Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments

Kuhn, Daniela; Osthus, Deryk

DOI:
10.1016/j.aim.2013.01.005

HAMILTON DECOMPOSITIONS OF REGULAR EXPANDERS: A PROOF OF KELLY'S CONJECTURE FOR LARGE TOURNAMENTS

DANIELA KÜHN AND DERYK OSTHUS

Abstract. A long-standing conjecture of Kelly states that every regular tournament on n vertices can be decomposed into $(n-1)/2$ edge-disjoint Hamilton cycles. We prove this conjecture for large n. In fact, we prove a far more general result, based on our recent concept of robust expansion and a new method for decomposing graphs. We show that every sufficiently large regular digraph G on n vertices whose degree is linear in n and which is a robust outexpander has a decomposition into edge-disjoint Hamilton cycles. This enables us to obtain numerous further results, e.g. as a special case we confirm a conjecture of Erdős on packing Hamilton cycles in random tournaments. As corollaries to the main result, we also obtain several results on packing Hamilton cycles in undirected graphs, giving e.g. the best known result on a conjecture of Nash-Williams. We also apply our result to solve a problem on the domination ratio of the Asymmetric Travelling Salesman problem, which was raised e.g. by Glover and Punnen as well as Alon, Gutin and Krivelevich.

1. Introduction

1.1. **Kelly's conjecture.** A graph or digraph G has a Hamilton decomposition if it contains a set of edge-disjoint Hamilton cycles which together cover all the edges of G. The study of Hamilton decompositions is one of the oldest and most natural problems in Graph Theory. For instance, in 1892 Walecki showed that the complete graph K_n on n vertices has a Hamilton decomposition if n is odd (see e.g. [4, 5, 39]). Tillson [47] solved the corresponding problem for complete digraphs. Here every pair of vertices is joined by an edge in each direction, and there is a Hamilton decomposition unless the number of vertices is 4 or 6.

However, though there are several deep conjectures in the area, little progress has been made so far in proving results on Hamilton decompositions for general classes of graphs. Possibly the most well known problem in this direction is Kelly's conjecture from 1968 (see e.g. the monographs and surveys [7, 10, 36, 40]), which states that every regular tournament has a Hamilton decomposition. Here a tournament is an orientation of a complete (undirected) graph. It is regular if the indegree of every vertex equals its outdegree. This condition is clearly necessary for a Hamilton decomposition. Here, we prove this conjecture for all large tournaments. In fact, it turns out that we can prove a much stronger result – we can obtain a Hamilton

Date: May 10, 2013.

The authors were supported by the EPSRC, grant no. EP/J008087/1. D. Kühn was also supported by the ERC, grant no. 258345.

decomposition of any regular orientation of a sufficiently dense graph. More precisely, an oriented graph G is obtained by orienting the edges of an undirected graph. So it contains no cycles of length two (whereas in a digraph this is permitted).

Theorem 1.1. *For every $\varepsilon > 0$ there exists n_0 such that every r-regular oriented graph G on $n \geq n_0$ vertices with $r \geq 3n/8 + \varepsilon n$ has a Hamilton decomposition. In particular, there exists n_0 such that every regular tournament on $n \geq n_0$ vertices has a Hamilton decomposition.*

It is not clear whether the lower bound on r in Theorem 1.1 is best possible. However, as discussed below, there are oriented graphs whose in- and outdegrees are all very close to $3n/8$ but which do not contain even a single Hamilton cycle. Moreover, for $r < (3n-4)/8$, it is not even known whether an r-regular oriented graph contains a single Hamilton cycle (this is related to a conjecture of Jackson, see the survey [36] for a more detailed discussion). Both these facts indicate that any improvement in the lower bound on r would be extremely difficult to obtain.

Regular tournaments obviously exist only if n is odd, but we still obtain an interesting corollary in the even case. Suppose that G is a tournament on n vertices where n is even and which is as regular as possible, i.e. the in- and outdegrees differ by 1. Then Theorem 1.1 implies that G has a decomposition into edge-disjoint Hamilton paths. Indeed, add an extra vertex to G which sends an edge to all vertices of G whose indegree is below $(n-1)/2$ and which receives an edge from all others. The resulting tournament G' is regular, and a Hamilton decomposition of G' clearly corresponds to a decomposition of G into Hamilton paths.

The difficulty of Kelly's conjecture is illustrated by the fact that even the existence of two edge-disjoint Hamilton cycles in a regular tournament is not obvious. The first result in this direction was proved by Jackson [22], who showed that every regular tournament on at least 5 vertices contains a Hamilton cycle and a Hamilton path which are edge-disjoint. Zhang [48] then demonstrated the existence of two edge-disjoint Hamilton cycles. These results were improved by considering Hamilton cycles in oriented graphs of large in- and outdegree by Thomassen [46], Häggkvist [20], Häggkvist and Thomason [21] as well as Kelly, Kühn and Osthus [25]. Keevash, Kühn and Osthus [24] then showed that every sufficiently large oriented graph G on n vertices whose in- and outdegrees are all at least $(3n-4)/8$ contains a Hamilton cycle. This bound on the degrees is best possible and confirmed a conjecture of Häggkvist [20] (as mentioned above, there are extremal constructions which are almost regular). Note that this result implies that every sufficiently large regular tournament on n vertices contains at least $n/8$ edge-disjoint Hamilton cycles, whereas Kelly's conjecture requires $(n-1)/2$ edge-disjoint Hamilton cycles. The conjecture has also been proved for small values of n and for several special classes of tournaments (see e.g. [5, 8] for somewhat outdated surveys).

Recently, Kühn, Osthus and Treglown [38] proved an approximate version of Theorem 1.1 by showing that every r-regular oriented graph G on $n \geq n_0(\varepsilon)$ vertices with $r \geq 3n/8+\varepsilon n$ has an approximate Hamilton decomposition (i.e. a set of edge-disjoint Hamilton cycles covering almost all edges).

1.2. **Robust outexpanders.** In fact, we prove a theorem which is yet more general than Theorem 1.1. Moreover, rather than being based on a degree condition, it uncovers an underlying structural property which guarantees a Hamilton decomposition. As discussed in the next subsection, this property is shared by several well known classes of digraphs. Roughly speaking, this notion of 'robust expansion' is defined as follows: for any set S of vertices, its robust outneighbourhood is the set of vertices which receive many edges from S. A digraph is a robust outexpander if for every set S which is not too small and not too large, its robust outneighbourhood is at least a little larger than S. This notion was introduced explicitly in [37], and was already used implicitly in the earlier papers [24, 25]. In these papers, we proved approximate and exact versions of several conjectures on Hamilton cycles in digraphs.

More precisely, let $0 < \nu \leq \tau < 1$. Given any digraph G on n vertices and $S \subseteq V(G)$, the *ν-robust outneighbourhood $RN^+_{\nu,G}(S)$ of S* is the set of all those vertices x of G which have at least νn inneighbours in S. G is called a *robust (ν, τ)-outexpander* if

$$|RN^+_{\nu,G}(S)| \geq |S| + \nu n \quad \text{for all} \quad S \subseteq V(G) \quad \text{with} \quad \tau n \leq |S| \leq (1-\tau)n.$$

Our main result states that every sufficiently large regular robust outexpander has a Hamilton decomposition.

Theorem 1.2. *For every $\alpha > 0$ there exists $\tau > 0$ such that for all $\nu > 0$ there exists $n_0 = n_0(\alpha, \nu, \tau)$ for which the following holds. Suppose that*

(i) *G is an r-regular digraph on $n \geq n_0$ vertices, where $r \geq \alpha n$;*
(ii) *G is a robust (ν, τ)-outexpander.*

Then G has a Hamilton decomposition. Moreover, this decomposition can be found in time polynomial in n.

Since Lemma 13.1 states that every oriented graph G on n vertices with minimum in- and outdegree at least $3n/8 + \varepsilon n$ is a robust outexpander (provided that n is sufficiently large compared to ε), Theorem 1.2 immediately implies Theorem 1.1.

Obviously, the condition that G is regular is necessary. The robust expansion property can be viewed as a natural strengthening of this property: Indeed, suppose that $\nu^{1/4} \leq \tau \leq \alpha$ and $|S| \geq \tau n$. Counting the edges from S to its ν-robust outneighbourhood shows that condition (i) already forces the ν-robust outneighbourhood of S to have size at least $(1 - \sqrt{\nu})|S|$. Condition (ii) then ensures (amongst others) that G is highly connected, which is obviously also necessary.

The following result of Osthus and Staden [42] gives an approximate version of Theorem 1.2 and will be used in its proof.

Theorem 1.3. *For every $\alpha > 0$ there exists $\tau > 0$ such that for all $\nu, \eta > 0$ there exists $n_0 = n_0(\alpha, \nu, \tau, \eta)$ for which the following holds. Suppose that*

(i) *G is an r-regular digraph on $n \geq n_0$ vertices, where $r \geq \alpha n$;*
(ii) *G is a robust (ν, τ)-outexpander.*

Then G contains at least $(1 - \eta)r$ edge-disjoint Hamilton cycles. Moreover, this set of Hamilton cycles can be found in time polynomial in n.

Note that Theorem 1.3 is a generalization of the approximate version of Theorem 1.1 proved in [38] (which was already mentioned at the end of the previous subsection). The approach in [38] is not algorithmic though. If we replace the use of Theorem 1.3 in our main proof by the result in [38], this yields exactly Theorem 1.1 (but not its algorithmic version). However, this would not result in a shorter proof of Theorem 1.2.

1.3. **Further applications.** In this section, we briefly discuss further applications of our main result.

1.3.1. *Regular digraphs and TSP tour domination.* As observed in Section 13, it is very easy to check that regular digraphs of sufficiently large degree are robust outexpanders. Together with Theorem 1.4 this implies the following result.

Theorem 1.4. *For every $\varepsilon > 0$ there exists n_0 such that every r-regular digraph G on $n \geq n_0$ vertices with $r \geq (1/2 + \varepsilon)n$ has a Hamilton decomposition. Moreover, such a decomposition can be found in time polynomial in n.*

Surprisingly, this has an immediate application to the area of TSP tour domination. More precisely, the Asymmetric travelling salesman problem (ATSP) asks for a Hamilton cycle of least weight in an edge-weighted complete digraph (where opposite edges are allowed to have different weight). An algorithm A for the ATSP has *domination ratio* $p(n)$ if it has the following property. For any problem instance I let $w(I)$ be the weight of the solution produced by A. Then for all n and for all instances I on n vertices, there are at least $p(n)(n-1)!$ solutions to instance I whose weight is also at least $w(I)$. (Note that the total number of possible solutions is $(n-1)!$.) This notion is of particular interest for the ATSP as it is not known whether there is a polynomial time algorithm for the ATSP whose approximation ratio is bounded by an absolute constant. Several well known TSP algorithms achieve a domination ratio of $\Omega(1/n)$ for the ATSP but no better results are known. In particular, a long-standing open problem (see e.g. Glover and Punnen [18], Gutin and Yeo [19] as well as Alon, Gutin and Krivelevich [2]) asks whether there is a polynomial time algorithm which achieves a constant domination ratio for the ATSP. Gutin and Yeo [19] proved that the existence of a polynomial time algorithm with domination ratio $1/2 - \varepsilon$ would follow from an algorithmic proof of Theorem 1.4. So the result of [19] together with Theorem 1.4 yields the following.

Corollary 1.5. *For any $\varepsilon > 0$, there is a polynomial time algorithm for the ATSP whose domination ratio is $1/2 - \varepsilon$.*

1.3.2. *Random tournaments.* Another application of Theorem 1.2 confirms a conjecture of Erdős (see [45]) which can be regarded as a probabilistic version of Kelly's conjecture. Given an oriented graph G, let $\delta^+(G)$ denote its minimum outdegree and $\delta^-(G)$ its minimum indegree. Clearly, the minimum of these two quantities is an upper bound on the number of edge-disjoint Hamilton cycles that G can have. Erdős conjectured that this bound is correct with high probability if G is a random tournament and one can use Theorem 1.1 to show this is indeed the case.

Theorem 1.6. *Almost all tournaments G contain $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$ edge-disjoint Hamilton cycles.*

More precisely, the term 'almost all' refers to the model where one considers the set $\mathcal{T}_n$ of all tournaments on n vertices and shows that the probability that a random tournament in $\mathcal{T}_n$ has the required number of Hamilton cycles tends to 1 as n tends to infinity. We prove Theorem 1.6 by showing that with high probability G contains a $\delta^0(G)$-regular spanning subdigraph G' and apply Theorem 1.1 to G' to find the required Hamilton cycles. As the first step requires some work, we defer this to a shorter companion paper [35].

The corresponding problem for the binomial random graph $G_{n,p}$ with edge probability p has a long history, going back to a result of Bollobás and Frieze [9], who showed that a.a.s. (asymptotically almost surely) $G_{n,p}$ contains $\lfloor \delta(G_{n,p})/2 \rfloor$ edge-disjoint Hamilton cycles in the range of p where the minimum degree $\delta(G_{n,p})$ is a.a.s. bounded. A striking conjecture of Frieze and Krivelevich [17] asserts that this result extends to arbitrary edge probabilities p. The range of p was extended in several papers, in particular due to recent results of Knox, Kühn and Osthus [26] as well as Krivelevich and Samotij [30], the conjecture remains open only in the (rather special) case when p tends to 1 fairly quickly. As we shall observe in [35], this case follows from Theorem 1.2 in a similar way as Theorem 1.6.

