a Hamilton decomposition. More precisely, the fact that Theorem 5 is algorithmic is used in [21] to provide an algorithm which finds the above Hamilton decomposition in polynomial time. This in turn is used to solve a problem on TSP tour domination which was posed by Glover and Punnen [11] as well as Alon, Gutin and Krivelevich [2] (see [21] for more details).

Our parallel algorithmic version of Theorem 2 is best possible not only in the sense that there are oriented graphs G with $\delta^0(G) = \lceil (3|G| - 4)/8 \rceil - 1$ which are not Hamiltonian, (see [14] for examples) but also in the following sense. Given an oriented graph G on n vertices with $\delta^0(G) \geqslant \eta n$ where $0 < \eta < 3/8$, it is NP-complete to decide whether G contains a Hamilton cycle. To see this, consider the graph G constructed as follows (see Figure 1). G has $(4+\alpha)n+1$ vertices partitioned into 5 parts A, B, C, D, H of sizes $|A| = |B| = |C| = n, |D| = n+1$ and $|H| = \alpha n$, where α is chosen so that $0 < \alpha < \frac{3-8\eta}{2\eta}$. Each of A and C span tournaments which are as regular as possible, B and D induce empty graphs, H is an arbitrary oriented graph and we add all possible edges from A to B and H, from B and H to C, from C to D and from D to A as well as bipartite tournaments between B and D and between D and H which are as (semi-)regular as possible (i.e. orientations of complete bipartite graphs such that, the in-degree and out-degree of each vertex differ by at most one.) It is easy to check that $\delta^0(G) \geqslant \frac{3n}{2} - 1 \geqslant \eta|G|$ (provided n is large enough). It is also easy to check that G contains a Hamilton cycle if

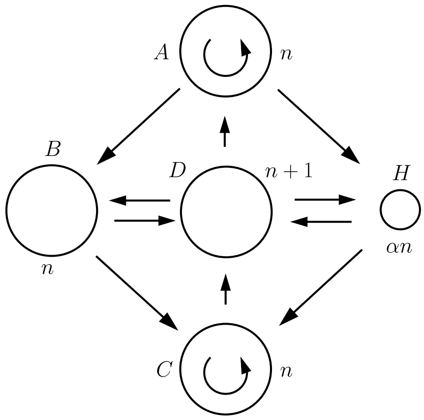

FIGURE 1.

and only if H contains a Hamilton path and it is well-known that to decide whether an arbitrary oriented graph H contains a Hamilton path is NP-complete.

Our paper is organized as follows. The next section contains some basic notation. In Section 3, we collect all the information we need about the Regularity Lemma and the Blow-up Lemma, and we state some simple facts about robust outexpanders. In Section 4, we give a brief overview of the proof. An important tool in our proof will be the notion of shifted walks. We explain how we obtain such walks in Section 5. Finally, in Section 6, we prove Theorem 5.

2. NOTATION

Given two vertices x and y of a digraph G, we write xy for the edge directed from x to y. The *order* $|G|$ of G is the number of its vertices. We write $N_G^+(x)$ and $N_G^-(x)$ for the out-neighbourhood and in-neighbourhood of x and $d_G^+(x)$ and $d_G^-(x)$ for its out-degree and in-degree. The *degree* of x is $d_G(x) = d_G^+(x) + d_G^-(x)$. The *minimum* and *maximum degree* of G are defined to be $\delta(G) = \min\{d(x) : x \in V(G)\}$ and $\Delta(G) = \max\{d(x) : x \in V(G)\}$ respectively. We usually drop the subscript G if this is unambiguous. Given a set A of vertices of G, we write $N_G^+(A)$ for the set of all out-neighbours of vertices of A, i.e. for the union of $N_G^+(x)$ over all $x \in A$. We define $N_G^-(A)$ analogously.

Given two vertices x and y on a directed cycle C we write xCy for the subpath of C from x to y. Similarly, given two vertices x and y on a directed path P such that x precedes y, we write xPy for the subpath of P from x to y. A *walk of length* ℓ in a digraph G is a sequence $v_0, v_1, \ldots, v_\ell$ of vertices of G such that $v_iv_{i+1} \in E(G)$ for all $0 \leqslant i \leqslant \ell - 1$. The walk is *closed* if $v_0 = v_\ell$. A *1-factor* of G is a collection of disjoint cycles which cover all vertices of G. Given a 1-factor F of G and a vertex x of G, we write x_F^+ and x_F^- for the successor and predecessor of x on the cycle in F containing x. We usually drop the subscript F if this is unambiguous.

Given disjoint vertex sets A and B in a graph G, we write $(A, B)_G$ for the induced bipartite subgraph of G with vertex classes A and B. We write $E_G(A, B)$ for the set of all edges ab with $a \in A$ and $b \in B$ and put $e_G(A, B) = |E_G(A, B)|$. As usual, we drop the subscripts when this is unambiguous.

Given a digraph G and a positive integer r, the *blow-up* of G by a factor of r is the digraph $G' = G \times E_r$ obtained from G by replacing every vertex x of G by r vertices $x_1, \ldots, x_r$ and replacing every edge xy of G by the r^2 edges x_iy_j $(1 \leqslant i, j \leqslant r)$.

To avoid unnecessarily complicated calculations we will sometimes omit floor and ceiling signs and treat large numbers as if they were integers.

3. The Main Tools

In this section, we collect all the information we need about the Regularity Lemma and the Blow-up Lemma and state some simple facts about outexpanders and robust outexpanders. For surveys on applications of the Regularity Lemma and the Blow-up Lemma, we refer the reader to [19, 16, 20].

3.1. **The Regularity Lemma.** The *density* of an undirected bipartite graph $G = (A, B)$ with vertex classes A and B is defined to be $d_G(A, B) = \frac{e_G(A,B)}{|A||B|}$. We often write $d(A, B)$ if this is unambiguous. Given $\varepsilon > 0$, we say that G is ε-*regular* if for all subsets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geqslant \varepsilon|A|$ and $|Y| \geqslant \varepsilon|B|$ we have that $|d(X, Y) - d(A, B)| < \varepsilon$. Given $d \in [0, 1]$, we say that G is (ε, d)-*regular* if it is ε-regular of density at least d. We also say that G is (ε, d)-*super-regular* if it is ε-regular and furthermore $d_G(a) \geqslant d|B|$ for all $a \in A$ and $d_G(b) \geqslant d|A|$ for all $b \in B$. Given partitions $V_0, V_1, \dots, V_k$ and $U_0, U_1, \dots, U_\ell$ of the vertex set of some graph, we say that $V_0, V_1, \dots, V_k$ *refines* $U_0, U_1, \dots, U_\ell$ if for all V_i with $1 \leqslant i \leqslant k$, there is some U_j with $0 \leqslant j \leqslant \ell$ which contains V_i. Note that this is weaker than the usual notion of refinement of partitions since V_0 need not be contained in any U_j.

Given a digraph G, and disjoint subsets A, B of $V(G)$, we say that the pair (A, B) is ε-regular, if the corresponding undirected bipartite graph consisting of all those edges of G which are directed from A to B is ε-regular. (So the order of A and B matters here.) We use a similar convention for super-regularity. The Diregularity Lemma is a version of the Regularity Lemma for digraphs due to Alon and Shapira [3]. We will use the degree form of the Diregularity Lemma which can be easily derived from the standard version, in exactly the same manner as the undirected degree form. (See e.g. [20] for a sketch proof.) We will also use the Diregularity Lemma in its algorithmic form. The algorithmic version of the Regularity Lemma is due to Alon, Duke, Lefmann, Rödl and Yuster [1]. Although we are not aware of any appearance of the algorithmic version of the Diregularity Lemma in print, it can be proved in much the same way as in [3], using instead the algorithmic ideas developed in [1]. For completeness, we include a sketch.

Lemma 6 (Diregularity Lemma; Algorithmic degree form)**.** *For every $\varepsilon \in (0, 1)$ and all positive integers M', M'', there are positive integers M and n_0 such that if G is a digraph on $n \geqslant n_0$ vertices, $d \in [0, 1]$ is any real number and $U_0, U_1, \dots, U_{M''}$ is a partition of the vertices of G, then there is an NC^1 algorithm that finds a partition of the vertices of G into $k + 1$ clusters $V_0, V_1, \dots, V_k$ and a spanning subdigraph G' of G with the following properties:*

- *$M' \leqslant k \leqslant M$;*
- *$V_0, V_1, \dots, V_k$ refines the partition $U_0, U_1, \dots, U_{M''}$;*
- *$|V_0| \leqslant \varepsilon n$, $|V_1| = \dots = |V_k| =: m$ and $G'[V_i]$ is empty for all $0 \leqslant i \leqslant k$;*
- *$d^+_{G'}(x) > d^+_G(x) - (d + \varepsilon)n$ and $d^-_{G'}(x) > d^-_G(x) - (d + \varepsilon)n$ for all $x \in V(G)$;*
- *all pairs $(V_i, V_j)_{G'}$ with $1 \leqslant i, j \leqslant k$ are ε-regular with density either 0 or at least d;*
- *all but at most εk^2 pairs $1 \leqslant i, j \leqslant k$ satisfy either $(V_i, V_j)_G = (V_i, V_j)_{G'}$ or $d_G(V_i, V_j) < d$.*

We call $V_1, \dots, V_k$ the *clusters* of the partition, V_0 the *exceptional set* and the vertices of G in V_0 the *exceptional vertices*. The fifth condition of the lemma says that all pairs of clusters are ε-regular in both directions (but possibly with different densities).