1.3.3. *Undirected robust expanders.* In [35], we also derive an undirected version of Theorem 1.2, where instead of the 'robust outneighbourhood' we consider the 'robust neighbourhood'. As an immediate corollary of this undirected version, we obtain the following approximate version of the 'Hamilton decomposition conjecture' of Nash-Williams [41].

Theorem 1.7. *For every $\varepsilon > 0$ there exists n_0 such that every r-regular graph G on $n \geq n_0$ vertices, where $r \geq (1/2 + \varepsilon)n$ is even, has a Hamilton decomposition.*

The conjecture of Nash-Williams asserts that the εn error term can be removed. Theorem 1.7 improves results by Christofides, Kühn and Osthus [13] as well as Perkovic and Reed [43].

Finally, the undirected version of Theorem 1.2 easily implies that every even-regular dense quasi-random graph has a Hamilton decomposition. An approximate version of this result was proved earlier by Frieze and Krivelevich [16]. Our undirected decomposition result implies e.g. a recent result of Alspach, Bryant and Dyer [6] that every Paley graph has a Hamilton decomposition (for the case of large graphs). Our undirected decomposition result also implies that with high probability, dense random regular graphs of even degree have a Hamilton decomposition. (Hamilton decompositions of random regular graphs of bounded degree have already been studied intensively.) These and other related results are discussed in more detail in [35].

1.3.4. *The robust decomposition lemma.* In a sequence of four papers [14, 15, 32, 33] we build on the results of the current paper to prove the long-standing '1-factorization conjecture': Suppose that n is even and sufficiently large, and that $D \geq 2\lceil n/4 \rceil - 1$. Then every D-regular graph G on n vertices has a decomposition into perfect matchings. (Equivalently, $\chi'(G) = D$.) Moreover, we improve Theorem 1.7 to completely solve the Hamilton decomposition conjecture of Nash-Williams from [41].

Finally, we also solve another problem of Nash-Williams on optimal packings of edge-disjoint Hamilton cycles in graphs of large minimum degree (the latter also uses results from [31]). The proofs are based on (the undirected version of) Theorem 1.2 as well as the 'robust decomposition lemma', which can be viewed as a version of Theorem 1.2 which is more technical but has the advantage of being more widely applicable (see Lemma 11.2 or 12.1).

In the next section, we give a brief outline of our methods. The approach is a very general one and we are certain that it will have significant further applications.

2. A brief outline of the argument

2.1. **The general approach.** The basic idea behind the proof of Theorem 1.2 can be described as follows. Let G be a robustly expanding digraph as in Theorem 1.2. Suppose that inside G we can find a sparse regular digraph $H^{\rm rob}$ which is robustly decomposable in the sense that it still has a Hamilton decomposition if we add a few edges to it. More precisely, $H^{\rm rob}$ is *robustly decomposable* if $H^{\rm rob} \cup H_0$ has a Hamilton decomposition whenever H_0 is a very sparse regular digraph which is edge-disjoint from $H^{\rm rob}$ and such that $V(H) = V(H^{\rm rob})$. Then Theorem 1.2 would be an immediate consequence of the existence of such an $H^{\rm rob}$. Indeed, first we remove the edges of $H^{\rm rob}$ from G to obtain G'. Then we apply Theorem 1.3 to G' to find edge-disjoint Hamilton cycles covering almost all edges of G'. Let H_0 be the leftover – i.e. the set of edges of G' which are not covered by one of these Hamilton cycles. Now apply the fact that $H^{\rm rob}$ is robustly decomposable to obtain a Hamilton decomposition of $H^{\rm rob} \cup H_0$ and thus of G. Essentially, this is what the 'robust decomposition lemma' (Lemma 11.2) achieves.

Unfortunately, we do not know how to construct such a digraph $H^{\rm rob}$ directly – the main problem is that we have almost no control over what H_0 might look like. However, one key idea is that we can define and find several digraphs $H_i^{\rm rob}$ which together play the role of $H^{\rm rob}$. Indeed, suppose that we remove the edges of several $H_i^{\rm rob}$ at the start and let H_0 be the leftover of the approximate decomposition as above. Then we can show that $H_1^{\rm rob} \cup H_0$ contains a set of edge-disjoint Hamilton cycles so that the resulting leftover H_1 has more structure than H_0 (i.e. it has some useful properties). This in turn means that we can improve on the previous step and now find a set of edge-disjoint Hamilton cycles in $H_2^{\rm rob} \cup H_1$ so that the resulting leftover H_2 has even more structure than H_1. After $\ell - 1$ steps, $H_{\ell-1}$ will be a sufficiently 'nice' digraph so that $H_\ell^{\rm rob} \cup H_{\ell-1}$ does have a Hamilton decomposition. This very general approach was first introduced in [26], where we used it to find optimal packings of edge-disjoint Hamilton cycles in random graphs. In [26], the aim was that successive H_i become sparser. In our setting, the density is less relevant – our aim is to obtain successively stronger structural properties for the H_i.

When finding Hamilton cycles in $H_i^{\rm rob} \cup H_{i-1}$, we usually proceed as follows (except for the final step, i.e. when $i = \ell$). First we decompose H_{i-1} into edge-disjoint 1-factors (where a 1-factor is a spanning union of vertex-disjoint cycles). Each of these 1-factors is then split into a set of paths in a suitable way (we call this a path system). In particular, the number of edges in each path system is small compared to n. Then we extend each path system into a suitable 1-factor F using edges of $H_i^{\rm rob}$. Then we

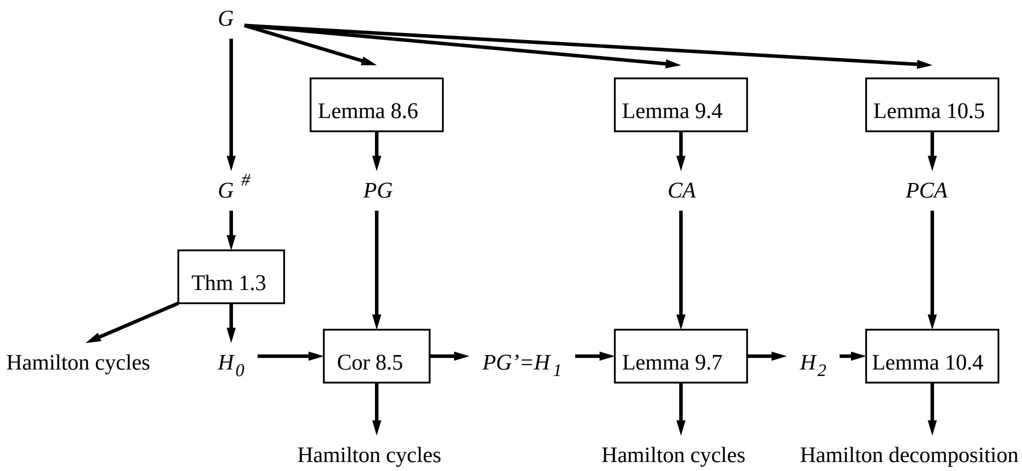

FIGURE 1. An illustration of the structure of the proof of Theorem 1.2. The purpose of Lemmas 8.6, 9.4 and 10.5 is to find the preprocessing graph PG, the chord absorber CA and the parity extended cycle absorber PCA respectively. The purpose of Corollary 8.5, Lemmas 9.7 and 10.4 is to use these graphs to find suitable Hamilton cycles. In Section 12, we combine Lemmas 9.4, 10.5, 9.7 and 10.4 into a single 'robust decomposition lemma' (see Lemma 11.2 or 12.1).

transform F into a Hamilton cycle C, again using edges of $H_i^{\rm rob}$. The resulting set of Hamilton cycles then covers all edges of H_{i-1}. In other words, H_{i-1} is 'absorbed' into the set of Hamilton cycles that we have constructed so far and the leftover H_i is a subgraph of $H_i^{\rm rob}$. In particular, H_i inherits any structural properties that $H_i^{\rm rob}$ has. When constructing the Hamilton cycle C, we will use the result of [37] that every robustly expanding digraph contains a Hamilton cycle, or one of its corollaries (see Section 6).

2.2. **The main steps.** The construction of the graphs $H_i^{\rm rob}$ involves Szemerédi's regularity lemma. We apply this to G to obtain a partition of its vertices into clusters $V_1, \dots, V_k$ and a small exceptional set V_0 so that almost all ordered pairs of clusters induce a pseudorandom subdigraph of G. As is well known, one can then define a 'reduced digraph' R whose vertex set consists of the clusters V_i, with an edge from V_i to V_j, if the subdigraph of G induced by the edges of G from V_i to V_j is pseudorandom and dense. R inherits many of the properties of G. In particular, R is also a robust outexpander. So it contains a Hamilton cycle C by the result mentioned above – without loss of generality $C = V_1V_2 \dots V_k$.

We can now define three digraphs playing the role of $H_1^{\rm rob}, H_2^{\rm rob}, H_3^{\rm rob}$ above (so we have $\ell = 3$ in our setting):

- the preprocessing graph PG;
- the chord absorber CA;
- the parity extended cycle absorber PCA.

We now describe the purpose of these three digraphs (see also Figure 1).

The *preprocessing graph* PG has the following property: let H_1 be the leftover of $PG \cup H_0$ after removing suitably chosen edge-disjoint Hamilton cycles. (So as discussed above, H_1 is a subdigraph of PG.) Then H_1 has no edges incident to V_0. Thus H_1 is not a regular digraph (but it will be a regular subdigraph of $G - V_0$). This degree irregularity will be compensated for by constructing the chord absorber CA in a suitable way.

The *chord absorber* CA has the following property: let H_2 be the leftover of $CA \cup H_1$ after removing suitably chosen edge-disjoint Hamilton cycles. Then H_2 is a blow-up of C. In other words, for every edge e of H_2, there is a j so that the initial vertex of e is in V_j and the final one in V_{j+1}. (So the edges of H_2 'wind around' C.) H_2 will be a subdigraph of CA. But CA itself will not only consist of a blow-up $\mathcal{B}(C)$ of C, it will also contain a set of suitably chosen 'chord edges' between clusters which are not adjacent on C. These chord edges lie in a digraph $\mathcal{B}(U')$, which is a blow-up of a 'universal walk' U on the clusters V_i. To absorb H_1, we will split its edges into path systems M as described above. We would like to extend each M into a Hamilton cycle. This may be impossible using the edges of $\mathcal{B}(C)$ alone, e.g. if M contains an edge e_1 from V_1 to V_3 but no other edges. So for each edge e of such a path system M, we then choose a set of chord edges from $\mathcal{B}(U')$ which 'balance out' this edge e to form a 'locally balanced sequence'. We extend (and balance) M in this way to obtain a path system M'. We then further extend M' to a Hamilton cycle using edges of $\mathcal{B}(C)$.

As an example, suppose again that $M = \{e_1\}$ with e_1 being an edge from V_1 to V_3. It turns out that a simple way of balancing e_1 would be to add an edge e_j from V_j to V_{j+2} for all j with $1 < j \le k$, so that $e_1, e_2, \dots, e_k$ form a matching M'. It is easy to see that one can extend M' into a Hamilton cycle using edges from $\mathcal{B}(C)$, i.e. edges which only wind around C. Indeed, start by traversing e_1, then wind around C to reach the initial vertex of e_2, then traverse e_2, then wind around C again, and eventually traverse e_k. Before returning to the initial vertex of e_1, wind around C sufficiently many times to visit every vertex in every cluster, thus obtaining a Hamilton cycle which contains M'. Unfortunately, we cannot guarantee the existence of such edges $e_2, \dots, e_k$ in an arbitrary robust outexpander. So we will use sequences of balancing edges which involve more edges but can be found in any robust outexpander. (They will be based on the concept of 'shifted walks' which we introduced in [25]).

The crucial point is that we can carry out the balancing in such a way that we use up *all* edges of $\mathcal{B}(U')$ in the process of balancing out the path systems of H_1. In particular, the surprising feature of the argument is that we can choose $\mathcal{B}(U')$ in advance (i.e. without knowing H_1) so that it has this property.

In the above, we did not discuss edges incident to the exceptional set V_0. Obviously a Hamilton cycle has to contain these, whereas neither of H_1, $\mathcal{B}(C)$ and $\mathcal{B}(U')$ have any edges incident to V_0. To deal with this, we introduce the following trick: suppose for example that V_0 contains a single exceptional vertex x. We find an outneighbour x^+ of x and an inneighbour x^- in $V(G) \setminus V_0$. We can then define an 'exceptional edge' x^-x^+ and add this edge x^-x^+ to the chord absorber CA. Then a Hamilton cycle of $(CA \cup H_1) - V_0$ containing this exceptional edge corresponds to a Hamilton cycle of G. The systematic use of exceptional edges in this way allows us to ignore

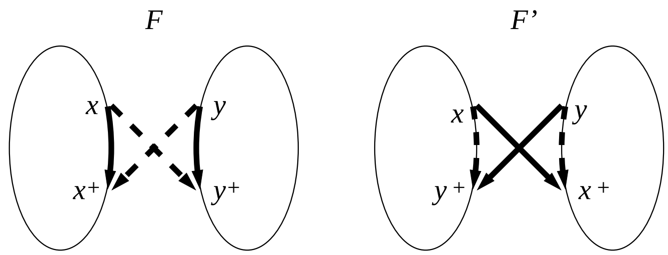

FIGURE 2. Transforming the 1-factors F and F' consisting of two cycles into Hamilton cycles by switching edges.

the exceptional set V_0 at many points of the argument. It might seem natural to apply this trick directly to H_0 (i.e. replace pairs of edges of H_0 incident to V_0 with exceptional edges) with the aim of making the preprocessing step unnecessary. However, this approach would run into considerable difficulties. It turns out that working inside H_0 (rather than the whole of G) when constructing the exceptional edges might not give us enough choice to find suitable sets of exceptional edges.

Finally, the *parity extended cycle absorber* PCA has the following property: let H_2 be the leftover of $CA \cup H_1$ after removing suitable Hamilton cycles. Then $H_2 \cup PCA$ has a Hamilton decomposition. We will find this Hamilton decomposition as follows: first we decompose $PCA \cup H_2$ into carefully chosen 1-factors (in particular, either half or all of the edges of each 1-factor are contained in PCA). When finding this decomposition, we use the fact that H_2 is a blow-up of C. (PCA will also be a blow-up of C.) Our aim is then to turn this 1-factorization of $PCA \cup H_2$ into a Hamilton decomposition of $PCA \cup H_2$ by successively switching edges between 1-factors. As an illustration, suppose that we are given two 1-factors F and F' so that F contains the edges xx^+ and yy^+. Similarly, suppose that F' contains the edges xy^+ and yx^+. Note that these edges form an orientation of a four-cycle C_4. Now perform a 'switch', which consists of removing xx^+ and yy^+ from F, adding xy^+ and yx^+ to F and proceeding similarly for F'. This yields two new 1-factors F_{new} and F'_{new}. Suppose that F consists of exactly two cycles and that xx^+ and yy^+ lie on different cycles. Then F_{new} is a Hamilton cycle. The same holds for F' (see Figure 2). These switches will always involve edges from PCA and not from H_2. So if we ensure that PCA has switches at the right places, we can eventually turn the 1-factorization of $PCA \cup H_2$ into a Hamilton decomposition after several switches.

This paper is organized as follows: In the next section, we introduce some notation. In Section 4, we collect tools which we will need in connection with Szemerédi's regularity lemma. Similarly, in Section 5 we collect general properties of robustly expanding digraphs. Section 6 is devoted to tools for finding Hamilton cycles (in robustly expanding digraphs). In Section 7, we introduce a systematic and convenient way of dealing with exceptional vertices which will be used throughout the remainder of the paper. This will be based on the concept of exceptional edges and 'balancing' edges via chord sequences. Section 8 deals with the preprocessing step, which involves the preprocessing graph PG. Then, in Section 9, we define, find and use the chord-absorber CA. Switches and the parity extended cycle absorber PCA

are then introduced in Section 10. In Section 11, we put everything together to prove Theorem 1.2. In Section 12, we state a standalone variant of the 'robust decomposition lemma', for use e.g. in [14, 32]. Finally, in Section 13 we derive Theorems 1.1 and 1.4 from Theorem 1.2.

3. Notation and probabilistic estimates

3.1. **Notation.** Given a graph or digraph G, we write $V(G)$ for its vertex set, $E(G)$ for its edge set, $e(G) := |E(G)|$ for the number of its edges and $|G|$ for the number of its vertices. Given $X \subseteq V(G)$, we write $G - X$ for the (di)graph obtained from G by deleting all vertices in X. Given $F \subseteq E(G)$, we write $G \setminus F$ for (di)graph obtained from G by deleting all edges in F. If H is a sub(di)graph of G, we write $G \setminus H$ for $G \setminus E(H)$.

Suppose that G is an undirected graph. We write $\delta(G)$ for the minimum degree of G and $\Delta(G)$ for its maximum degree. Whenever $X, Y \subseteq V(G)$, we write $e_G(X, Y)$ for the number of all those edges of G which have one endvertex in X and the other endvertex in Y. If $X \cap Y = \emptyset$, we denote by $G[X, Y]$ the bipartite subgraph of G with vertex sets X and Y whose edges are all the edges of G between X and Y. If G is a bipartite graph with vertex classes A and B, we often write $G = (A, B)$.

If G is a digraph, we write xy for an edge directed from x to y. Unless stated otherwise, when we refer to paths and cycles in digraphs, we mean directed paths and cycles, i.e. the edges on these paths/cycles are oriented consistently. Given two vertices x and y on a directed cycle C, we write xCy for the (directed) subpath of C from x to y. If x is a vertex of a digraph G, then $N^+_G(x)$ denotes the *outneighbourhood* of x, i.e. the set of all those vertices y for which $xy \in E(G)$. Similarly, $N^-_G(x)$ denotes the *inneighbourhood* of x, i.e. the set of all those vertices y for which $yx \in E(G)$. We write $d^+_G(x) := |N^+_G(x)|$ for the *outdegree* of x and $d^-_G(x) := |N^-_G(x)|$ for its *indegree*. We denote the *minimum outdegree* of G by $\delta^+(G) := \min_{x \in V(G)} d^+_G(x)$, the *minimum indegree* by $\delta^-(G) := \min_{x \in V(G)} d^-_G(x)$, the *minimum degree* by $\delta(G) := \min_{x \in V(G)}(d^+(x) + d^-(x))$ and the *maximum degree* by $\Delta(G) := \max_{x \in V(G)}(d^+(x) + d^-(x))$. The *minimum semidegree* of G is $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$. Whenever $X, Y \subseteq V(G)$, we write $e_G(X, Y)$ for the number of all those edges of G which have their initial vertex in X and their final vertex in Y. If $X \cap Y = \emptyset$, we denote by $G[X, Y]$ the bipartite subdigraph of G with vertex sets X and Y whose edges are all the edges of G directed from X to Y. In all these definitions we often omit the subscript G if the graph or digraph G is clear from the context. A subdigraph H of G is an *r-factor* of G if the outdegree and the indegree of every vertex of H is r. A *path system* is the union of vertex-disjoint directed paths.

Given a digraph R and a positive integer r, the *r-fold blow-up of R* is the digraph $R \times E_r$ obtained from R by replacing every vertex x of R by r vertices and replacing every edge xy of R by the oriented complete bipartite graph $K_{r,r}$ between the two sets of r vertices corresponding to x and y in which all the edges are oriented towards the r vertices corresponding to y. Now consider the case when $V_1, \dots, V_k$ is a partition of some set V of vertices and R is a digraph whose vertices are $V_1, \dots, V_k$. Then a *blow-up $\mathcal{B}(R)$ of R* is obtained from R by replacing every vertex V_i of R by the vertices in V_i and replacing every edge V_iV_j of R by a certain bipartite graph with vertex classes V_i and V_j in which all the edges are oriented towards the vertices in V_j. Usually, these bipartite graphs will be ε-regular or superregular (as defined in Section 4). If R is a directed cycle, say $R = C = V_1 \dots V_k$ and G is a digraph with $V(G) \subseteq V = V_1 \cup \dots \cup V_k$, we say that *(the edges of) G wind(s) around C* if for every edge xy of G there exists an index j such that $x \in V_j$ and $y \in V_{j+1}$. So if $V(G) = V$ then G winds around C if and only if G is a blow-up of C.

In order to simplify the presentation, we omit floors and ceilings and treat large numbers as integers whenever this does not affect the argument. The constants in the hierarchies used to state our results have to be chosen from right to left. More precisely, if we claim that a result holds whenever $0 < 1/n \ll a \ll b \ll c \leq 1$ (where n is the order of the graph or digraph), then this means that there are non-decreasing functions $f : (0, 1] \to (0, 1]$, $g : (0, 1] \to (0, 1]$ and $h : (0, 1] \to (0, 1]$ such that the result holds for all $0 < a, b, c \leq 1$ and all $n \in \mathbb{N}$ with $b \leq f(c)$, $a \leq g(b)$ and $1/n \leq h(a)$. We will not calculate these functions explicitly. Hierarchies with more constants are defined in a similar way. Moreover, we will often assume that certain numbers involving these constants, e.g. b/a^2 or an are integers. We will only make this assumption if our hierarchy guarantees that these numbers are sufficiently large since then by adjusting the constants slightly one can actually guarantee that these numbers are integers. However, all our results will hold if n is sufficiently large, i.e. we will make no divisibility assumptions on n. (Note that if we assume that an is an integer then this can be achieved by adjusting the constant a slightly.)

3.2. **Probabilistic estimates, derandomization and algorithmic aspects.** We will use the following standard Chernoff type bound (see e.g. Corollary 2.3 in [23] and Theorem 2.2 in [44]).

Proposition 3.1. *Suppose X has binomial distribution and $0 < a < 1$. Then*

$$\mathbb{P}(X \geq (1 + a)\mathbb{E}X) \leq \mathrm{e}^{-\frac{a^2}{3}\mathbb{E}X} \text{ and } \mathbb{P}(X \leq (1 - a)\mathbb{E}X) \leq \mathrm{e}^{-\frac{a^2}{3}\mathbb{E}X}.$$

To obtain an algorithmic version of Theorem 1.2, we need to 'derandomize' our applications of Proposition 3.1. This can be done via the well known 'method of conditional probabilities', which is based on an idea of Erdős and Selfridge, and which was further developed e.g. by Spencer as well as Raghavan. The following result of Srivastav and Stangier (Theorem 2.10 in [44]) is also based on this method. Given a probabilistic existence proof of some structure based on polynomially many applications of Proposition 3.1, it guarantees an algorithm which finds this structure.

Suppose we are given N independent 0/1 random variables $X_1, \dots, X_N$ where $\mathbb{P}(X_j = 1) = p$ and $\mathbb{P}(X_j = 0) = 1 - p$ for some rational $0 \leq p \leq 1$. Suppose that $1 \leq i \leq m$. Let $w_{ij} \in \{0, 1\}$. Denote by ϕ_i the random variables $\phi_i := \sum_{j=1}^{N} w_{ij} X_j$. Fix β_i with $0 < \beta_i < 1$. Now let E_i^+ denote the event that $\phi_i \geq (1 + \beta_i)\mathbb{E}[\phi_i]$ and let E_i^- denote the event that $\phi_i \leq (1 - \beta_i)\mathbb{E}[\phi_i]$. Let E_i be either E_i^+ or E_i^-. Suppose that

$$\sum_{i=1}^{m} \mathrm{e}^{-\beta_i^2 \mathbb{E}(\phi_i)/3} \leq 1/2. \tag{3.1}$$

Theorem 3.2. [44] *Let $E_1, \dots, E_m$ be events such that (3.1) holds. Then*

$$\mathbb{P}\left(\bigcap_{i=1}^{m} E_i\right) \geq 1/2$$

and a vector $x \in \bigcap_{i=1}^{m} E_i$ can be constructed in time $O(mN^2 \log(mN))$.

In general, it will usually be clear that the proofs can be translated into polynomial time algorithms. Where this is not obvious, we will add a corresponding remark. We make no attempt to prove an explicit bound on the time needed to find the Hamilton decomposition (beyond the fact that it is polynomial in n).

4. Regularity

4.1. **The Regularity Lemma.** We will use of a directed version of Szemerédi's regularity lemma. If $G = (A, B)$ is an undirected bipartite graph with vertex classes A and B, then the *density* of G is defined as

$$d(A, B) := \frac{e_G(A, B)}{|A||B|}.$$

For any $\varepsilon > 0$, we say that G is *ε-regular* if for any $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$ we have $|d(A', B') - d(A, B)| < \varepsilon$. We say that G is *$(\varepsilon, \geq d)$-regular* if it is ε-regular and has density d' for some $d' \geq d - \varepsilon$.

Given disjoint vertex sets X and Y in a digraph G, recall that $G[X, Y]$ denotes the bipartite subdigraph of G whose vertex classes are X and Y and whose edges are all the edges of G directed from X to Y. We often view $G[X, Y]$ as an undirected bipartite graph. In particular, we say $G[X, Y]$ is *ε-regular* or *$(\varepsilon, \geq d)$-regular* if this holds when $G[X, Y]$ is viewed as an undirected graph.

Next we state the degree form of the regularity lemma for digraphs. A regularity lemma for digraphs was proved by Alon and Shapira [3]. The degree form follows from this in the same way as the undirected version (see [34] for a sketch of the latter). An algorithmic version of the (undirected) regularity lemma was proved in [1]. An algorithmic version of the directed version can be proved in essentially the same way (see [12] for a sketch of the argument proving a similar statement).

Lemma 4.1 (Regularity Lemma for digraphs)**.** *For all $\varepsilon, M' > 0$ there exist M, n_0 such that if G is a digraph on $n \geq n_0$ vertices and $d \in [0, 1]$, then there exists a partition of $V(G)$ into $V_0, \dots, V_k$ and a spanning subdigraph G' of G satisfying the following conditions:*

(i) *$M' \leq k \leq M$.*
(ii) *$|V_0| \leq \varepsilon n$.*
(iii) *$|V_1| = \dots = |V_k| =: m$.*
(vi) *$d^+_{G'}(x) > d^+_G(x) - (d + \varepsilon)n$ for all vertices $x \in V(G)$.*
(v) *$d^-_{G'}(x) > d^-_G(x) - (d + \varepsilon)n$ for all vertices $x \in V(G)$.*
(vi) *For all $i = 1, \dots, k$ the digraph $G'[V_i]$ is empty.*
(vii) *For all $1 \leq i, j \leq k$ with $i \neq j$ the pair $G'[V_i, V_j]$ is either empty or ε-regular of density at least d. Moreover, if $G'[V_i, V_j]$ is nonempty then $G'[V_i, V_j] = G[V_i, V_j]$.*

We refer to V_0 as the *exceptional set* and to $V_1, \dots, V_k$ as *clusters*. $V_0, V_1, \dots, V_k$ as above is also called a *regularity partition* for G. Given a digraph G on n vertices, we form the *reduced digraph R of G with parameters ε, d and M'* by applying the regularity lemma to G with these parameters to obtain $V_0, \dots, V_k$. R is then the digraph whose vertices are the clusters $V_1, \dots, V_k$ and whose edges are those (ordered) pairs V_iV_j of clusters for which $G'[V_i, V_j]$ is non-empty.