Sketch proof of Lemma 6. To prove an algorithmic version of the standard form of the Diregularity Lemma we follow the proof of Lemma 3.1 in [3]. To refine a partition $P = (V_1, \dots, V_k)$, instead of applying Lemma 3.4 of [3] which merely asserts the existence a refinement with some given properties we proceed as follows. Corollary 3.3 of [1] gives an NC^1 algorithm which either certifies that P is ε-regular (meaning that it produces a list of at least $\binom{k}{2} - \varepsilon k^2$ pairs which are ε-regular), or certifies that at least $\frac{\varepsilon^4}{16}$ pairs are not $\frac{\varepsilon^4}{16}$-regular (meaning that it returns subsets of the vertex classes of the pair which verify the non-regularity of the pair). Given these certificates, Lemma 3.4 of [1] gives an NC^1 algorithm which produces a refinement P' of P with the same properties as the refinement whose existence is guaranteed by Lemma 3.4 of [3] (but with slightly worse constants). Note that each time we apply Lemma 3.4 of [1] we apply it to one of the undirected graphs $\overrightarrow{G}(P)$, $\overleftarrow{G}(P)$ or $\overline{G}(P)$ which have the same vertex set as G and in which there is an edge between $x \in V_i$ and $y \in V_j$ with $i < j$, if and only if xy is an edge of G, yx is an edge of G, both xy and yx are edges of G respectively. Given the partition P, these undirected graphs can be constructed in NC^1. The proof of Lemma 3.1 of [3] shows that we only need to repeat this a constant number of times before Corollary 3.3 of [1] proves that we have arrived at an ε-regular partition. Although Lemma 3.1 of [3] does not mention the refinement property stated in Lemma 6 it is obvious that the same proof works for this property as well. It remains to show how to obtain the degree form of the Diregularity Lemma. This is obtained in a similar way from the standard version as in the undirected case: one applies the Diregularity lemma with a parameter $\varepsilon' \ll \varepsilon$ and then deletes a small proportion of the edges (in particular all edges between pairs which are not ε'-regular or have density less than $d+\varepsilon'$) and moves a small proportion of the vertices into V_0. (See [20] for a sketch of this for the undirected case.) One important difference is that in our case we do not know whether each pair is ε'-regular or not. However, for most ε'-regular pairs, we do have certificates confirming the ε'-regularity of the pair. So, instead of removing all edges between non ε'-regular pairs, we remove all edges between all pairs which are not known to be ε'-regular. The calculations remain unchanged. Finally, we just need to check that whenever we delete edges, or we remove vertices from the clusters into the exceptional cluster, we only need knowledge of the degrees of the vertices in the various clusters and there is an NC^1 algorithm for finding these degrees. □

The *reduced digraph* R of G' with parameters ε, d, M' (with respect to the above partition) is the digraph whose vertices are the clusters $V_1, \dots, V_k$ and in which V_iV_j is an edge precisely when $(V_i, V_j)_{G'}$ has density at least d (and thus is also ε-regular).

In various stages of our proof of Theorem 5, we will want to make some pairs of clusters super-regular, while retaining the regularity of all other pairs. This can be achieved by the following folklore lemma.

Lemma 7. *Let $\varepsilon \ll d, 1/\Delta$ and let R be a reduced digraph of G as given by Lemma 6. Let H be a subdigraph of R of maximum degree Δ. Then, we can move exactly $\Delta\varepsilon m$ vertices from each cluster V_i into V_0 such that each pair of clusters corresponding to an edge of H becomes $(2\varepsilon, \frac{d}{2})$-super-regular, while each pair of clusters corresponding to an edge of R becomes 2ε-regular with density at least $d - \varepsilon$. Moreover, there is an NC^1 algorithm for finding the set of vertices to be removed.*

Proof. For each cluster V of the partition, let

$$A(V) = \left\{ x \in V : \begin{array}{l} |N^+(x) \cap W| < (d-\varepsilon)m \text{ for some out-neighbour } W \text{ of } V \text{ in } H \\ \text{or } |N^-(x) \cap W| < (d-\varepsilon)m \text{ for some in-neighbour } W \text{ of } V \text{ in } H \end{array} \right\}$$

The definition of regularity implies that $|A(V)| \leqslant \Delta\varepsilon m$. Remove from each cluster V a set of size exactly $\Delta\varepsilon m$ containing $A(V)$. Since $\Delta\varepsilon \leqslant \frac{1}{2}$, it follows easily that all pairs corresponding to edges of R are 2ε-regular of density at least $d-\varepsilon$. Moreover, the minimum degree of each pair corresponding to an edge of H is at least $(d - (\Delta+1)\varepsilon)m \geqslant \frac{d}{2}m$, as required. Finally, for each cluster V and each vertex $x \in V$, to check whether $x \in A(V)$ we only need to compute the out-degrees and in-degrees of x in all the other clusters W so the parallelization claim follows. □

3.2. **A parallel algorithm for finding maximal matchings and systems of paths.** At several steps of our algorithm, we will need to produce matchings in certain bipartite graphs. It will turn out that if we only needed to find a sequential polynomial time algorithm, then we could find these matchings greedily. To find them in parallel, we will use the following result of Lev [23].

Theorem 8. *There exists an NC^4 algorithm for finding a maximal matching (i.e. a matching which cannot be extended) in a bipartite graph.*

We will also use the following result. Note that, using the definition of super-regularity, it is easy to greedily find the paths guaranteed by this result. The point is that one can find those paths efficiently in parallel.

Lemma 9. *Suppose k, m are integers and $\varepsilon_2, \varepsilon, d$ are real numbers such that $0 < \frac{1}{m} \ll \varepsilon_2 \ll \varepsilon, 1/k \ll d \leqslant 1$. Let R be a graph on $[k]$, let $V_1, \dots, V_k$ be pairwise disjoint sets of size m and let G be a graph with vertex set $V = V_1 \cup \dots \cup V_k$ obtained from R by replacing every vertex i of R with the set V_i $(1 \leqslant i \leqslant k)$ and replacing every edge ij of R by an (ε, d)-super-regular pair between V_i and V_j. Let $s \leqslant \varepsilon_2 m$ be a positive integer and for each $1 \leqslant i \leqslant s$, let $W_i = i_1 i_2 \dots i_{\ell(i)}$ be a walk in R with $4 \leqslant \ell(i) \leqslant k^3$. Suppose also that any closed subwalk of any W_i has length at least 4. Let $x_1, y_1, \dots, x_s, y_s$ be distinct vertices of V such that $x_i \in V_{i_1}$ and $y_i \in V_{i_{\ell(i)}}$ for each $1 \leqslant i \leqslant s$. Then, there is an NC^4 algorithm (wrt m) which finds s disjoint paths $P_1, \dots, P_s$ in G such that each P_i joins x_i to y_i, has the same length as W_i and such that whenever ab is an edge of W_i, the corresponding edge of P_i joins the sets V_a and V_b.*

Proof. We begin by finding the first edge of all paths P_i for which $\ell(i) \geqslant 5$. To find these edges consider the bipartite graph with vertex classes $A = \{x_i : \ell(i) \geqslant 5\}$ and $B = V \setminus \{x_i, y_i : 1 \leqslant i \leqslant s\}$. In this graph we join $x_i \in A$ to $v \in B$ if and only if $v \in V_{i_2}$ and x_i is adjacent to v in G. By super-regularity of the pair (V_{i_1}, V_{i_2}), each $x_i \in A$ has at least $(d/2 - 2\varepsilon_2)m$ neighbours in this bipartite graph. Since $\frac{1}{m} \ll \varepsilon_2 \ll d$, it follows that any maximal matching in this bipartite graph covers every vertex of A. Thus Theorem 8 implies that we can find the required edges. Repeating this at most $\frac{1}{k^3}$ times, we may find the first $\ell(i) - 4$ edges of each path P_i. Indeed, at each application of Theorem 8 we know that at most $sk^3 \leqslant \varepsilon_2 k^3 m \ll dm$ vertices have been used from each V_i and so a similar argument as above shows that the paths can be extended. To avoid introducing more notation, from now on we will assume that each walk W_i has length exactly 3 (and so is a path) and keep in mind the extra restriction that each V_i contains a subset U_i of size at most $\varepsilon_2 k^3 m$ of vertices which are not allowed to be used when creating the paths P_i. For each $1 \leqslant i \leqslant k$, let $V_i' = V_i \setminus U_i$. We now want to find distinct $w_1, z_1, \dots, w_s, z_s \in V' = V_1' \cup \dots \cup V_k'$ such that for each i, $x_i w_i, w_i z_i, z_i y_i$ are edges of G, $w_i \in V_{i_2}$ and $z_i \in V_{i_3}$. Then, the $P_i := x_i w_i z_i y_i$ will be the required paths in G. To find these w_i's and z_i's we proceed as follows. For each i, consider $N(x_i) \cap V_{i_2}'$ and $N(y_i) \cap V_{i_3}'$. By super-regularity of the pairs (V_{i_1}, V_{i_2}) and (V_{i_3}, V_{i_4}) we have that $|N(x_i) \cap V_{i_2}|, |N(y_i) \cap V_{i_3}| \geqslant dm/2$ and so $|N(x_i) \cap V_{i_2}'|, |N(y_i) \cap V_{i_3}'| \geqslant (d/2 - \varepsilon_2 k^3)m$. Since $\varepsilon_2 k^3 \ll d$, the regularity of the pair (V_{i_2}, V_{i_3}) implies that for each $1 \leqslant i \leqslant s$, we can find subsets $W_i \subseteq N(x_i) \cap V_{i_2}'$ and $Z_i \subseteq N(y_i) \cap V_{i_3}'$ such that each $w_i \in W_i$ has at least $\frac{d^2 m}{3}$ neighbours in Z_i and each $z_i \in Z_i$ has at least $\frac{d^2 m}{3}$ neighbours in W_i. In particular, $|W_i|, |Z_i| \geqslant \frac{d^2 m}{3}$. We claim that we can pick distinct $w_1, \dots, w_s$ such that $w_i \in W_i$ for each $1 \leqslant i \leqslant s$. Since $\varepsilon_2 \ll d$ and so $s \ll d^2 m$, this follows by applying Theorem 8 in the natural auxiliary bipartite graph. Finally, we claim that we can pick distinct $z_1, \dots, z_s$ such that $z_i \in Z_i \cap N(w_i)$ for each $1 \leqslant i \leqslant s$. This follows again by applying Theorem 8 in the natural auxiliary bipartite graph. This completes the proof of the lemma. $\square$

3.3. **The Blow-up Lemma.** The Blow-up Lemma implies that dense super-regular pairs behave like complete bipartite graphs with respect to containing bounded degree graphs as subgraphs. In our proof of Theorem 5, we will need the algorithmic version of the Blow-up Lemma [18].

Lemma 10 (Blow-up Lemma; Algorithmic form)**.** *For any graph R of order k and any positive parameters d, Δ, there exists an $\varepsilon_0 = \varepsilon_0(d, \Delta, k) > 0$ such that whenever $0 < \varepsilon \leqslant \varepsilon_0$, the following holds. Let n be a positive integer and let us replace the vertices of R with pairwise disjoint sets $V_1, \dots, V_k$ of size n (blowing-up). We construct two graphs on the same vertex set $V_1 \cup \cdots \cup V_k$. The graph $R(n)$ is obtained by replacing all edges of R with copies of the complete bipartite graph $K_{n,n}$ and a sparser graph G is obtained by replacing the edges of R with some (ε, d)-super-regular pairs. If a graph H with maximum degree $\Delta(H) \leqslant \Delta$ is embeddable into $R(n)$ then it is already embeddable into G. Moreover, there is an NC^5 algorithm for finding such a copy of H in G.*

In fact, we will only use the following consequence of the Blow-up Lemma.

Lemma 11. *For every real number $d \in [0, 1]$, there exists an $\varepsilon'_0 = \varepsilon'_0(d) > 0$ such that whenever $0 < \varepsilon \leqslant \varepsilon'_0$, the following holds. Let k, n be positive integers with $k \geqslant 4$, $V_1, \dots, V_k$ be pairwise disjoint sets of size n and suppose G is a digraph on $V_1 \cup \cdots \cup V_k$ such that each $(V_i, V_{i+1})_G$ is (ε, d)-super-regular. (Here, $V_{k+1} := V_1$.) Take any $x \in V_1$ and any $y \in V_k$. Then there is an NC^5 algorithm which finds a Hamilton path P in G, starting with x and ending with y. Moreover, for every vertex $v \in V_i$, the successor of v on P lies in V_{i+1}.*

Proof. We claim that we may take $\varepsilon'_0(d) = \min\{\frac{1}{2}\varepsilon_0(d/2, 2, \ell) : \ell \leqslant 6\}$. We show that this ε'_0 works as follows. By deleting edges if necessary we may assume that for every edge vw of G there is an i such that $v \in V_i$ and $w \in V_{i+1}$. Consider $(V_k, V_1)_G - \{x, y\}$. By the Blow-up Lemma (applied to the corresponding undirected graph), there is an NC^5 algorithm giving a perfect matching from $V_k \setminus y$ to $V_1 \setminus x$. Let us write $V_i = \{x_{i1}, \dots, x_{in}\}$ for each $1 \leqslant i \leqslant k$. We may assume $x_{11} = x$, $x_{kn} = y$ and the edges of the matching are all edges of the form $x_{ki}x_{1(i+1)}$ for $1 \leqslant i \leqslant n-1$. Hence, it is enough to give an NC^5 algorithm which produces n vertex disjoint paths of length $k-1$, connecting x_{1i} with x_{ki} for each $1 \leqslant i \leqslant n$. By fixing some intermediate vertices we can partition the edge set of the path of length $k-1$ corresponding to the graph G into paths of length at least 3 and at most 5. By considering these paths instead, we may assume that $4 \leqslant k \leqslant 6$. We now define a new undirected graph G' by identifying V_1 with V_k via the identification of x_{1i} with x_{ki} and by ignoring the orientation of the edges. Applying the Blow-up Lemma to G' we obtain n disjoint cycles of length k in G'. The result now follows since these cycles in G' correspond to the required paths of length $k-1$ in G. □

3.4. **Properties of Outexpanders.** In this subsection, we gather some simple properties about outexpanders that will be needed in the proof of Theorem 5. We assume throughout that $0 < \nu \leqslant \tau \leqslant \frac{1}{2}$.

Lemma 12. *Let G be a digraph of order n with $\delta^0(G) \geqslant \tau n$ and suppose G is a (ν, τ)-outexpander. Then G contains a 1-factor.*

Proof. We claim that for every $S \subseteq V(G)$, we have $|N^+(S)| \geqslant |S|$. Indeed, if $0 \neq |S| < \tau n$, then $|N^+(S)| \geqslant \delta^+(G) \geqslant \tau n > |S|$, if $\tau n \leqslant |S| \leqslant (1-\tau)n$, then $|N^+(S)| \geqslant |S| + \nu n$ by the outexpansion properties of G, and finally, if $|S| > (1-\tau)n$, then $|S| + \delta^-(G) > n$ and so $N^+(S) = V(G)$, hence $|N^+(S)| \geqslant |S|$. The result now follows by applying Hall's theorem to the bipartite graph H with vertex classes A and B, where A and B are both copies of the vertex set of G and there is an edge joining $a \in A$ to $b \in B$ if and only if there is a directed edge from a to b in G. Indeed, by Hall's theorem H has a perfect matching and by the definition of H this corresponds to a 1-factor of G. □

Lemma 13. *Let G be a (ν,τ)-outexpander of order n and let G' be a graph obtained from G by adding at most $\frac{\nu}{2}n$ isolated vertices. Then G' is a $(\frac{\nu}{4},2\tau)$-outexpander.*

Proof. Take $S'\subseteq V(G')$ with $2\tau|G'|\leqslant|S'|\leqslant(1-2\tau)|G'|$ and let $S=S'\cap V(G)$. Then $\tau n\leqslant|S|\leqslant(1-\tau)n$, hence

$$|N^+_{G'}(S')|\geqslant|N^+_G(S)|\geqslant|S|+\nu n\geqslant|S'|+\frac{\nu}{2}n\geqslant|S'|+\frac{\nu}{4}|G'|,$$

as required. □

Lemma 14. *Let G be a (ν,τ)-outexpander and let G' be a blow-up of G. Then G' is also a (ν,τ)-outexpander.*

Proof. Let us denote the order of G by n and suppose G' is the blow-up of G by a factor of r. Take $S'\subseteq V(G')$ with $\tau rn\leqslant|S'|\leqslant(1-\tau)rn$ and consider

$$S=\{x\in G: S' \text{ contains a copy of } x\}.$$

Since G is a (ν,τ)-outexpander, it follows that:

(i) Either $|N^+(S)|\geqslant|S|+\nu n$;
(ii) or $|S|\geqslant(1-\tau)n$, in which case (considering a subset of S of size $(1-\tau)n$) we have $|N^+(S)|\geqslant(1-\tau+\nu)n$.

Note that if a vertex x of G belongs to $N^+(S)$, then any copy x' of x in G' belongs to $N^+(S')$. It follows that $|N^+(S')|\geqslant r|N^+(S)|$. Thus, in case (i) we have

$$|N^+(S')|\geqslant r|N^+(S)|\geqslant r|S|+r\nu n\geqslant|S'|+\nu rn,$$

while in case (ii) we have

$$|N^+(S')|\geqslant r|N^+(S)|\geqslant(1-\tau)rn+\nu rn\geqslant|S'|+\nu rn,$$

as required. □

We will also use the following lemma from [22, Lemma 11]. This is the only place where, for our proof to work, we do need our digraphs to be robust outexpanders rather than just outexpanders.