Given $d \in [0,1]$ and a bipartite graph $G = (A, B)$, we say that G is (ε, d)-*superregular* if it is ε-regular and furthermore $d_G(a) \geq (d-\varepsilon)|B|$ for every $a \in A$ and $d_G(b) \geq (d-\varepsilon)|A|$ for every $b \in B$. (This is a slight variation of the standard definition of (ε, d)-superregularity where one requires $d_G(a) \geq d|B|$ and $d_G(b) \geq d|A|$.)

We say that a bipartite graph $G = (A, B)$ is $[\varepsilon, d]$-*superregular* if it is ε-regular and $d_G(a) = (d \pm \varepsilon)|B|$ for every $a \in A$ and $d_G(b) = (d \pm \varepsilon)|A|$ for every $b \in B$. So if G is $[\varepsilon, d]$-*superregular*, then it is (ε, d)-*superregular*. We say that G is $[\varepsilon, \geq d]$-*superregular* if it is $[\varepsilon, d']$-*superregular* for some $d' \geq d$. As for ε-regularity, these definitions extend naturally to bipartite graphs where all edges are oriented towards the same vertex class.

The following well known observation states that in an ε-regular bipartite graph almost all vertices have the expected degree and almost all pairs of vertices have the expected codegree (i.e. the expected number of common neighbours). Its proof follows immediately from the definition of regularity.

Proposition 4.2. *Suppose that $0 < \varepsilon \leq d \leq 1$. Let G be an ε-regular bipartite graph of density d with vertex classes A and B of size m. Then the following conditions hold.*

- *All but at most $2\varepsilon m$ vertices in A have degree $(d \pm \varepsilon)m$.*
- *All but at most $4\varepsilon m^2$ pairs $a \neq a'$ of distinct vertices in A satisfy $|N(a) \cap N(a')| = (d^2 \pm \varepsilon)m$.*
- *The vertices in B satisfy the analogues of these statements.*

The following simple observation states that the removal of a small number of edges and vertices from a bipartite graph does not affect its ε-regularity (and superregularity) too much.

Proposition 4.3. *Suppose that $0 < 1/m \ll \varepsilon \leq d' \leq d \ll 1$. Let G be a bipartite graph with vertex classes A and B of size m. Suppose that G' is obtained from G by removing at most $d'm$ vertices from each vertex class and at most $d'm$ edges incident to each vertex from G.*

(i) *If G is ε-regular of density at least d then G' is $2\sqrt{d'}$-regular of density at least $d - 2\sqrt{d'}$.*
(ii) *If G is (ε, d)-superregular then G' is $(2\sqrt{d'}, d)$-superregular.*
(iii) *If G is $[\varepsilon, d]$-superregular then G' is $[2\sqrt{d'}, d]$-superregular.*

Proof. Let us first prove (i). Let d^* denote the density of G. Let $A' \subseteq A$ and $B' \subseteq B$ denote the vertex classes of G'. Suppose that $S \subseteq A'$, $T \subseteq B'$ are such that

$|S| \ge 2\sqrt{d'}|A'|$ and $|T| \ge 2\sqrt{d'}|B'|$. So $|S|,|T| \ge \sqrt{d'}m \ge \varepsilon m$ and thus

$$\begin{aligned} e_{G'}(S,T) &\ge (d^*-\varepsilon)|S||T| - |S|d'm \ge (d^*-\varepsilon)|S||T| - |S|d'\cdot|T|/\sqrt{d'} \\ &\ge (d^*-2\sqrt{d'})|S||T|. \end{aligned}$$

Since clearly $e_{G'}(S,T) \le e_G(S,T) \le (d^*+\varepsilon)|S||T|$, (i) follows. To see (ii) and the lower bound on the vertex degrees for (iii), note that in G' the degrees of the vertices in A' are still at least $(d-\varepsilon)m - 2d'm \ge (d-\varepsilon-2d')|B'| \ge (d-\sqrt{d'})|B'|$. Similarly, the degrees in G' of the vertices in B' are still at least $(d-\sqrt{d'})|A'|$.

To see (iii), note that in G' the degrees of the vertices in A' are still at most $(d+\varepsilon)m \le (d+\varepsilon)|B'|/(1-d') \le (d+\sqrt{d'})|B'|$. Similarly, the degrees in G' of the vertices in B' are still at most $(d+\sqrt{d'})|A'|$. □

The following lemma is also well known in several variations.

Lemma 4.4. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll d, 1/\Delta \le 1$. Let G be a digraph on n vertices. Let $\mathcal{P}_0$ be a partition of $V(G)$ into k clusters $V'_1,\dots,V'_k$ and an exceptional set V'_0 such that $m' := |V'_1| = \dots = |V'_k|$. Let R be a digraph whose vertices are $V'_1,\dots,V'_k$ and such that whenever $V'_iV'_j \in E(R)$ the pair $G[V'_i,V'_j]$ is $(\varepsilon, \ge d)$-regular. Let H be a subdigraph of R of maximum degree Δ. Then there is a partition of $V(G)$ into $V_0,\dots,V_k$ such that the following holds:*

(i) *For each $i=1,\dots,k$, V_i is obtained from V'_i by moving exactly $\sqrt{\varepsilon}m'$ vertices into V'_0. V_0 is then the set consisting of V'_0 and these additional vertices.*
(ii) *Whenever $V'_iV'_j \in E(H)$, the pair $G[V_i,V_j]$ is $[2\varepsilon^{1/4}, \ge d]$-superregular.*

Proof. For each edge $V'_iV'_j$ of H, let d_{ij} denote the density of $G[V'_i,V'_j]$. So $d_{ij} \ge d-\varepsilon$. Call a vertex x in V'_i *bad* if at least one of the following two conditions hold:

- There is an edge $V'_iV'_j$ of H so that the degree of x in $G[V'_i,V'_j]$ is not $(1\pm 2\varepsilon)d_{ij}m'$.
- There is an edge $V'_jV'_i$ of H so that the degree of x in $G[V'_j,V'_i]$ is not $(1\pm 2\varepsilon)d_{ji}m'$.

Note that V'_i contains at most $2\Delta\varepsilon m' \le \sqrt{\varepsilon}m'$ bad vertices. Let V_i be obtained from V'_i by removing all bad vertices (and some additional ones if there are fewer than $\sqrt{\varepsilon}m'$ of these). Now suppose that $V'_iV'_j$ is an edge of H. Then Proposition 4.3(i) implies that $G[V_i,V_j]$ is $2\varepsilon^{1/4}$-regular. Together with the choice of V_i and V_j this implies that $G[V_i,V_j]$ is $[2\varepsilon^{1/4}, \ge d]$-superregular. □

The following result is proved in Section 3.1 of [16] by a simple application of the Max-Flow-Min-Cut theorem (similarly to that in Lemma 5.2 below). In particular, the subgraph of G guaranteed by this result can be found in time polynomial in m.

Lemma 4.5. *Let $0 < 1/m \ll \varepsilon \ll d \le 1$. Suppose that G is a bipartite graph with vertex classes U and V of size m and with minimum degree at least dm. Also suppose that $|d(A,B)-d| \le \varepsilon m$ for all $A \subseteq U$ and $B \subseteq V$ with $|A|,|B| \ge \varepsilon m$. Define $d' := (1-4\varepsilon)d$. Then G contains a spanning $d'm$-regular subgraph.*

Lemma 4.6. *Let $0 < 1/m \ll \varepsilon \ll d \ll 1$. Let $d' := (1-12\varepsilon)(d-\varepsilon)$. Suppose that G is a $[\varepsilon, d]$-superregular bipartite graph with vertex classes of size m. Then G contains a spanning $d'm$-regular subgraph which is also $[4\sqrt{\varepsilon}, d']$-superregular.*

Proof. Note that the assumptions imply that G satisfies the conditions of Lemma 4.5 with ε replaced by 3ε and d replaced by $d-\varepsilon$. So we can apply Lemma 4.5 to obtain a $d'm$-regular subgraph G' with $d' = (1-12\varepsilon)(d-\varepsilon)$. Note that G' is obtained from G by removing at most $3\varepsilon m$ edges at every vertex. Thus Proposition 4.3(i) (with 3ε playing the role of d') implies that G' is also $[4\sqrt{\varepsilon}, d']$-superregular. □

4.2. **Uniform refinements.** Let G be a digraph and let $\mathcal{P}$ be a partition of $V(G)$ into an exceptional set V_0 and clusters of equal size. Suppose that $\mathcal{P}'$ is another partition of $V(G)$ into an exceptional set V_0' and clusters of equal size. We say that $\mathcal{P}'$ is an *ℓ-refinement of $\mathcal{P}$* if $V_0 = V_0'$ and if the clusters in $\mathcal{P}'$ are obtained by partitioning each cluster in $\mathcal{P}$ into ℓ subclusters of equal size. (So if $\mathcal{P}$ contains k clusters then $\mathcal{P}'$ contains $k\ell$ clusters.) $\mathcal{P}'$ is an *ε-uniform ℓ-refinement* of $\mathcal{P}$ if it is an ℓ-refinement of $\mathcal{P}$ which satisfies the following condition:

(URef) Whenever x is a vertex of G, V is a cluster in $\mathcal{P}$ and $|N_G^+(x)\cap V| \geq \varepsilon|V|$ then $|N_G^+(x) \cap V'| = (1 \pm \varepsilon)|N_G^+(x) \cap V|/\ell$ for each cluster $V' \in \mathcal{P}'$ with $V' \subseteq V$. The inneighbourhoods of the vertices of G satisfy an analogous condition.

Lemma 4.7. *Suppose that $0 < 1/m \ll 1/k, \varepsilon \ll \varepsilon', d, 1/\ell \leq 1$ and that $m/\ell \in \mathbb{N}$. Suppose that G is a digraph on $n \leq 2km$ vertices and that $\mathcal{P}$ is a partition of $V(G)$ into an exceptional set V_0 and k clusters of size m. Then there exists an ε-uniform ℓ-refinement of $\mathcal{P}$. Moreover, any ε-uniform ℓ-refinement $\mathcal{P}'$ of $\mathcal{P}$ automatically satisfies the following conditions:*

(i) *Suppose that V, W are clusters in $\mathcal{P}$ and V', W' are clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. If $G[V,W]$ is $[\varepsilon, d']$-superregular for some $d' \geq d$ then $G[V', W']$ is $[\varepsilon', d']$-superregular.*

(ii) *Suppose that V, W are clusters in $\mathcal{P}$ and V', W' are clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. If $G[V,W]$ is $(\varepsilon, \geq d)$-regular then $G[V', W']$ is $(\varepsilon', \geq d)$-regular.*

Proof. To prove the existence of an ε-uniform ℓ-refinement of $\mathcal{P}$, let $\mathcal{P}^*$ be a partition obtained by splitting each cluster $V \in \mathcal{P}$ uniformly at random into ℓ subclusters. More precisely, the probability that a vertex $x \in V$ is assigned to the ith subcluster is $1/\ell$, independently of all other vertices. Consider a fixed vertex x of G and a cluster $V \in \mathcal{P}$ with $d^+ := |N_G^+(x) \cap V| \geq \varepsilon m$. Given a cluster $V' \in \mathcal{P}^*$ with $V' \subseteq V$, we say that x is *out-bad for V'* if the outdegree of x into V' is not $(1 \pm \varepsilon/2)d^+/\ell$. Then Proposition 3.1 implies that the probability that x is out-bad for V' is at most $2\mathrm{e}^{-\varepsilon^2 d^+/3\cdot 4\ell} \leq 2\mathrm{e}^{-\varepsilon^4 m}$. Since $\mathcal{P}^*$ contains $k\ell \leq n$ clusters, the probability that G contains some vertex which is out-bad for at least one cluster $V' \in \mathcal{P}^*$ is at most $n^2\mathrm{e}^{-\varepsilon^4 m} < 1/8$. We argue analogously for the inneighbourhoods of the vertices in G (by considering 'in-bad' vertices).

We now say that a cluster V' of $\mathcal{P}^*$ is *good* if $|V'| = (1 \pm \varepsilon^2/2)m/\ell$. A similar argument as above shows that the probability that $\mathcal{P}^*$ has a cluster which is not good is at most 1/4. So with probability at least 1/2, all clusters of $\mathcal{P}^*$ are good, and no vertices are out-bad or in-bad.

Now obtain $\mathcal{P}'$ from $\mathcal{P}^*$ as follows: for each cluster V of $\mathcal{P}$, equalize the sizes of the corresponding ℓ subclusters in $\mathcal{P}$ by moving at most $\varepsilon^2 m/2\ell$ vertices from one subcluster to another. So whenever x is a vertex of G, V is a cluster in $\mathcal{P}$ and $|N^+_G(x) \cap V| \geq \varepsilon|V|$, it follows that we have

$$|N^+_G(x) \cap V'| = (1 \pm \varepsilon/2)|N^+_G(x) \cap V|/\ell \pm \varepsilon^2 m/2\ell$$

for each cluster $V' \in \mathcal{P}'$ with $V' \subseteq V$. The inneighbourhoods of the vertices of G satisfy an analogous condition. So (URef) holds and so $\mathcal{P}'$ is an ε-uniform ℓ-refinement of $\mathcal{P}$.