Lemma 15. *Let M',n_0 be integers and let $\varepsilon,d,\nu,\tau,\beta$ be constants such that $0<\frac{1}{n_0}\ll\varepsilon\ll d\ll\nu\leqslant\tau,\beta<1/2$ and such that $0<M'\ll n_0$. Let G be a digraph on $n\geqslant n_0$ vertices with $\delta^0(G)\geqslant\beta n$ and such that G is a robust (ν,τ)-outexpander. Let R be the reduced digraph of G with parameters ε,d and M'. Then $\delta^0(R)\geqslant\frac{\beta}{2}|R|$ and R is a robust $(\frac{\nu}{2},2\tau)$-outexpander.*

4. Overview of the Proof of Theorem 5

We now give a rough overview of the proof of Theorem 5, which is worth keeping in mind when following the details of the proof. By applying the Diregularity Lemma to G with parameters ε_1,d_1 and $M_1'=\frac{1}{\varepsilon_1}$, we obtain a reduced graph R_1 of order k_1 and an exceptional set V_0^1. By Lemma 15 R_1 is an outexpander and so by Lemma 12 it contains a 1-factor F_1. By Lemma 7, we may assume that the edges of F_1 correspond to super-regular pairs. Let R_1^* be the graph obtained from R_1 by adding the set V_0^1 of exceptional vertices and for each $x\in V_0^1$ and each $V\in R_1$ adding the edge xV if x has many out-neighbours in V and the edge Vx if x has many in-neighbours in V. We would like to find a closed walk W in R_1^* such that

(a) For each cycle C_1 of F_1, W visits every vertex of C_1 the same number of times;
(b) W visits every cluster of R_1 at least once but not too many times;
(c) W visits every vertex of V_0^1 exactly once;
(d) any two vertices of V_0^1 are at distance at least 3 along W.

Having obtained W, we would then find a corresponding cycle W' such that whenever W visits a vertex of V_0^1, W' visits the same vertex, and whenever W visits a cluster V_i of R_1, then W' visits a vertex $x \in V_i$. We would then be able to use Lemma 11 to transform W' to a Hamilton cycle of G. Property (a) is required because we want to ensure that whenever we apply Lemma 11, all clusters have the same sizes. Property (b) is required to ensure that whenever we apply Lemma 11, all pairs of clusters we are interested in are indeed super-regular. Property (c) is required so that the Hamilton cycle does indeed cover all vertices of V_0^1 (exactly once) and finally property (d) is required in order to construct W' with the properties described above. Unfortunately, since V_0^1 might have size $\varepsilon_1 n$, this simple approach can only guarantee that W visits each cluster of R_1 at most $O(\frac{\varepsilon_1}{\nu} n)$ times. This however is far too large to allow the use of Lemma 11 (as it is larger than the number of vertices in each cluster). So, instead of considering R_1^*, we proceed as follows.

We refine our partition by applying the Diregularity Lemma with new parameters $\varepsilon_2, d_2 \ll \varepsilon_1$ and $M_2' = \frac{1}{\varepsilon_2}$ to obtain a new reduced graph R_2 whose clusters are subclusters of the V_i and a new exceptional set V_0^2. Fix $0 < \theta < 1$. Using the fact that the blow-up of R_1 is an outexpander, we can find a union F_2 of disjoint cycles covering all subclusters of V_0^1 as well as a θ-proportion of the subclusters of each cluster V_i of R_1 (provided θ is large enough.) As before, we may assume that the edges of F_2 correspond to super-regular pairs. For each cycle C_2 of F_2, Lemma 11 gives a Hamilton path in the subgraph of G corresponding to C_2. Now let R^* be the graph obtained from R_1 by adding the set V_0^2 of exceptional vertices and a vertex for each cycle C_2 of F_2. For each $x \in V_0^2$ and each $V \in R_1$ add the edge xV if x has many out-neighbours in V and the edge Vx if x has many in-neighbours in V. Given a cycle C_2 of F_2, suppose that the application of Lemma 11 yields a Hamilton path in the corresponding subgraph of G, starting at x and ending at y, where x belongs to the cluster V_k. Then add an edge in R^* from C_2 to V_k and an edge from V_k^- (the predecessor of V_k in F_1) to C_2. Provided θ is not too large, we can find a closed walk W in R^* such that

(a) For each cycle C_1 of F_1, W visits every vertex of C_1 the same number of times;
(b) W visits every cluster of R_1 at least once but not too many times;
(c) W visits every vertex of V_0^2 exactly once;
(d) W visits every cycle C_2 of F_2 exactly once;
(e) any two vertices of V_0^2 are at distance at least 3 along W.

With this approach, we can now guarantee that the number of times that W visits a cluster V_i of R_1 is $\ll |V_i|$, and this is small enough to allow the use of Lemma 11 in order to transform W into a Hamilton cycle of G.

5. Shifted Walks

To achieve property (a) above, we will build up W from certain special walks, each of them satisfying property (a). Given vertices $a, b \in R_1$, a *shifted walk* from a to b is a walk $W(a, b)$ of the form

$$W(a, b) = x_1 C_1 x_1^- x_2 C_2 x_2^- \dots x_t C_t x_t^- x_{t+1},$$

where $x_1 = a$, $x_{t+1} = b$, $C_1, \dots, C_t$ are (not necessarily distinct) cycles of F_1, and for each $1 \leqslant i \leqslant t$, x_i^- is the predecessor of x_i on C_i. We call $C_1, \dots, C_t$ the cycles which are *traversed* by $W(a, b)$. So even if the cycles $C_1, \dots, C_t$ are not distinct, we say that W traverses t cycles. Note that for every cycle C of F_1, the walk $W(a, b) - b$ visits the vertices of C an equal number of times.

Our next lemma will guarantee that between any two vertices a, b of R_1 there will be a shifted walk $W(a, b)$ which does not traverse too many cycles.

Lemma 16. *Let R be a (ν,τ)-outexpander with $\delta^0(R) \geqslant 3\tau|R|$ and let F be a 1-factor in R. Then, for any $a, b \in V(R)$, there is a shifted walk $W(a,b)$ from a to b traversing at most $\frac{1}{\nu}$ cycles.*

Proof. Let $S_1 = N^+_R(a^-)$ and for $i \geqslant 1$, let $S_{i+1} = N^+_R(N^-_F(S_i))$. Note that for every $i \geqslant 1$, S_i is the set of vertices x of R for which there exists a shifted walk from a to x traversing at most i cycles. In particular, $S_i \subseteq S_{i+1}$. Note also that $|S_1| \geqslant \delta^+(R) \geqslant 3\tau|R|$. Since R is a (ν,τ)-outexpander, it follows that either $|S_1| \leqslant (1-\tau)|R|$ in which case we have $|S_2| \geqslant |S_1| + \nu|R| \geqslant 4\nu|R|$, or $|S_1| \geqslant (1-\tau)|R|$ in which case we have $|S_2| \geqslant (1-\tau+\nu)|R|$. In both cases, it follows that $|S_2| \geqslant \min\{4\nu|R|, (1-\tau+\nu)|R|\}$ and inductively, $|S_i| \geqslant \min\{(i+2)\nu|R|, (1-\tau+\nu)|R|\}$ for every $i \geqslant 1$. In particular, $|S_{\lfloor 1/\nu \rfloor -1}| \geqslant (1-\tau+\nu)|R|$. But then $|S_{\lfloor 1/\nu \rfloor -1}| + \delta^-(R) > n$, and so $S_{\lfloor 1/\nu \rfloor} = V(R)$ as required. □

6. Proof of Theorem 5

We begin by defining additional constants such that

$$\frac{1}{n_0} \ll \varepsilon_2 \ll d_2 \ll \varepsilon_1 \ll \theta \ll d_1 \ll \nu \leqslant \tau \ll \beta \leqslant 1.$$

Recall that this means that we can choose the constants from right to left as explained before the statement of Theorem 5.

Apply the Diregularity Lemma with parameters ε_1, d_1 and $M'_1 = \frac{1}{\varepsilon_1}$ to obtain an exceptional set V^1_0, a spanning subdigraph G'_1 of G and a reduced graph R'_1. By Lemma 15, R'_1 is a $(\frac{\nu}{2}, 2\tau)$-outexpander with $\delta^0(R'_1) \geqslant \frac{\beta}{2}|R'_1|$. So by Lemma 12, R'_1 contains a 1-factor F'_1.

For technical reasons, it will be convenient to be able to assume that each cycle of F'_1 has length at least 4. (This is because Lemma 9 fails if one of the walks W_i has length less than 3.) To achieve this, we arbitrarily partition each cluster of G into 2 parts of equal size. (If the sizes of the clusters are odd then we move one vertex from each cluster to V^1_0.) Consider the graph R''_1 whose vertices correspond to the parts and where two vertices are joined by an edge if the corresponding bipartite subgraph of G'_1 is $(3\varepsilon_1, \frac{2d_1}{3})$-regular. Note that we may not be able to construct R''_1 in NC. This is because deciding whether a given pair is ε-regular is co-NP-complete (see [3]). For this reason, we will instead work with the subgraph $R_1 = R'_1 \times E_2$ of R''_1. We will denote the order of R_1 by k_1 and write $V_1, \dots, V_{k_1}$ for its clusters (which were the parts of the original clusters). Note that $\delta^0(R_1) \geqslant \frac{\beta}{2}k_1$. Note also that by Lemma 14, R_1 is a $(\frac{\nu}{2}, 2\tau)$-outexpander. The size of the exceptional set is now at most $\varepsilon_1 n + |R'_1| \leqslant 2\varepsilon_1 n$. Each cycle of length ℓ of F'_1 now becomes a copy of $C_\ell \times E_2$, which contains a cycle of length 2ℓ. This yields a 1-factor F_1 of R_1 so that all cycles of F_1 have length at least 4.

Our next step is to make the pairs of clusters corresponding to edges of F_1 $(6\varepsilon_1, \frac{d_1}{3})$-super-regular, rather than just regular. By Lemma 7 we can achieve this by moving exactly $6\varepsilon_1|V_i|$ vertices from each cluster V_i into V^1_0 and thus increasing the size of V^1_0 to at most $8\varepsilon_1 n$. We will still refer to the new clusters as $V_1, \dots, V_{k_1}$ and to the new exceptional set as V^1_0. We will denote the size of the V_i by m_1. Note that R_1 has not been altered in any way and all edges of R_1 correspond to $6\varepsilon_1$-regular pairs of density at least $\frac{d_1}{3}$.

As explained in the overview, we will now need to refine our partition. Before doing so, we define a new graph G_1 obtained from G'_1 by removing, for each $x \in V^1_0$, all edges from x into V_i (for $1 \leqslant i \leqslant k_1$) unless $|N^+_G(x) \cap V_i| \geqslant \frac{\beta}{4}m_1$, and all edges from V_i into x unless $|N^-_G(x) \cap V_i| \geqslant \frac{\beta}{4}m_1$. Since $|V^1_0| \leqslant 8\varepsilon_1 n$, it is immediate that for every $x \in V(G) \setminus V^1_0$, we have $d^+_{G_1}(x) > d^+_{G'_1}(x) - 8\varepsilon_1 n \geqslant d^+_G(x) - (d_1 + 9\varepsilon_1)n$ and similarly

$d^-_{G_1}(x) > d^-_G(x) - (d_1 + 9\varepsilon_1)n$. Moreover, for each $x \in V_0^1$ we have

$$d^+_{G_1}(x) \geqslant d^+_{G'_1}(x) - \beta m_1 k_1/4 \geqslant d^+_G(x) - (d_1 + \varepsilon_1 + \beta/4)n > \beta n/4.$$

Similarly $d^-_{G_1}(x) > \frac{\beta}{4}n$. We now apply the Diregularity Lemma to G_1 with parameters ε_2, d_2 and $M'_2 = \frac{1}{\varepsilon_2}$ to obtain a partition refining $V_0^1, V_1, \ldots, V_{k_1}$. Observe that since k_1 is bounded by a function of ε_1, we may assume that the constant d_2 is chosen in such a way that $d_2 \ll \frac{1}{k_1}$ holds. We denote the exceptional set by V_0^2, the spanning subdigraph by G'_2, the reduced graph by R'_2, its order by k_2 and the size of the clusters of R'_2 by m'_2. For each $1 \leqslant i \leqslant k_1$, we denote the clusters of R'_2 contained in V_i by V_{ij} and call them the *subclusters* of V_i. Since $(1 - 8\varepsilon_1)\frac{n}{k_1} \leqslant m_1 \leqslant \frac{n}{k_1}$ and $(1 - \varepsilon_2)\frac{n}{k_2} \leqslant m'_2 \leqslant \frac{n}{k_2}$ we have for all $i \geqslant 1$ that

$$(1 - 9\varepsilon_1)\frac{k_2}{k_1} \leqslant (m_1 - |V_0^2|)\frac{k_2}{n} \leqslant |\{V_{ij} : j \geqslant 1\}| \leqslant \frac{1}{(1 - \varepsilon_2)}\frac{k_2}{k_1}. \tag{1}$$

Note however that distinct V_i may have different number of subclusters. Finally, we denote the clusters of R'_2 contained in V_0^1 by V_{0j} and call them the subclusters of V_0^1.

Our next aim is to find a union F_2 of cycles in R'_2 covering all subclusters of V_0^1 and exactly $\theta\frac{k_2}{k_1}$ subclusters of every other V_i. Before doing that, it will be convenient to collect some results about the edge distribution in R'_2. The next lemma states that every subcluster of V_0^1 has significant degree in R'_2.

Lemma 17. *Every subcluster V_{0i} of V_0^1 satisfies $d^+_{R'_2}(V_{0i}), d^-_{R'_2}(V_{0i}) \geqslant \frac{\beta}{5}k_2$.*

Proof. Suppose this is not the case, say $d^+_{R'_2}(V_{0i}) < \frac{\beta}{5}k_2$ for some i and consider any $x \in V_{0i}$. Then

$$\frac{\beta}{4}n \leqslant d^+_{G_1}(x) \leqslant d^+_{R'_2}(V_{0i})m'_2 + |V_0^2| + (d_2 + \varepsilon_2)n < \left(\frac{\beta}{5} + d_2 + 2\varepsilon_2\right)n,$$

a contradiction. □

We now remove some edges from G'_2 to obtain a new digraph G_2. The reason for doing this, is to guarantee later that any two subclusters of V_0^1 are at distance at least 3 in the union F_2 of cycles. For each subcluster V_{ij} with $i \geqslant 1$, we either remove all edges from V_{ij} into all subclusters V_{0k} of V_0, or we remove all edges from all subclusters V_{0k} of V_0 into V_{ij}. We also let $R_2 \subseteq R'_2$ be the reduced digraph of G_2 with respect to the same partition. We can randomly remove the edges in such a way that the conclusion of the following lemma holds.

Lemma 18. *There is a subdigraph G_2 obtained from G'_2 as above, such that for every subcluster V_{0k} of V_0^1 we have $d^+_{R_2}(V_{0k}), d^-_{R_2}(V_{0k}) \geqslant \frac{\beta}{20}k_2$. Moreover, G_2 can be obtained from G'_2 in constant parallel time.*

Proof. For each V_{ij} with $i \geqslant 1$, either remove all edges from V_{ij} into all subclusters V_{0k} of V_0, or remove all edges from all subclusters V_{0k} of V_0 into V_{ij} choosing either option with probability 1/2, independently at random. For each subcluster V_{0k} of V_0^1 denote by X_k^+ the random variable $d^+_{R_2}(V_{0k})$ and by X_k^- the random variable $d^-_{R_2}(V_{0k})$. Lemma 17 implies that $\mathbb{E}X_k^+ \geqslant \frac{\beta}{10}k_2$ and so by Chernoff's inequality (see e.g. [4, Theorem A.1.4])

$$\mathbb{P}\left(X_k^+ \leqslant \frac{\beta}{20}k_2\right) \leqslant \mathbb{P}\left(X_k^+ \leqslant \frac{1}{2}\mathbb{E}X_k^+\right) \leqslant \exp\left\{-\frac{\beta}{40}k_2\right\}.$$

A similar inequality holds for X_k^- and so the probability that G_2 does not satisfy the required properties of the lemma is at most $2k_2 \exp\left\{-\frac{\beta}{40}k_2\right\}$. Since $k_2 \geqslant M'_2 \gg 1$, there is a positive probability that G_2 has the required properties. Finally, to see that G_2 can be obtained from G_2' in constant time, note that the size of the probability space used depends only on k_2 and not on n. $\square$

We proceed by showing that every subcluster of V_0^1 forms an edge of R_2 (and thus an (ε_2, d_2)-regular pair) with many subclusters of many clusters of R_1.

Lemma 19. *For every subcluster V_{0i} of V_0^1*

(i) *there are at least $\frac{\beta}{50}k_1$ clusters V_j such that (V_{0i}, V_{jk}) is an edge of R_2 for at least $\frac{\beta}{50}\frac{k_2}{k_1}$ subclusters V_{jk} of V_j;*

(ii) *there are at least $\frac{\beta}{50}k_1$ clusters V_j such that (V_{jk}, V_{0i}) is an edge of R_2 for at least $\frac{\beta}{50}\frac{k_2}{k_1}$ subclusters V_{jk} of V_j.*

Proof. If (i) is not true, then by (1) there is an i such that

$$d^+_{R_2}(V_{0i}) \leqslant \left(\frac{\beta}{50}k_1\right)\left(\frac{1}{1-\varepsilon_2}\frac{k_2}{k_1}\right) + k_1\left(\frac{\beta}{50}\frac{k_2}{k_1}\right) < \frac{\beta}{20}k_2,$$

contradicting Lemma 18. Part (ii) of the lemma is proved in a similar way. $\square$

The last result we need in order to produce the union F_2 of cycles is that if (V_i, V_j) is an edge of R_1 then in G_2 most subclusters of V_i form an edge of R_2 with many subclusters of V_j.