To prove (i), suppose that $\mathcal{P}'$ is any ε-uniform ℓ-refinement of $\mathcal{P}$ and that $G[V,W]$ is $[\varepsilon, d']$-superregular for some $d' \geq d$ (where V and W are clusters in $\mathcal{P}$). Let V' and W' be clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. Then $G[V', W']$ is $\varepsilon\ell$-regular and thus ε'-regular. Consider any $x \in V'$ and let $d^+ := |N^+_G(x) \cap W|$. Thus $d^+ = (d' \pm \varepsilon)m$ since $G[V,W]$ is $[\varepsilon, d']$-superregular. Together with the ε-uniformity of $\mathcal{P}'$ this implies that $|N^+_G(x) \cap W'| = (1 \pm \varepsilon)d^+/\ell = (d' \pm \varepsilon')m/\ell$. The inneighbourhoods in V' of the vertices in W' satisfy the analogous property. Thus $G[V', W']$ is $[\varepsilon', d']$-superregular.

The proof of (ii) is almost the same. $\square$

Note that Theorem 3.2 (with $p = 1/\ell$) implies that the above applications of Proposition 3.1 can be derandomized to find $\mathcal{P}'$ in polynomial time.

4.3. **A sparse notion of ε-regularity.** We will also use a 'sparse' version of ε-(super)-regularity, which is defined below. In particular, this definition allows for $d < \varepsilon$. We will need this notion mainly in Section 8, where we will have to work with graphs for which we cannot guarantee $(\varepsilon, \geq d)$-regularity with $\varepsilon \leq d$. In general, one useful consequence of $(\varepsilon, \geq d)$-regularity is that sets of size between εm and $(1-\varepsilon)m$ expand robustly. With our sparse version, we will also be able to guarantee that even sets of size less than εm expand robustly. This will follow from condition (Reg2) below.

More precisely, let G be a bipartite graph with vertex classes U and V, both of size m. Given $0 < \varepsilon, d, c < 1$, we say that G is (ε, d, c)-*regular* if the following conditions are satisfied:

(Reg1) Whenever $A \subseteq U$ and $B \subseteq V$ are sets of size at least εm, then $d(A,B) = (1 \pm \varepsilon)d$.

(Reg2) For all $u, u' \in U$ we have $|N(u) \cap N(u')| \leq c^2 m$. Similarly, for all $v, v' \in V$ we have $|N(v) \cap N(v')| \leq c^2 m$.

(Reg3) $\Delta(G) \leq cm$.

We say that G is (ε, d, d^*, c)-*superregular* if it is (ε, d, c)-regular and in addition the following condition holds:

(Reg4) $\delta(G) \geq d^* m$.

The next result gives an analogue of Proposition 4.3 for the above notion of (super)-regularity.

Proposition 4.8. *Suppose that $0 < 1/m \ll d^*, d, \varepsilon, c \ll 1$. Let G be a bipartite graph with vertex classes U and V of size m. Suppose that G' is obtained from G by removing at most $\varepsilon^2 dm$ edges incident to each vertex from G.*

(i) *If G is (ε, d, c)-regular then G' is $(2\varepsilon, d, c)$-regular.*
(ii) *If G is (ε, d, d^*, c)-superregular then G' is $(2\varepsilon, d, d^* - \varepsilon^2 d, c)$-superregular.*

Proof. Let us first prove (i). Clearly G' still satisfies (Reg2) and (Reg3). So we only need to check that it also satisfies (Reg1). Suppose that $S \subseteq U$, $T \subseteq V$ are such that $|S|, |T| \geq \varepsilon m$. Then

$$e_{G'}(S,T) \geq (1-\varepsilon)d|S||T| - |S|\varepsilon^2 dm \geq (1-\varepsilon)d|S||T| - |S|\varepsilon^2 d \cdot |T|/\varepsilon = (1-2\varepsilon)d|S||T|.$$

Since clearly $e_{G'}(S,T) \leq e_G(S,T) \leq (1+\varepsilon)d|S||T|$, (i) follows. To see (ii) note that the degrees in G' are still at least $d^* m - \varepsilon^2 dm$. □

We will construct sparse (ε, d, c)-regular graphs in the proof of Lemma 4.10. To verify (Reg1) in the proof of Lemma 4.10, we will use a variant of the well known characterization in terms of codegrees of pairs of vertices which was proved as Lemma 3.2 in [1] (the version in [1] gives more precise bounds but the statement is not suitable for sparse regularity).

Suppose that $G = (U, V)$ is a bipartite graph with vertex classes U and V of size m. We say that G is *$\{\varepsilon, d\}$-regular* if for all $A \subseteq U$ and $B \subseteq V$ with $|A|, |B| \geq \varepsilon m$ we have $d(A,B) = (1 \pm \varepsilon)dm$. (So (Reg1) is equivalent to saying that G is $\{\varepsilon, d\}$-regular.) Note that this notion allows for $d < \varepsilon$. Moreover, it is stronger than ε-regularity in the sense that if $\varepsilon < d \ll 1$ then every $\{\varepsilon, d\}$-regular pair is also ε-regular of density $(1 \pm \varepsilon)d$.

Call a pair of distinct vertices in V *bad* if the number of common neighbours in U is at least $(1+\varepsilon)d^2 m$.

Lemma 4.9. *Suppose that $1/m \ll \varepsilon, d \leq 1/C \leq 1$ and $\varepsilon \ll 1/C$. Let $G = (U, V)$ be a bipartite graph with vertex classes U and V of size m. Suppose that all but at most εm vertices in V have degree at least $(1-\varepsilon)dm$ and for all pairs of distinct vertices in V the number of common neighbours is at most $Cd^2 m$. Suppose also that the number of bad pairs of distinct vertices in V is at most εm^2. Then G is $\{\varepsilon^{1/6}, d\}$-regular.*

Proof. The proof follows the argument in [1]. Let $\varepsilon_0 := \varepsilon^{1/6}$. It is easy to see that it suffices to check that $d(X,Y) = (1 \pm \varepsilon_0)d$ for all pairs $X \subseteq U$ and $Y \subseteq V$ with $|X| = |Y| = \varepsilon_0 m$. For any pair of vertices $y_1, y_2 \in Y$, let $\sigma(y_1, y_2) := |N(y_1) \cap N(y_2)| - d^2 m$. Then we always have $\sigma(y_1, y_2) \leq Cd^2 m$ and can improve this to $\sigma(y_1, y_2) \leq \varepsilon d^2 m$ if the pair y_1, y_2 is not bad. Also define

$$\sigma(Y) := \frac{1}{|Y|^2} \sum_{y_1, y_2 \in Y,\ y_1 \neq y_2} \sigma(y_1, y_2).$$

Then our assumption on $|Y|$ and on the number of bad pairs implies that

$$\sigma(Y) \leq \frac{1}{(\varepsilon_0 m)^2}\left((\varepsilon m^2)Cd^2 m + (\varepsilon_0 m)^2 \varepsilon d^2 m\right) = C\varepsilon_0^4 d^2 m + \varepsilon d^2 m \leq \varepsilon_0^3 d^2 m/3. \tag{4.1}$$

We claim that

$$\sum_{x\in X} (|N(x)\cap Y| - d|Y|)^2 \le \varepsilon_0^3 d^2 m|Y|^2. \tag{4.2}$$

To prove the claim, we use that the left hand side (as shown in [1]) is at most

$$e(U,Y) + \sigma(Y)|Y|^2 + 2d^2|Y|^2 m - 2e(U,Y)d|Y|.$$

But our assumption on the vertex degrees implies that

$$e(U,Y) \ge (|Y| - \varepsilon m)(1-\varepsilon)dm = |Y|(1-\varepsilon/\varepsilon_0)(1-\varepsilon)dm \ge |Y|(1-\varepsilon_0^3/6)dm$$

and so

$$2d^2|Y|^2 m - 2e(U,Y)d|Y| \le \varepsilon_0^3 d^2|Y|^2 m/3.$$

So together with (4.1) this implies that the left hand side of (4.2) is at most

$$e(U,Y) + \frac{2}{3}\varepsilon_0^3 d^2 m|Y|^2 \le |Y|m + \frac{2}{3}\varepsilon_0^3 d^2 m|Y|^2 \le \varepsilon_0^3 d^2 m|Y|^2.$$

This proves the claim. On the other hand, the Cauchy-Schwarz inequality implies that

$$\sum_{x\in X} (|N(x)\cap Y| - d|Y|)^2 \ge \frac{1}{|X|}\left(\left(\sum_{x\in X}|N(x)\cap Y|\right) - d|X||Y|\right)^2.$$

So together with (4.2), this implies that

$$\left(\left(\sum_{x\in X}|N(x)\cap Y|\right) - d|X||Y|\right)^2 \le |X|\left(\varepsilon_0^3 d^2 m|Y|^2\right).$$

Thus dividing both sides by $|X|^2|Y|^2$ yields

$$|d(X,Y) - d|^2 \le \frac{1}{|X|}\left(\varepsilon_0^3 d^2 m\right) = \varepsilon_0^2 d^2,$$

as required. □

The first part of the following lemma implies that inside an ε-regular pair we can find sparse subgraphs which satisfy (Reg1)–(Reg3) with good bounds on the parameters. Assertion (iii) will only be used in [35, 42].

Lemma 4.10. *Suppose that $0 < 1/m \ll \varepsilon, d' \le d \le 1$ and $\varepsilon \ll d$.*

- (i) *If G is an ε-regular bipartite graph of density d with vertex classes of size m, then it contains an $(\varepsilon^{1/12}, d', 3d'/2d)$-regular spanning subgraph.*
- (ii) *If G is an (ε, d)-superregular bipartite graph with vertex classes of size m, then it contains an $(\varepsilon^{1/12}, d', d'/2, 3d'/2d)$-superregular spanning subgraph.*
- (iii) *If $\varepsilon \ll d'$ and G is an ε-regular bipartite graph of density d with vertex classes of size m, then it contains an $\{\varepsilon^{1/12}, d'\}$-regular spanning subgraph J. Moreover, if $x \in V(G)$ satisfies $d_G(x) = (d \pm \varepsilon)m$, then $d_J(x) = (d' \pm \sqrt{\varepsilon})m$.*
- (iv) *If $\varepsilon \ll d'$ and G is an $[\varepsilon, d]$-superregular bipartite graph with vertex classes of size m, then it contains an $[\varepsilon^{1/12}, d']$-superregular spanning subgraph.*

Proof. We only prove (i). Since a $(\varepsilon^{1/12}, d', 3d'/2d)$-regular pair is $\{\varepsilon^{1/12}, d'\}$-regular, (iii) follows from (i) (and the 'moreover' part follows from the proof of (i)). The argument for (ii) and (iv) is similar to the proof of (i). So suppose that G is ε-regular of density d with vertex classes U and V of size m. Let G' be the spanning subgraph obtained from G by picking every edge of G with probability $p := d'/d$, independently from all other edges. Consider any vertex $v \in V$ with $d_G(v) = (d \pm \varepsilon)m$. Then the expected degree of v in G' is $p(d \pm \varepsilon)m = (1 \pm \sqrt{\varepsilon}/2)d'm$. So Proposition 3.1 implies that

$$\begin{aligned}\mathbb{P}\left(d_{G'}(v) \neq (1 \pm \sqrt{\varepsilon})d'm\right) &\leq \mathbb{P}\left(|d_{G'}(v) - \mathbb{E}(d_{G'}(v))| \geq \frac{\sqrt{\varepsilon}}{3}\mathbb{E}(d_{G'}(v))\right)\\ &\leq 2\mathrm{e}^{-\varepsilon\mathbb{E}(d_{G'}(v))/27} \leq 2\mathrm{e}^{-\varepsilon d'm/28}.\end{aligned}$$

Similarly, consider any vertex $x \in U \cup V$ (with no restriction on its degree $d_G(x)$). If $d_G(x) \leq 3d'm/2d$, then clearly $d_{G'}(x) \leq 3d'm/2d$. So suppose that $d_G(x) \geq 3d'm/2d$. Then $3(d')^2m/2d^2 = p \cdot 3d'm/2d \leq \mathbb{E}(d_{G'}(x)) \leq pm = d'm/d$ and so

$$\begin{aligned}\mathbb{P}\left(d_{G'}(x) \geq \frac{3d'm}{2d}\right) &\leq \mathbb{P}\left(d_{G'}(x) \geq \frac{3}{2}\mathbb{E}(d_{G'}(x))\right)\\ &\leq \mathrm{e}^{-\mathbb{E}(d_{G'}(x))/12} \leq \mathrm{e}^{-(d')^2m/8d^2}.\end{aligned}$$

For the remainder of the proof, we let the *codegree* $d_G(x, x')$ of a pair x, x' of vertices in G be the number of common neighbours of x and x'. Consider any pair $v, v' \in V$ of distinct vertices with codegree $d_G(v, v') = (d^2 \pm \varepsilon)m$. Then the expected codegree of v, v' in G' is $\mathbb{E}(d_{G'}(v, v')) = p^2(d^2 \pm \varepsilon)m = (1 \pm \sqrt{\varepsilon}/2)(d')^2m$. So Proposition 3.1 implies that

$$\begin{aligned}\mathbb{P}\left(d_{G'}(v, v') \geq (1 \pm \sqrt{\varepsilon})(d')^2m\right) &\leq \mathbb{P}\left(d_{G'}(v, v') \geq \left(1 + \sqrt{\varepsilon}/3\right)\mathbb{E}(d_{G'}(v, v'))\right)\\ &\leq \mathrm{e}^{-\varepsilon\mathbb{E}(d_{G'}(v,v'))/27} \leq \mathrm{e}^{-\varepsilon(d')^2m/28}.\end{aligned}$$