Lemma 20. *Let (V_i, V_j) be an edge of R_1. Let S_i and S_j be unions of s_i and s_j subclusters of V_i and V_j respectively, where $s_i, s_j \geqslant \sqrt{\varepsilon_1}\frac{k_2}{k_1}$. Call a subcluster V_{ik} of V_i* ***bad for*** *S_j, if there are at most $d_1^2 s_j$ subclusters $V_{j\ell}$ belonging to S_j such that $(V_{ik}, V_{j\ell})$ is an edge of R_2. Then S_i has at most $\sqrt[4]{\varepsilon_1}s_i$ subclusters which are bad for S_j.*

Proof. Suppose S_i has $b \geqslant \sqrt[4]{\varepsilon_1}s_i$ subclusters which are bad for S_j. Let B be the union of these bad subclusters and consider the bipartite graph $(B, S_j)_{G_1}$. Since $|B|, |S_j| \gg \varepsilon_1 m_1$ and $(V_i, V_j)_{G_1}$ is $(6\varepsilon, d_1/3)$-regular, we have $d_{G_1}(B, S_j) \geqslant \frac{d_1}{3} - 6\varepsilon_1 \geqslant \frac{d_1}{4}$. However, by our assumption, at least $(1 - d_1^2)bs_j$ pairs of subclusters $V_{ik}, V_{j\ell}$ belonging to B and S_j do not form an edge of R_2. The last property of Lemma 6 implies that at most $\varepsilon_2 k_2^2$ of these have density at least d_2 in G_1. Since $\varepsilon_2 k_2^2 \ll d_1^2 bs_j$ as $\varepsilon_2 \ll \frac{1}{k_1}$, it follows that at least $(1 - 2d_1^2)bs_j$ of the pairs $V_{ik}, V_{j\ell}$ belonging to B and S_j have density less than d_2 in G_1. But then $(B, S_j)_{G_1}$ must have density at most $2d_1^2 + d_2 < \frac{d_1}{4}$, a contradiction. $\square$

We can now find the promised union F_2 of cycles in R_2.

Lemma 21. *R_2 contains a union F_2 of cycles covering all subclusters of V_0^1 and exactly $\theta\frac{k_2}{k_1}$ subclusters of every V_i with $1 \leqslant i \leqslant k_1$. Furthermore, every cycle in F_2 has length at least 4 and contains two consecutive subclusters, say V_{ij} followed by $V_{k\ell}$, such that neither of them is a subcluster of V_0^1 and moreover, V_{ij} is not bad for V_k.*

Proof. We begin by finding a 1-factor F_2^A in an auxiliary graph A, and then use F_2^A to create F_2. We define A as follows: We blow up R_1 by a factor of $\theta\frac{k_2}{k_1}$ and add to this blow-up all subclusters of V_0^1. Moreover we add edges from V_{0i} to all copies of V_j in the blow-up if and only if (V_{0i}, V_{jk}) is an edge of R_2 for at least $\frac{\beta}{50}\frac{k_2}{k_1}$ subclusters V_{jk} of V_j and similarly we add edges from all copies of V_j to V_{0i} if and only if (V_{jk}, V_{0i}) is an edge of R_2 for at least $\frac{\beta}{50}\frac{k_2}{k_1}$ subclusters V_{jk} of V_j. By Lemma 14, the blow-up of R_1 is $(\frac{\nu}{2}, 2\tau)$-outexpander. Hence, by Lemma 13, A is a $(\frac{\nu}{8}, 4\tau)$-outexpander. This follows because we can assume $\varepsilon_1 \ll \nu\theta$. Moreover, Lemma 19 implies that $\delta^0(A) \geqslant \frac{\beta}{51}|A|$, and so, by Lemma 12, A contains a 1-factor F_2^A. We claim that we may assume that F_2^A contains no cycles of length 2. Indeed, if such a cycle appears, then by definition of R_2 it cannot contain a subcluster of V_0^1. So suppose that this cycle is A_iA_j where A_i is a copy of V_i and A_j is a copy of V_j. Then remove this cycle, find any other copy B_i of V_i on some other cycle, and replace the appearance of B_i by $A_iA_jB_i$. By the construction of A, we still have a union of cycles, with one fewer cycle of length 2. A similar argument also shows that we may assume that F_2^A contains no cycles of length 3. Moreover, the fact that every two vertices corresponding to subclusters of V_0^1 have distance at least 3 in R_2 implies that every cycle of F_2^A contains two consecutive vertices, say A_i and A_j, which correspond to clusters V_i and V_j with $i, j \geqslant 1$.

We now use F_2^A to induce the required union F_2 of cycles in R_2. To do this, we will find for each cycle $A_1A_2 \dots A_rA_1$ of F_2^A, a cycle $V_{i_1j_1}V_{i_2j_2} \dots V_{i_rj_r}V_{i_1j_1}$ of R_2 such that:

- If A_ℓ is a subcluster of V_0^1, then $V_{i_\ell j_\ell} = A_\ell$ (and so $i_\ell = 0$). If A_ℓ is a copy of some cluster V_i with $i \neq 0$, then $V_{i_\ell j_\ell}$ is a subcluster of V_i (and so $i_\ell = i$).
- If both i_ℓ and $i_{\ell+1}$ (addition done modulo r) are not equal to 0, then $V_{i_\ell j_\ell}$ is not bad for $V_{i_{\ell+1}}$.
- Every subcluster V_{ij} of R_2 is used in at most one such cycle.

Clearly, if we can do this, we obtain the required union F_2 of cycles.

Suppose first that A_1 and A_s are subclusters of V_0^1 but $A_2, \dots, A_{s-1}$ are not (possibly with $A_1 = A_s$, i.e. $r = s-1$). Note that we must have $s \geqslant 4$. For $\ell = 2, 3, \dots, s-3$, given $V_{i_{\ell-1},j_{\ell-1}}$ we choose V_{i_ℓ,j_ℓ} such that $(V_{i_{\ell-1},j_{\ell-1}}, V_{i_\ell,j_\ell})$ is an edge of R_2 and V_{i_ℓ,j_ℓ} is not bad for $V_{i_{\ell+1}}$. To see that this can be done note that if $\ell = 2$, then by definition of A there are at least $\frac{\beta}{50}\frac{k_2}{k_1} \geqslant \frac{d_1^2}{2}\frac{k_2}{k_1}$ choices for V_{i_2,j_2} such that $(V_{i_1,j_1}, V_{i_2,j_2})$ is an edge of R_2. If $\ell \geqslant 3$, then by Lemma 20 and (1) there are also at least $(1-9\varepsilon_1)d_1^2\frac{k_2}{k_1} \geqslant \frac{d_1^2}{2}\frac{k_2}{k_1}$ choices for V_{i_ℓ,j_ℓ} such that $(V_{i_{\ell-1},j_{\ell-1}}, V_{i_\ell,j_\ell})$ is an edge of R_2. By Lemma 20 and (1) again, at most $\frac{1}{1-\varepsilon_2}\sqrt[4]{\varepsilon_1}\frac{k_2}{k_1}$ of those choices are bad for $V_{i_{\ell+1}}$. Of those remaining, at most $\theta\frac{k_2}{k_1}$ have been already used in our construction so far. Since $d_1 \gg \varepsilon_1, \theta$, it follows that there is such a choice for V_{i_ℓ,j_ℓ}. (If $s = 4$, then $V_{i_{s-3}j_{s-3}}$ has already been chosen so we do nothing.) It remains to choose $V_{i_{s-2}j_{s-2}}$ and $V_{i_{s-1}j_{s-1}}$ so that $(V_{i_{s-3}j_{s-3}}, V_{i_{s-2}j_{s-2}}), (V_{i_{s-2}j_{s-2}}, V_{i_{s-1}j_{s-1}})$ and $(V_{i_{s-1}j_{s-1}}, A_s)$ are edges of R_2 and moreover $V_{i_{s-2}j_{s-2}}$ is not bad for $V_{i_{s-1}}$. To see that this can be done, note that as above, there are at least $\frac{d_1^2}{2}\frac{k_2}{k_1}$ choices for $V_{i_{s-2}j_{s-2}}$ so that $(V_{i_{s-3}j_{s-3}}, V_{i_{s-2}j_{s-2}})$ is an edge of R_2 (whether $s = 4$ or not). By Lemma 20 and (1), at most $\frac{1}{1-\varepsilon_2}\sqrt[4]{\varepsilon_1}\frac{k_2}{k_1}$ of those are bad for $V_{i_{s-1}}$. Of those remaining, at most $\theta\frac{k_2}{k_1}$ have been already used. In particular, we have at least $\frac{d_1^2}{3}\frac{k_2}{k_1}$ choices for $V_{i_{s-2}j_{s-2}}$ so that $(V_{i_{s-3}j_{s-3}}, V_{i_{s-2}j_{s-2}})$ is an edge of R_2 and $V_{i_{s-2}j_{s-2}}$ is not bad for $V_{i_{s-1}}$. By the definition of A, there are at least $\frac{\beta}{50}\frac{k_2}{k_1}$ choices for $V_{i_{s-1}j_{s-1}}$, so that $(V_{i_{s-1}j_{s-1}}, A_s)$ is an edge of R_2 and of those at most $\theta\frac{k_2}{k_1}$ have been already used. It remains to show that among all possible choices for $V_{i_{s-2}j_{s-2}}$ and $V_{i_{s-1}j_{s-1}}$ as above, there is such a choice such that $(V_{i_{s-2}j_{s-2}}, V_{i_{s-1}j_{s-1}})$ is an edge of R_2. But this follows from Lemma 20 since $d_1, \beta \gg \varepsilon_1, \theta$.

Repeated application of this argument shows that we can create a cycle of R_2 having the required properties for each cycle of F_2^A containing at least one subcluster of V_0^1. Similarly, we can also create such a cycle for each cycle of F_2^A not containing a subcluster of V_0^1. (For this we need that the length of such a cycle is at least 3, but we already guaranteed that this will be the case.) $\square$

Our next step is to make the pairs of clusters corresponding to edges of F_2 $(2\varepsilon_2, \frac{d_2}{2})$-super-regular, rather than just regular. By Lemma 7 we can achieve this by moving exactly $2\varepsilon_2 m_2'$ vertices from each cluster of R_2 into V_0^2 and thus increasing the size of V_0^2 to at most $3\varepsilon_2 n$. We still write V_0^2 for the new exceptional set and V_{ij} for these altered clusters of R_2 and we denote the sizes of V_{ij} by m_2. So $m_2 = (1-2\varepsilon_2)m_2'$. Note that R_2 has not been altered in any way and all edges of R_2 correspond to $2\varepsilon_2$-regular pairs of density at least $\frac{d_2}{2}$.

For each cycle C of F_2 we now use Lemma 11 to obtain a Hamilton path P_C in the subgraph of G_1 corresponding to C. Note that for the endpoints of this path we may choose any two vertices which lie in any two consecutive clusters of C. We make this choice as follows. First we pick two consecutive clusters, say V_{ij} followed by $V_{k\ell}$ of C such that none of them is a subcluster of V_0^1 and moreover V_{ij} is not bad for V_k. The existence of these two subclusters is guaranteed by Lemma 21. We choose any x_C in $V_{k\ell}$ as the initial vertex of the path. For the endvertex of the path we choose any vertex $y_C \in V_{ij}$ which maximizes $|N^+_G(y_C) \cap V'_k|$, where by V'_k we denote the union of all subclusters of V_k not used in F_2.

Lemma 22. *Let C be a cycle of F_2 and let y_C be chosen as above. Then $|N^+_G(y_C) \cap V'_k| \geqslant \frac{1}{5}d_1^2 d_2 m_1$.*

Proof. Suppose y_C belongs to the subcluster V_{ij} and let $V_{k\ell}$ be the successor of V_{ij} in F_2. By our choice of V_{ij} and $V_{k\ell}$, V_{ij} is not bad for V_k. By (1) and the definition of a subcluster being bad, it follows that there are at least $(1-9\varepsilon_1)d_1^2\frac{k_2}{k_1}$ subclusters $V_{k\ell'}$ of V_k such that $(V_{ij}, V_{k\ell'})$ is an edge of R_2. From this, using Lemma 21 and the fact that $\theta \ll d_1$ we conclude that at least $\frac{d_1^2}{2}\frac{k_2}{k_1}$ of these $V_{k\ell'}$'s are not used in any of the cycles of F_2 and so they belong to V'_k. The result follows since every edge of R_2 corresponds to a $(2\varepsilon_2, d_2/2)$-regular pair in G'_2. $\square$

Let $V_0^2 = \{v_1, \dots, v_r\}$ and let $\{C_1, \dots, C_s\}$ be the set of cycles of F_2. For each $v_i \in V_0^2$, since $\delta^0_G(v_i) \geqslant \beta n$, we can find distinct clusters U_i and W_i in R_1 such that $|N^-_G(v_i) \cap U_i| \geqslant \frac{\beta}{2}m_1$ and $|N^+_G(v_i) \cap W_i| \geqslant \frac{\beta}{2}m_1$. We write P_i for the path from U_i^+ to U_i in the 1-factor F_1 of R_1. For each cycle C_j of F_2, we denote the cluster of R_1 containing x_{C_j} by B_j and write Q_j for the path from B_j to B_j^- in F_1. Define a graph R^* by adding to the vertex set of R_1 all vertices of V_0^2 and one vertex for each cycle C_j of F_2 as follows. For each $1 \leqslant i \leqslant r$ we add the edges $U_i v_i$ and $v_i W_i$ and for each $1 \leqslant j \leqslant s$ we add the edges $B_j^- C_j$ and $C_j B_j$. Now we define the closed walk W described in Section 4. For each $1 \leqslant i, j \leqslant k_1$ we apply Lemma 16 to R_1 to obtain a shifted walk $W(V_i, V_j)$ from V_i to V_j traversing at most $\frac{2}{\nu}$ cycles. We start at V_1 and we incorporate the vertices of V_0^2 by following the walks

$$W(V_1, U_1^+), P_1, U_1v_1W_1, W(W_1, U_2^+), P_2, U_2v_2W_2, \dots, W(W_{r-1}, U_r^+), P_r, U_rv_rW_r.$$

Then we incorporate the cycles of F_2 by following the walks

$$W(W_r, B_1), Q_1, B_1^-C_1B_1, W(B_1, B_2), Q_2, B_2^-C_2B_2, \dots, W(B_{s-1}, B_s), Q_s, B_s^-C_sB_s.$$

Finally, to close the walk and to make sure that W visits every cluster of R_1, we follow the walks

$$W(B_s, V_2), W(V_2, V_3), \dots, W(V_{k_1-1}, V_{k_1}), W(V_{k_1}, V_1)$$

Note that the walk W thus defined visits every v_i and every C_j exactly once, for each cycle C of F_1 it visits every vertex of C the same number of times and for each cluster V of R_1 it visits V at least once and at most

$$\left(\frac{2}{\nu}+1\right)r + \left(\frac{2}{\nu}+1\right)s + \frac{2k_1}{\nu} \leqslant \frac{7\varepsilon_2 n}{\nu} \leqslant \frac{7\varepsilon_2 k_1 m_1}{(1-8\varepsilon_1)\nu} \leqslant \frac{d_1 m_1}{4} \tag{2}$$

times. The last inequality follows since $\varepsilon_2 \ll 1/k_1, d_1, \nu$.

It remains to show how to transform W into a Hamilton cycle of G. Initially, we will transform W to a cycle W' of G with the following properties:

- Each cluster V in W is replaced by an $x \in V' \subseteq V$ in W'. (Recall that V' is the union of all the subclusters of V not used in F_2. Of course, to ensure that W' is a cycle, different x's will be chosen for each appearance of V in W'.)
- Each $v_i \in V_0^2$ in W is left unchanged in W'.
- For each $C_j \in F_2$, we replace C_j in W with the path P_{C_j} in W'.

To achieve this we proceed as follows. For each $1 \leqslant i \leqslant r$ we choose $u_i \in U_i$ and $w_i \in W_i$ such that all of them are distinct and do not belong to the subclusters in F_2 and moreover $u_i v_i$ and $v_i w_i$ are edges of G. To see that this can be done, consider the (undirected) bipartite graph with vertex classes D_1 and D_2 defined as follows. For every $1 \leqslant i \leqslant r$, D_1 contains 2 vertices corresponding to $v_i \in V_0^2$ which we call v_i^+ and v_i^-, while D_2 is the set of all vertices of G lying in some V_k' with $1 \leqslant k \leqslant k_1$. We join v_i^+ to a vertex w of D_2 if and only if $w \in W_i$ and $v_i w$ is an edge of G and we join v_i^- to a vertex u of D_2 if and only if $u \in U_i$ and uv_i is an edge of G. We use Theorem 8 to find a maximal matching in this graph. We claim that this matching covers all vertices of D_1. Indeed, the size of D_1 is at most $6\varepsilon_2 n$, the degree of every vertex of D_1 is at least $(\frac{\beta}{2} - 2\theta)m_1$ and so any matching which does not cover a vertex in D_1 can be extended to a larger matching as $\varepsilon_2 k_1 \ll \theta \ll \beta$. Given this matching from D_1 to D_2, we now take u_i to be the unique vertex in D_2 adjacent to v_i^- and w_i to be the unique vertex in D_2 adjacent to v_i^+ in this matching. Now, for each $1 \leqslant j \leqslant s$ we choose $b_j \in B_j$ and $b_j^- \in B_j^-$ such that all of them are distinct, they are distinct from the u_i, w_i $(1 \leqslant i \leqslant r)$, they do not belong to the subclusters used in F_2 and moreover $b_j^- x_{C_j}$ and $y_{C_j} b_j$ are edges of G. To achieve this, consider the bipartite graph with vertex classes D_3 and D_4 defined as follows: $D_3 = \{x_{C_j}, y_{C_j} : 1 \leqslant j \leqslant s\}$, $D_4 = D_2 \setminus \{u_i, w_i : 1 \leqslant i \leqslant r\}$, with x_{C_j} adjacent to $b^- \in D_4$ if and only if $b^- \in B_j^-$ and $b^- x_{C_j}$ is an edge of G, and y_{C_j} adjacent to $b \in D_4$ if and only if $b \in B_j$ and $y_{C_j} b$ is an edge of G. As before, we use Theorem 8 to find a maximal matching in this graph and claim that this matching covers all vertices of D_3. Indeed, if there was a vertex v of D_3 not covered by the matching, then we could extend the matching either by Lemma 22 if $v = y_{C_j}$ for some j, or by super-regularity of the pair $(B_j^-, B_j)_{G_1}$ if $v = x_{C_j}$ for some j. Given this matching, we can now take b_j to be the unique vertex adjacent to y_{C_j} and b_j^- to be the unique vertex adjacent to x_{C_j} in this matching.