Similarly, consider any pair $x \neq x'$ of vertices in G (with no restriction on the codegree $d_G(x, x')$). If $d_G(x, x') \leq 3(d')^2m/2d^2$, then clearly $d_{G'}(x, x') \leq 3(d')^2m/2d^2$. So suppose that $d_G(x, x') \geq 3(d')^2m/2d^2$. Then $3(d')^4m/2d^4 \leq \mathbb{E}(d_{G'}(x, x')) \leq p^2m = (d')^2m/d^2$ and so

$$\begin{aligned}\mathbb{P}\left(d_{G'}(x, x') \geq \frac{3(d')^2m}{2d^2}\right) &\leq \mathbb{P}\left(d_{G'}(x, x') \geq \frac{3}{2}\mathbb{E}(d_{G'}(x, x'))\right)\\ &\leq \mathrm{e}^{-\mathbb{E}(d_{G'}(x,x'))/12} \leq \mathrm{e}^{-(d')^4m/8d^4}.\end{aligned}$$

Proposition 4.2 implies that V contains at most $2\varepsilon m$ vertices whose degree in G is not $(d \pm \varepsilon)m$ as well as at most $4\varepsilon m^2$ pairs of distinct vertices whose codegree in G is not $(d^2 \pm \varepsilon)m$. Thus a union bound implies that with probability at least

$$1 - 4m\mathrm{e}^{-\varepsilon d'm/28} + 2m\mathrm{e}^{-(d')^2m/8d^2} + m^2\mathrm{e}^{-\varepsilon(d')^2m/28} + m^2\mathrm{e}^{-(d')^4m/8d^4} \geq 1/2$$

all of the following properties are satisfied:

- All but at most $2\varepsilon m$ vertices $v \in V$ satisfy $d_{G'}(v) = (1 \pm \sqrt{\varepsilon})d'm$.
- All but at most $4\varepsilon m^2$ pairs $v \neq v'$ of vertices in V satisfy $d_{G'}(v, v') = |N_{G'}(v) \cap N_{G'}(v')| \leq (1 + \sqrt{\varepsilon})(d')^2m$.
- All pairs $v \neq v'$ of vertices in V satisfy $d_{G'}(v, v') \leq 3(d')^2m/2d^2$.

- $\Delta(G') \le 3d'm/2d$.

Thus we can choose G' to satisfy (Reg2) and (Reg3). Moreover, Lemma 4.9 (applied with $\sqrt{\varepsilon}$, d', $3/2d^2$ playing the roles of ε, d, C) together with the first two properties implies that G' also satisfies (Reg1) with ε replaced by $\varepsilon^{1/12}$. □

Theorem 3.2 implies that the applications of Proposition 3.1 in the proof of Lemma 4.10 can be derandomized to find the spanning subgraphs guaranteed by the lemma in polynomial time. Note that instead of using Lemma 4.9 to check (Reg1) in our proof of Lemma 4.10, one could have checked (Reg1) directly. But this would have involved a union bound over exponentially many sets. So we would not have been able to apply Theorem 3.2.

Let G be a $(\varepsilon, d, d^*, d/\mu)$-superregular bipartite graph with vertex classes $U = \{u_1, \dots, u_m\}$ and $V = \{v_1, \dots, v_m\}$. The following lemma implies that the digraph obtained from G by orienting all the edges from U to V and identifying u_i and v_i for all $i = 1, \dots, m$ is a robust (ν, τ)-outexpander of minimum semidegree at least d^*m. Together with Theorem 6.2 below this will imply that this 'contracted' digraph has a Hamilton cycle, which will in turn be used in the proof of Lemma 6.5.

Lemma 4.11. *Let $0 < 1/m \ll \nu \ll \tau \ll d \le \varepsilon \ll \mu, \zeta \le 1/2$ and let G be a $(\varepsilon, d, \zeta d, d/\mu)$-superregular bipartite graph with vertex classes U and V of size m. Let $A \subseteq U$ be such that $\tau m \le |A| \le (1-\tau)m$. Let $B \subseteq V$ be the set of all those vertices in V which have at least νm neighbours in A. Then $|B| \ge |A| + \nu m$.*

Proof. Let us first prove the following claim.

Claim. *Let $U' \subseteq U$ be such that $|U'| \ge \tau m/2$ and let $RN(U')$ be the set of all those vertices in V which have at least νm neighbours in U'. Then $|RN(U')| \ge \min\{10\varepsilon m, |U'|/\sqrt{\varepsilon}\}$. Similarly, let $V' \subseteq V$ be such that $|V'| \ge \tau m/2$ and let $RN(V')$ be the set of all those vertices in U which have at least νm neighbours in V'. Then $|RN(V')| \ge \min\{10\varepsilon m, |V'|/\sqrt{\varepsilon}\}$.*

We only prove the first part of the claim. The argument for the second part is identical. Suppose that $|RN(U')| \le 10\varepsilon m$. Given $V', V'' \subseteq V$, let $f(V', V'')$ denote the number of paths of length two which have their midpoint in U', one endpoint in V' and the other endpoint in V''. Since $\delta(G) \ge \zeta dm$ by (Reg4) we have that

$$f(V,V) \ge |U'|\binom{\zeta dm}{2} \ge |U'|\frac{\zeta^2 d^2 m^2}{3}.$$

On the other hand, $f(V \setminus RN(U'), V) \le \nu |V \setminus RN(U')| m^2$ since every vertex in $V \setminus RN(U')$ has at most νm neighbours in U'. Thus

$$\begin{aligned} f(V,V) &= f(RN(U'), RN(U')) + f(V \setminus RN(U'), V) \\ &\le \sum_{v,v' \in RN(U'),\ v \neq v'} |N(v) \cap N(v')| + \nu |V \setminus RN(U')| m^2 \\ &\overset{\text{(Reg2)}}{\le} |RN(U')|^2 \frac{d^2 m}{\mu^2} + \nu m^3 \le |RN(U')| \frac{10\varepsilon d^2 m^2}{\mu^2} + \nu m^3. \end{aligned}$$

(To obtain the final inequality we use that $|RN(U')| \leq 10\varepsilon m$.) Altogether this implies that

$$|RN(U')| \geq \frac{\mu^2}{10\varepsilon d^2 m^2}\left(|U'|\frac{\zeta^2 d^2 m^2}{3} - \nu m^3\right)$$
$$\geq \frac{\mu^2}{10\varepsilon d^2 m^2} \cdot |U'|\frac{\zeta^2 d^2 m^2}{4} = \frac{\mu^2\zeta^2}{40\varepsilon}|U'| \geq \frac{|U'|}{\sqrt{\varepsilon}},$$

which proves the claim.

In order to prove the lemma, we distinguish several cases according to the size of A. Suppose first that $\tau m \leq |A| \leq \varepsilon m$. Then the claim implies that $B = RN(A)$ has size at least $\min\{10\varepsilon m, |A|/\sqrt{\varepsilon}\} \geq |A| + \nu m$, as required.

Suppose next that $\varepsilon m \leq |A| \leq (1-2\varepsilon)m$. Since

$$e(A, V \setminus B) \leq \nu m|V \setminus B| < (1-\varepsilon)d|A||V \setminus B|,$$

together with (Reg1) this implies that $|V \setminus B| < \varepsilon m$. Thus $|B| \geq (1-\varepsilon)m \geq |A| + \nu m$.

Finally, consider the case when $(1-2\varepsilon)m \leq |A| \leq (1-\tau)m$. Suppose for a contradiction that $|B| < |A| + \nu m$. From the previous case it follows that $|B| \geq (1-2\varepsilon)m + \nu m$. Let $V' := V \setminus B$. Then

$$\tau m/2 \leq m - |A| - \nu m < m - |B| = |V'| \leq 2\varepsilon m. \tag{4.3}$$

Let $A' := A \cap RN(V')$. Since every vertex in A' has at least νm neighbours in V', but every vertex in V' has less than νm neighbours in $A \supseteq A'$, it follows that $|A'|\nu m \leq e(A', V') \leq |V'|\nu m$. Thus $|A'| \leq |V'|$ and so

$$|RN(V')| \leq |U \setminus A| + |A'| \leq m - |A| + |V'| \leq 2|V'| + \nu m < 3|V'| \leq 6\varepsilon m.$$

(Here we used (4.3) in the final three inequalities.) On the other hand, the claim implies that $|RN(V')| \geq \min\{10\varepsilon m, |V'|/\sqrt{\varepsilon}\}$, a contradiction. □

We will also use the above lemma to show that bipartite graphs satisfying (Reg1)–(Reg3) have large matchings.

Lemma 4.12. *Suppose that $0 < 1/m \ll d' \ll 1/k \ll \varepsilon \ll d \ll \zeta \leq 1/2$ and let G be a bipartite graph with vertex classes U and V of size m.*

(i) *If G is $(\varepsilon, d'/k, d'/dk)$-regular, then it contains a matching of size at least $(1-\varepsilon)m$.*
(ii) *If G is $(\varepsilon, d', \zeta d', d'/d)$-superregular, then it has a perfect matching.*

Since maximum matchings can be found in polynomial time, this is also the case for the matchings guaranteed by the lemma.

Proof. To prove (i), note that (Reg1) implies that $|N(A)| \geq (1-\varepsilon)m$ for every set $A \subseteq U$ with $|A| \geq \varepsilon m$. Together with the defect version of Hall's theorem this implies that G has a matching of size at least $(1-\varepsilon)m$.

To prove (ii), we use Hall's theorem. Lemma 4.11 implies that Hall's condition holds for all sets $A \subseteq U$ with $\tau m \leq |A| \leq (1-\tau)m$ (where we apply the lemma with d', d playing the roles of d, μ and with some ν, τ satisfying $1/m \ll \nu \ll \tau \ll d'$). (Reg4) implies that G has minimum degree at least $\zeta d' m \geq \tau m$, so Hall's condition also holds for sets $A \subseteq U$ with $|A| \leq \tau m$ or $|A| \geq (1-\tau)m$. □

4.4. **The Blow-up Lemma.** We will use the blow-up lemma of Komlós, Sárközy and Szemerédi [27] (see [28] for their proof of an algorithmic version). Roughly speaking, it states that superregular pairs behave like complete bipartite graphs with respect to embedding subgraphs of bounded degree.

Lemma 4.13 (Blow-up Lemma)**.** *Suppose that* $0 < 1/m \ll \varepsilon \ll 1/k, d, 1/\Delta \leq 1$. *Let* R *be a graph with vertex set* $\{1,\dots,k\}$. *Let* $V_1,\dots,V_k$ *be pairwise disjoint sets of vertices, each of size* m. *Let* $R(m)$ *be the graph on* $V_1 \cup \dots \cup V_k$ *which is obtained from* R *by replacing each edge* ij *of* R *with a complete bipartite graph* $K_{m,m}$ *between* V_i *and* V_j. *Let* G *be a graph on* $V_1 \cup \dots \cup V_k$ *which is obtained from* R *by replacing each edge* ij *of* R *with some* (ε, d)*-superregular pair* $G[V_i, V_j]$. *If a graph* H *with maximum degree* $\Delta(H) \leq \Delta$ *is embeddable into* $R(m)$ *then it is already embeddable into* G.

The following proposition is a very special case of the blow-up lemma. It is also easy to prove it in the same way as Lemma 4.12(ii).

Proposition 4.14. *Suppose that* $0 < 1/m \ll \varepsilon \ll d \leq 1$. *Suppose that* G *is a* (ε, d)*-superregular bipartite graph with vertex classes of size* m. *Then* G *contains a perfect matching.*

The following consequence of the blow-up lemma states that we can link up arbitrary sets of vertices which are joined by a 'blown-up' path.

Corollary 4.15. *Suppose that* $0 < 1/m \ll \varepsilon \ll d \leq 1$ *and that* $q \geq 4$. *Let* $V_1,\dots,V_q$ *be pairwise disjoint sets of vertices, each of size* m. *Let* G *be a graph on* $V_1 \cup \dots \cup V_q$ *such that* $G[V_i, V_{i+1}]$ *is* (ε, d)*-superregular for each* $i = 1,\dots,q-1$. *Let* $a_1,\dots,a_m$ *be an arbitrary enumeration of the vertices in* V_1 *and let* $b_1,\dots,b_m$ *be an arbitrary enumeration of the vertices in* V_q. *Then* G *contains a set of* m *vertex-disjoint paths connecting* a_i *to* b_i *for every* i.

Proof. First suppose that $q = 4$. Consider the graph G' which is obtained from $\bigcup_{i=1}^{q-1} G[V_i, V_{i+1}]$ by identifying the vertices a_i and b_i. Thus G' can be viewed as the blow-up of a triangle. The blow-up lemma implies that G' contains a set of disjoint triangles covering all vertices of G' (and where each triangle has a vertex in each V_i). This corresponds to the desired set of paths connecting the a_i and b_i for all $i = 1,\dots,m$.

If $q > 4$, we simply apply Proposition 4.14 $q-4$ times to get m vertex-disjoint paths joining all vertices in V_4 to V_q. Then we proceed as in the case when $q = 4$. □

5. Robust Outexpanders

The next result (Lemma 14 from [37]) implies that the property of a digraph G being a robust outexpander is 'inherited' by the reduced digraph of G. For this, we need that G is a robust outexpander, rather than just an outexpander.

Lemma 5.1. *Suppose that $0 < 1/n \ll \varepsilon \ll d \ll \alpha, \nu, \tau < 1$ and $M'/n \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq \alpha n$ 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 $\delta^0(R) \geq \alpha|R|/2$ and R is a robust $(\nu/2, 2\tau)$-outexpander.*

The following result shows that in a robust outexpander, we can guarantee a spanning subdigraph with a given degree sequence (as long as the required degrees are not too large and do not deviate too much from each other). We will state a more general version of this lemma for multidigraphs, which will be used in [42]. In the current paper, we will only use Lemma 5.2 in the case when $Q = G$. If x is a vertex of a multidigraph Q, we write $d^+_Q(x)$ for the number of edges in Q whose initial vertex is x and $d^-_Q(x)$ for the number of edges in Q whose final vertex is x.

Lemma 5.2. *Let $q \in \mathbb{N}$. Suppose that $0 < 1/n \ll \varepsilon \ll \nu \leq \tau \ll \alpha < 1$ and that $1/n \ll \xi \leq q\nu^2/3$. Let G be a digraph on n vertices with $\delta^0(G) \geq \alpha n$ which is a robust (ν, τ)-outexpander. Suppose that Q is a multidigraph on $V(G)$ such that whenever $xy \in E(G)$ then Q contains at least q edges from x to y. For every vertex x of G, let $n^+_x, n^-_x \in \mathbb{N}$ be such that $(1-\varepsilon)\xi n \leq n^+_x, n^-_x \leq (1+\varepsilon)\xi n$ and such that $\sum_{x\in V(G)} n^+_x = \sum_{x\in V(G)} n^-_x$. Then Q contains a spanning submultidigraph Q' such that $d^+_{Q'}(x) = n^+_x$ and $d^-_{Q'}(x) = n^-_x$ for every $x \in V(G) = V(Q)$.*

Proof. Our aim is to apply the Max-Flow-Min-Cut theorem. So let H be the (unoriented) bipartite multigraph whose vertex classes A and B are both copies of $V(G)$ and in which $a \in A$ is joined to $b \in B$ once for every (directed) edge ab of Q. Give every edge of H capacity 1. Add a source s^* which is joined to every vertex $a \in A$ with an edge of capacity n^+_a. Add a sink t^* which is joined to every vertex $b \in B$ with an edge of capacity n^-_b. Let $r := \sum_{x\in V(G)} n^+_x/n = \sum_{x\in V(G)} n^-_x/n$. Note that an integer-valued rn-flow corresponds to the desired spanning submultigraph Q' of Q. Thus by the Max-Flow-Min-Cut theorem it suffices to show that every cut has capacity at least rn.

So consider a minimal cut $\mathcal{C}$. Let S be the set of all those vertices $a \in A$ for which $s^*a \notin \mathcal{C}$. Similarly, let T be the set of all those vertices $b \in B$ for which $bt^* \notin \mathcal{C}$. Let $S' := A \setminus S$ and $T' := B \setminus T$. Thus the capacity of $\mathcal{C}$ is

$$c := \sum_{s\in S'} n^+_s + e_B(S,T) + \sum_{t\in T'} n^-_t.$$

Suppose first that $|T'| \geq |S| + \nu n/2$. But then

$$c \geq \sum_{s\in S'} n^+_s + \sum_{t\in T'} n^-_t \geq (1-\varepsilon)\xi n(|S'|+|T'|) \geq (1-\varepsilon)\xi n(n+\nu n/2) \geq (1+\varepsilon)\xi n^2 \geq rn,$$

as required.

So suppose next that $|T'| \leq |S| + \nu n/2$ and $\tau n < |S| < (1-\tau)n$. Then at least $\nu n/2$ vertices from T lie in $RN^+_{\nu,G}(S)$ and each such vertex receives at least νn edges from S (in the digraph G). Thus

$$c \geq e_B(S,T) \geq q \cdot e_G(S,T) \geq q\nu^2 n^2/2 \geq (1+\varepsilon)\xi n^2 \geq rn,$$

as required.

Now suppose that $|T'| \leq |S| + \nu n/2$ and $|S| \leq \tau n$. Then in G every vertex in S sends at least $\alpha n/2$ edges to T as $\delta^0(G) \geq \alpha n$. Thus in Q every vertex in S sends at least $q\alpha n/2$ edges to T and so

$$c \geq \sum_{s \in S'} n_s^+ + |S| \cdot q\alpha n/2 \geq \sum_{s \in S'} n_s^+ + |S|(1+\varepsilon)\xi n \geq rn,$$

as required.

Finally, suppose that $|S| \geq (1-\tau)n$. Then in G every vertex in T receives at least $\alpha n/2$ edges from S. Thus in Q every vertex in T receives at least $q\alpha n/2$ edges from S and so

$$c \geq \sum_{t \in T'} n_t^- + |T| \cdot q\alpha n/2 \geq \sum_{t \in T'} n_t^- + |T|(1+\varepsilon)\xi n \geq rn,$$

as required. □

Recall from Section 3 that the r-fold blow-up of a digraph G is obtained from G by replacing every vertex x of G by r vertices and replacing every edge xy of G by the complete bipartite graph $K_{r,r}$ between the two sets of r vertices corresponding to x and y such that all the edges of $K_{r,r}$ are oriented towards the r vertices corresponding to y.

Lemma 5.3. *Let $r \geq 3$ and let G be a robust (ν, τ)-outexpander with $0 < 3\nu \leq \tau < 1$. Let G' be the r-fold blow-up of G. Then G' is a robust $(\nu^3, 2\tau)$-outexpander.*

Proof. Let $n := |G|$. So $|G'| = rn$. Call two vertices in G' *friends* if they correspond to the same vertex of G. (In particular, every vertex is a friend of itself.) Consider any $S' \subseteq V(G')$ with $2\tau rn \leq |S'| \leq (1-2\tau)rn$. Call a vertex $x \in S'$ *bad* if S' contains at most $\nu^2 r$ friends of x. So if $\nu^2 r < 1$ then no vertex in S' is bad. Let b denote the number of bad vertices in S'. Then S' contains a set S^* of at least $b/\nu^2 r$ bad vertices corresponding to different vertices of G. (So no two vertices in S^* are friends.) But every $x \in S^*$ has at least $r - 1 - \nu^2 r \geq r/2$ friends outside S'. Thus

$$\frac{b}{\nu^2 r} \cdot \frac{r}{2} \leq |S^*| \cdot \frac{r}{2} \leq |G'| = rn$$

and so $b \leq 2\nu^2 rn$. Let $S'' \subseteq S'$ be the set of all those vertices in S' which are not bad and let S be the set of all those vertices x in G for which S'' contains a copy of x. Thus

$$|S| \geq |S''|/r = (|S'| - b)/r \geq |S'|/2r \geq \tau n. \tag{5.1}$$

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

(i) Either $|RN^+_{\nu,G}(S)| \geq |S| + \nu n$;
(ii) or $|S| \geq (1-\tau)n$, in which case (considering a subset of S of size $(1-\tau)n$) we have $|RN^+_{\nu,G}(S)| \geq (1 - \tau + \nu)n$.

Note that if a vertex x of G belongs to $RN^+_{\nu,G}(S)$, then any copy x' of x in G' has at least $\nu^2 r \cdot \nu n = \nu^3 |G'|$ inneighbours in S'' (since no vertex in S'' is bad) and so

$x' \in RN^+_{\nu^3,G'}(S')$. It follows that $|RN^+_{\nu^3,G'}(S')| \geq r|RN^+_{\nu,G}(S)|$. Thus, in case (i) we have

$$|RN^+_{\nu^3,G'}(S')| \geq r|RN^+_{\nu,G}(S)| \geq r|S| + r\nu n \overset{(5.1)}{\geq} |S'| - b + r\nu n \geq |S'| + \nu^3 rn,$$

while in case (ii) we have

$$|RN^+_{\nu^3,G'}(S')| \geq r|RN^+_{\nu,G}(S)| \geq (1-\tau)rn + \nu rn \geq (1-\tau)rn \geq |S'| + \nu^3 rn,$$

as required. □

6. Tools for finding Hamilton cycles

The following well known observation states that every regular multidigraph G has a 1-factorization, i.e. the edges of G can be decomposed into edge-disjoint 1-factors. (In a multidigraph G we allow multiple edges between any two vertices. G is r-regular if every vertex sends out r edges and receives r edges.)

Proposition 6.1. *Let G be a regular multidigraph. Then G has a 1-factorization.*

Proof. Consider an auxiliary bipartite multigraph G' whose vertex classes A and B are copies of the vertex set $V(G)$ of G and in which the number of (undirected) edges between $a \in A$ and $b \in B$ equals the number of (directed) edges from a to b in G. Then G' is regular and so Hall's theorem (which holds for bipartite multigraphs as well) implies that G' has a decomposition into edge-disjoint perfect matchings. But every perfect matching in G' corresponds to a 1-factor of G. □

As mentioned earlier, there are several well known polynomial time algorithms for finding maximum matchings – these can be applied repeatedly to find the above factorization.

The following result (Theorem 16 from [37]) guarantees a Hamilton cycle in a robust outexpander G, as long as the minimum semidegree of G is not too small. As shown in [12], this Hamilton cycle can be found in polynomial time.

Theorem 6.2. *Suppose that $0 < 1/n \ll \nu \leq \tau \ll \alpha < 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq \alpha n$ which is a robust (ν,τ)-outexpander. Then G contains a Hamilton cycle.*

The *kth power* of a cycle C is obtained from C by adding an edge between every pair of vertices whose distance on C is at most k. We will also use the following result of Kómlos, Sárközy and Szemerédi [29] (where they proved the so-called Pósa-Seymour conjecture for sufficiently large graphs). We do not use the full strength of the result: any bound of the form $(1-\varepsilon)n$ instead of $10n/11$ would be sufficient for our proof. Moreover, we will only apply the result to a graph of bounded size (in the proof of Lemma 8.7), so an algorithmic version is not necessary.

Theorem 6.3. *There exists an $n_0 \in \mathbb{N}$ such that every graph G on $n \geq n_0$ vertices with minimum degree at least $10n/11$ contains the 10th power of a Hamilton cycle.*

The next two lemmas will be used to turn a 1-regular digraph F into a cycle on the same vertex set. More precisely, suppose that G is a blow-up of a cycle $C = V_1 \dots V_k$ so that for any successive clusters V_i, V_{i+1} the subdigraph $G[V_i, V_{i+1}]$ of G is superregular. Suppose also that F is a 1-regular digraph on $V_1 \cup \dots \cup V_k$ which 'winds around' C, i.e. each edge goes from V_i to V_{i+1} for some i. Then we can replace the edges of F from e.g. V_1 to V_2 with edges from G between these clusters to turn F into a Hamilton cycle. (Actually, the lemma is more general and does not require all edges of F to wind around C.)

To prove the lemmas, we will use an idea from [11].

Lemma 6.4. *Suppose that $0 < 1/m \ll d' \ll \varepsilon \ll d \ll \zeta, 1/t \le 1/2$. Let $V_1, \dots, V_k$ be pairwise disjoint clusters, each of size m and let $C = V_1 \dots V_k$ be a directed cycle on these clusters. Let G be a digraph on $V_1 \cup \dots \cup V_k$ and let $J \subseteq E(C)$. For each edge $V_iV_{i+1} \in J$, let $V_i^1 \subseteq V_i$ and $V_{i+1}^2 \subseteq V_{i+1}$ be such that $|V_i^1| = |V_{i+1}^2| \ge m/100$ and such that $G[V_i^1, V_{i+1}^2]$ is $(\varepsilon, d', \zeta d', td'/d)$-superregular. Suppose that F is a 1-regular digraph with $V_1 \cup \dots \cup V_k \subseteq V(F)$ such that the following properties hold:*

- (i) *For each edge $V_iV_{i+1} \in J$ the digraph $F[V_i^1, V_{i+1}^2]$ is a perfect matching.*
- (ii) *For each cycle D in F there is some edge $V_iV_{i+1} \in J$ such that D contains a vertex in V_i^1.*
- (iii) *Whenever $V_iV_{i+1}, V_jV_{j+1} \in J$ are such that J avoids all edges in the segment $V_{i+1}CV_j$ of C from V_{i+1} to V_j, then F contains a path P_{ij} joining some vertex $u_{i+1} \in V_{i+1}^2$ to some vertex $u'_j \in V_j^1$ such that P_{ij} winds around C.*

Then we can obtain a cycle on $V(F)$ from F by replacing $F[V_i^1, V_{i+1}^2]$ with a suitable perfect matching in $G[V_i^1, V_{i+1}^2]$ for each edge $V_iV_{i+1} \in J$. Moreover, if $J = E(C)$ then (iii) can be replaced by

- (iii$'$) *$V_i^1 \cap V_i^2 \neq \emptyset$ for all $i = 1, \dots, k$.*

Proof. For any edge $V_iV_{i+1} \in J$, let Old_i be the perfect matching $F[V_i^1, V_{i+1}^2]$. We will first prove the following:

For any edge $V_iV_{i+1} \in J$, we can find a perfect matching New_i in $G[V_i^1, V_{i+1}^2] =: G_i$ so that if we replace Old_i in F with New_i, then all vertices of G_i will lie on a common cycle in the new 1-factor F' thus obtained from F. Moreover any pair of vertices of F that were formerly on a common cycle in F are still on a common cycle in F'. (†)

To prove (†), we proceed as follows. Pick ν and τ such that $1/m \ll \nu \ll \tau \ll d'$. For every $u \in V_{i+1}^2$, we move along the cycle C_u of F containing u (starting at u) and let $f(u)$ be the first vertex on C_u in V_i^1 (note that $f(u)$ exists by (i)). Define an auxiliary digraph A on V_{i+1}^2 such that $N_A^+(u) := N_{G_i}^+(f(u))$. So A is obtained by identifying each pair $(u, f(u))$ into one vertex with an edge from $(u, f(u))$ to $(v, f(v))$ if G_i has an edge from $f(u)$ to v. So Lemma 4.11 applied with d', d/t playing the roles of d, μ implies that A is a robust (ν, τ)-outexpander. Moreover, $\delta^0(A) = \delta^0(G_i) \ge \zeta d'|V_{i+1}^2| = \zeta d'|A|$ by (Reg4). Thus Theorem 6.2 implies that A has a Hamilton cycle, which clearly corresponds to a perfect matching New_i in G_i with the desired property.