Now we use W to join up the vertices u_i, w_i, b_j, b_j^- by disjoint paths whose edges join clusters corresponding to the relevant edges of W. (For example, the path joining up w_1 to u_2 moves through the clusters in the subwalk $W(W_1, U_2^+)P_2U_2$ of W.) Delete all the vertices in V_0^2 as well as $C_1, \dots, C_s$ from W to obtain a set $\mathcal{W}$ of subwalks of W. So each walk in $\mathcal{W}$ corresponds to one of the paths joining up the vertices u_i, w_i, b_j, b_j^- we are looking for. To choose these paths we first fix edges in G corresponding to all those edges of the walks in $\mathcal{W}$ that do not lie within a cycle of F_1. This can be done by looking at all ordered pairs (V_i, V_j) with $V_j \neq V_i^+$ in turn. Let w_{ij} be the number of times the edge $V_i V_j$ is used by walks in $\mathcal{W}$. We need to choose a matching in G that avoids all previously chosen vertices and uses w_{ij} edges from V_i' to V_j'. (Recall that V_i' is the union of all subclusters of V_i not used in F_2.) To see that this matching exists, recall that the pair (V_i, V_j) is $(6\varepsilon_1, \frac{d_1}{3})$-regular and so the pair obtained from (V_i', V_j') by deleting all the previously chosen vertices is still $(7\varepsilon_1, \frac{d_1}{4})$-regular. Since $w_{ij} \ll d_1 m_1$ by (2), this implies the existence of the required matching from V_i' to V_j'. Theorem 8 now implies that there is a NC^4 algorithm for finding such a matching. After considering all such pairs (V_i, V_j) we have found edges in G corresponding to all those edges of the walks in $\mathcal{W}$ that do not lie within a cycle of F_1. Finally, we can apply Lemma 9 with F_1 playing the role of R and with the subgraph of G_1 which corresponds to F_1 playing the role of G to find paths that connect all the vertices chosen so far. (So these paths correspond to the set $\mathcal{W}'$ of walks obtained from the walks in $\mathcal{W}$ by deleting the edges outside F_1. Lemma 9 can be applied since $|\mathcal{W}'| \le (r+s+k_1)\frac{3}{\nu} \le \sqrt{\varepsilon_2}n$ and since each walk in $\mathcal{W}'$ has length at least 3 and at most k_1.) Together with the previously chosen edges of G and the paths P_{C_j} covering the vertices lying in the subclusters belonging to F_2, this yields a cycle W' in G as required.

Finally, we extend W' to a Hamilton cycle of G. For this note that by (2) for each cycle C of F_1, W' has visited every cluster of C exactly m_C times for some $m_C \leqslant \frac{d_1 m_1}{4}$. Fix one particular occasion on which W 'winds around' C. It is enough to show that we can replace the corresponding path P in W' by a new path with the same endpoints exhausting all vertices in the clusters of C which do not appear in W'. To do this, remove all vertices from the clusters of C which are used in W' apart from the ones used in P. Since exactly $m_C - 1 \leqslant \frac{m_1 d_1}{4}$ vertices have been removed and since the pairs of clusters corresponding to the edges of F_1 are $(6\varepsilon_1, \frac{d_1}{3})$-super-regular, the modified clusters are now $(12\varepsilon_1, \frac{d_1}{12})$-super-regular and so we can use Lemma 11 to replace P by a new path with the required property.

To see that the algorithm is in NC^5, note that at most steps of the algorithm we either use one of Lemmas 6,7,9,11,18 or we use Theorem 8 or we work entirely within one of the reduced digraphs (which have constant size). The only other steps of the algorithm which we need to check are when obtaining G_1 from G_1' and when defining the vertices y_C for each cycle C of F_2. To obtain G_1 from G_1' we only need knowledge of the in-degrees and out-degrees of each vertex x within each cluster V_k, which can be found in NC^1. Similarly, to define each y_C we only need knowledge of the out-degrees of each vertex y within each set V_k' which can again be found in NC^1. This completes the proof of Theorem 5.

References

[1] N. Alon, R. A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms* **16** (1994), 80–109.

[2] N. Alon, G. Gutin and M. Krivelevich, Algorithms with large domination ratio, *J. Algorithms* **50** (2004), 118–131.

[3] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *J. Comput. System Sci.* **69** (2004), 353–382.

[4] N. Alon and J. H. Spencer, *The probabilistic method*, Second edition, Wiley-Intersci., 2000.

[5] J. A. Bondy and V. Chvátal, A method in graph theory, *Discrete Math.* **15** (1976), 111–135.

[6] V. Chvátal, On Hamilton's ideals, *J. Combinatorial Theory B* **12** (1972), 163–168.

[7] B. Csaba, On the Bollobás-Eldridge conjecture for bipartite graphs, *Combin. Probab. Comput.* **16** (2007), 661–691.

[8] E. Dahlhaus, P. Hajnal and M. Karpinski, On the parallel complexity of Hamiltonian cycle and matching problem on dense graphs, *J. Algorithms* **15** (1993), 367–384.

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

[10] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *C. R. Acad. Sci. Paris* **251** (1960), 495–497.

[11] F. Glover and A.P. Punnen, The Travelling Salesman Problem: New solvable cases and linkages with the development of approximation algorithms, *Journal of the Operational Research Society* **48** (1997), 502–510.

[12] R. M. Karp and V. Ramachandran, Parallel algorithms for shared-memory machines, in *Handbook of theoretical computer science, Vol. A*, Elsevier, Amsterdam (1990), 869–941.

[13] R. Häggkvist, Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **2** (1993), 25–32.

[14] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, *J. London Math. Soc.* **79** (2009), 144–166.

[15] L. Kelly, D. Kühn and D. Osthus, A Dirac type result on Hamilton cycles in oriented graphs, *Combin. Probab. Comput* 17 (2008), 689–709.

[16] J. Komlós, The blow-up lemma, *Combin. Probab. Comput.* **8** (1999), 161–176.

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

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

[19] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, in *Combinatorics, Paul Erdős is eighty, Vol. 2*, Bolyai Soc. Math. Stud. **2**, János Bolyai Math. Soc., Budapest (1996), 295–352.
[20] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, in *Surveys in Combinatorics* (S. Huczynska, J.D. Mitchell and C.M. Roney-Dougal eds.), *London Math. Soc. Lecture Notes* **365**, 137–167, Cambridge University Press, 2009.
[21] D. Kühn and D. Osthus, Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments, preprint.
[22] D. Kühn, D. Osthus and A. Treglown, Hamiltonian degree sequences in digraphs, *J. Combinatorial Theory B* **100** (2010), 367–380.
[23] G. Lev, Size bounds and parallel algorithms for networks, Tech. Rep. CST-8-80, Dept. of Comput. Sci., Univ. of Edinburgh, 1980.
[24] C. St. J. A. Nash-Williams, Hamiltonian circuits, in *Studies in graph theory, Part II*, Studies in Math. **12**, Math. Assoc. Amer., Washington (1975), 301–360.
[25] G. N. Sárközy, A fast parallel algorithm for finding Hamiltonian cycles in dense graphs, to appear in *Discrete Math.*
[26] E. Szemerédi, Regular partitions of graphs, in *Problèmes combinatoires et théorie des graphes*, Colloq. Internat. CNRS **260**, Paris (1978), 399–401.
[27] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in combinatorics*, Cambridge Univ. Press, Cambridge (1981), 211–228.

Peter Keevash & Demetres Christofides
School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London
E1 4NS
UK

Daniela Kühn & Deryk Osthus
School of Mathematics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

E-mail addresses: p.keevash@qmul.ac.uk, christofidesdemetres@gmail.com, {d.kuhn,d.osthus}@bham.ac.uk