Now we apply (†) to every edge $V_iV_{i+1} \in J$ sequentially. We claim that after repeating this for every such edge, the resulting 1-regular digraph F'' is a cycle. To see this, note that (ii) and the last part of (†) together imply that every cycle of F'' contains a vertex in V_i^1 for some edge $V_iV_{i+1} \in J$. Moreover, by the first part of (†), all the vertices in $V_i^1 \cup V_{i+1}^2$ lie on a common cycle of F'', C_i say. So all the C_i together form F''. Consider any two edges $V_iV_{i+1}, V_jV_{j+1} \in J$ such that J avoids all edges in $V_{i+1}CV_j$ and let $P_{ij} = u_{i+1} \dots u'_j$ be the path guaranteed by (iii). Since P_{ij} winds around C, it follows that $P_{ij} \subseteq F''$. Thus u_{i+1} and u'_j lie on a common cycle in F''. But $u_{i+1} \in C_i$ and $u'_j \in C_j$. Thus $C_i = C_j$ and so all the C_i are identical. Thus F'' is a cycle.

It remains to check that if $J = E(C)$ then (iii) can be replaced by (iii′). But this is clear since in this case we have $i + 1 = j$ in (iii) and so we can take P_{ij} to be any vertex in $V_j^2 \cap V_j^1$. □

We will also use the following slightly different version of Lemma 6.4, which involves the usual notion of ε-regularity. In this paper, we will only use Lemmas 6.4 and 6.5 in the special case when $J = E(C)$. (So in Lemma 6.5 condition (iii) can be omitted.) The more general version of Lemma 6.5 will be used in [42].

Lemma 6.5. *Let $0 < 1/m \ll \varepsilon \ll d < 1$. Let $V_1, \dots, V_k$ be pairwise disjoint clusters, each of size m and let $C = V_1 \dots V_k$ be a directed cycle on these clusters. Let $J \subseteq E(C)$. Let G be a digraph on $V_1 \cup \dots \cup V_k$ such that $G[V_i, V_{i+1}]$ is (ε, d)-superregular for every $V_iV_{i+1} \in J$. For each edge $V_iV_{i+1} \in J$ let $V_i^1 \subseteq V_i$ and $V_{i+1}^2 \subseteq V_{i+1}$ be sets of size at least $(1 - d/2)m$ such that $|V_i^1| = |V_{i+1}^2|$. Suppose that F is a 1-regular digraph with $V_1 \cup \dots \cup V_k \subseteq V(F)$ such that the following properties hold:*

- (i) *For each edge $V_iV_{i+1} \in J$ the digraph $F[V_i^1, V_{i+1}^2]$ is a perfect matching.*
- (ii) *For each cycle D in F there is some edge $V_iV_{i+1} \in J$ such that D contains a vertex in V_i^1.*
- (iii) *Whenever $V_iV_{i+1}, V_jV_{j+1} \in J$ are such that J avoids all edges in the segment $V_{i+1}CV_j$ of C from V_{i+1} to V_j, then F contains a path P_{ij} joining some vertex $u_{i+1} \in V_{i+1}^2$ to some vertex $u'_j \in V_j^1$ such that P_{ij} winds around C.*

Then we can obtain a cycle on $V(F)$ from F by replacing $F[V_i^1, V_{i+1}^2]$ with a suitable perfect matching in $G[V_i^1, V_{i+1}^2]$ for each edge $V_iV_{i+1} \in J$. Moreover, if $J = E(C)$ then (iii) can be omitted.

Proof. The proof is very similar to that of Lemma 6.4. As before, for any edge $V_iV_{i+1} \in J$ let $G_i := G[V_i^1, V_{i+1}^2]$. Since $|V_i^1|, |V_{i+1}^2| \geq (1 - d/2)m$, the (ε, d)-superregularity of $G[V_i, V_{i+1}]$ implies that G_i is still $(2\varepsilon, d/2)$-superregular. Now we proceed as before to define A. Using the superregularity of G_i, it is easy to show that A is a robust $(\varepsilon d, 3\varepsilon)$-outexpander with $\delta^0(A) \geq d|A|/3$. (Indeed, for the outexpansion it suffices to observe that for all $U \subseteq V_i^1$ with $|U| \geq 3\varepsilon|A|$, the number of vertices in V_{i+1}^2 which receive at least $(d/2 - 2\varepsilon)|U| \geq \varepsilon d|A|$ edges from U in G_i is at least $(1 - 2\varepsilon)|A|$.) So as before we can apply Theorem 6.2 to find a Hamilton cycle in A, which corresponds to a matching as required in (†). The remainder of the argument is now identical. □

7. Schemes, consistent systems, chord sequences, and exceptional factors

7.1. **Schemes and consistent systems.** In order to simplify (and shorten) the statements of our lemmas throughout the paper, we will introduce the notions of a (k, m, ε, d)-scheme and of a consistent $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$-system. $(G, \mathcal{P}, R, C)$ is called a (k, m, ε, d)-*scheme* if the following properties are satisfied:

(Sch1) G and R are digraphs. $\mathcal{P}$ is a partition of $V(G)$ into an exceptional set V_0 of size at most $\varepsilon|G|$ and into k clusters of size m. The vertex set of R consists of these clusters.
(Sch2) For every edge UW of R the corresponding pair $G[U, W]$ is $(\varepsilon, \geq d)$-regular.
(Sch3) C is a Hamilton cycle in R and for every edge UW of C the corresponding pair $G[U, W]$ is $[\varepsilon, \geq d]$-superregular.
(Sch4) V_0 forms an independent set in G.

So roughly speaking, a scheme consists of a digraph G with a regularity partition $\mathcal{P}$ where the corresponding reduced digraph R contains a Hamilton cycle C. A consistent system has several additional features: mainly, the digraph G needs to be a robust outexpander and the definition involves an additional partition $\mathcal{P}_0$ which is coarser than $\mathcal{P}$. More precisely, $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ is called a *consistent* $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$-*system* if the following properties are satisfied (for (CSys3), recall that an ℓ^*-refinement was defined before Lemma 4.7):

(CSys1) Each of the digraphs G, R_0 and R is a robust (ν, τ)-outexpander. Moreover $\delta^0(G) \geq \alpha n$, where $n := |G|$, $\delta^0(R_0) \geq \alpha|R_0|$ and $\delta^0(R) \geq \alpha|R|$.
(CSys2) $\mathcal{P}_0$ is a partition of $V(G)$ into an exceptional set V_0^0 of size at most εn and into k/ℓ^* equal sized clusters. The vertex set of R_0 consists of these clusters. So $|R_0| = k/\ell^*$. Similarly, $\mathcal{P}$ is a partition of $V(G)$ into an exceptional set V_0 of size at most εn and into k clusters of size m. The vertex set of R consists of these clusters. So $|R| = k$.
(CSys3) $\mathcal{P}$ is obtained from an ℓ^*-refinement of $\mathcal{P}_0$ by removing some vertices from each cluster of this refinement and adding them to the exceptional set V_0^0. So V_0 is the union of V_0^0 with the set of these vertices.
(CSys4) For every edge UW of R the corresponding pair $G[U, W]$ is $(\varepsilon, \geq d)$-regular.
(CSys5) C_0 is a Hamilton cycle in R_0. Similarly, C is a Hamilton cycle in R and for every edge UW of C the corresponding pair $G[U, W]$ is $[\varepsilon, \geq d]$-superregular.
(CSys6) Suppose that W, W' are clusters in $\mathcal{P}_0$ and V, V' are clusters in $\mathcal{P}$ with $V \subseteq W$ and $V' \subseteq W'$. Then $WW' \in E(R_0)$ if and only if $VV' \in E(R)$.
(CSys7) C can be viewed as obtained from C_0 by winding ℓ^* times around C_0, i.e. for every edge WW' of C_0 there are precisely ℓ^* edges VV' of C such that $V \subseteq W$ and $V' \subseteq W'$.
(CSys8) Whenever W is a cluster in $\mathcal{P}_0$ and $x \in V(G)$ is a vertex with $n^+ \geq \tau|W|$ outneighbours in W, then x has at least $\theta n^+/\ell^*$ outneighbours in each cluster V in $\mathcal{P}$ with $V \subseteq W$. A similar condition holds for inneighbours of the vertices of G.
(CSys9) V_0 forms an independent set in G.

Note that (CSys3) and (CSys6) together imply that R is an ℓ^*-fold blow-up of R_0. Moreover, if $(G,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ is a consistent $(\ell^*,k,m,\varepsilon,d,\nu,\tau,\alpha,\theta)$-system then $(G,\mathcal{P},R,C)$ is a (k,m,ε,d)-scheme. We will usually denote the clusters in $\mathcal{P}$ by $V_1,\dots,V_k$ and assume that they are labelled in such a way that $C=V_1\dots V_k$.

The next result states that if $(G,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ is a consistent system and one deletes only a few edges at every vertex of G, then one still has a consistent system with slightly worse parameters. The analogue also holds for schemes.

Lemma 7.1. *Suppose that* $0<1/n\ll 1/k\ll\varepsilon\le\varepsilon'\ll d\ll\nu\ll\tau\ll\alpha,\theta\le 1$.

(i) *Let* $(G,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ *be a consistent* $(\ell^*,k,m,\varepsilon,d,\nu,\tau,\alpha,\theta)$*-system with* $|G|=n$. *Let* G' *be a digraph obtained from* G *by deleting at most* $\varepsilon' m$ *outedges and at most* $\varepsilon' m$ *inedges at every vertex of* G. *Then*
$$(G',\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$$
is still a consistent $(\ell^*,k,m,3\sqrt{\varepsilon'},d,\nu/2,\tau,\alpha/2,\theta/2)$*-system.*

(ii) *Let* $(G,\mathcal{P},R,C)$ *be a* (k,m,ε,d)*-scheme with* $|G|=n$. *Let* G' *be a digraph obtained from* G *by deleting at most* $\varepsilon' m$ *outedges and at most* $\varepsilon' m$ *inedges at every vertex of* G. *Then* $(G',\mathcal{P},R,C)$ *is still a* $(k,m,3\sqrt{\varepsilon'},d)$*-scheme.*

Proof. We only prove (i). The argument for (ii) is similar. It is easy to see that $(G',\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ still satisfies (CSys2), (CSys3), (CSys6), (CSys7) and (CSys9). Proposition 4.3 (applied with $d':=\varepsilon'$) implies that every edge of R still corresponds to an $2\sqrt{\varepsilon'}$-regular pair in G' of density at least $d-\varepsilon-2\sqrt{\varepsilon'}\ge d-3\sqrt{\varepsilon'}$. Similarly, every edge of C still corresponds to an $[2\sqrt{\varepsilon'},d]$-superregular pair in G'. Thus (CSys4) and (CSys5) hold with ε replaced by $3\sqrt{\varepsilon'}$. Moreover, it is easy to check that G' is still a robust $(\nu/2,\tau)$-outexpander with $\delta^0(G)\ge\alpha n/2$. So (CSys1) holds with ν and α replaced by $\nu/2$ and $\alpha/2$. Finally, suppose that x, n^+, W and V are as in (CSys8). Note that $|W|\ge\ell^* m$ and so
$$\varepsilon' m\le\theta\tau m/2\le\theta\tau|W|/2\ell^*\le\theta n^+/2\ell^*.$$
Thus in the digraph G' the number of outneighbours of x in V is at least $\theta n^+/\ell^*-\varepsilon' m\ge\theta n^+/2\ell^*$. So (CSys8) holds with θ replaced by $\theta/2$. □

7.2. **Shifted walks and chord sequences.** Roughly speaking, the Hamilton cycles we will find usually wind around a blown-up cycle $C=V_1\dots V_k$. Here the V_i are clusters. However, we also need to incorporate the vertices of an exceptional set V_0 into the cycle. For each $x\in V_0$, Lemma 7.5(i) below will give suitable in- and outneighbours x^- and x^+ which attach x to the blown-up cycle. However, to build a Hamilton cycle, we need additional edges: Suppose for example that $V_0=\{x\}$ and x^+ is not in the cluster succeeding the cluster containing x^-. Then it is impossible to extend the path x^-xx^+ into a Hamilton cycle in which all other edges wind around C. So we need additional edges which will 'balance out' the edges x^-x and xx^+. These additional edges are found via so-called 'shifted walks' and their associated chord sequences, which we define next. Shifted walks were first introduced in [25], also in order to find Hamilton cycles in directed graphs.

Let R be a digraph and let C be a Hamilton cycle in R. Given a vertex V of R, let V^+ denote the vertex succeeding V on C and let V^- denote the vertex preceding

V. (Later on, the vertices of R will be clusters, so we use capital letters to denote them.) A *shifted walk* from a vertex A to a vertex B in R is a walk $SW(A,B)$ of the form

$$SW(A,B) = V_1CV_1^-V_2CV_2^- \dots V_tCV_t^-V_{t+1},$$

where $V_1 = A$, $V_{t+1} = B$ and the edge $V_i^-V_{i+1}$ belongs to R for each $i = 1, \dots, t$. (Here we write $V_iCV_i^-$ for the path obtained from C by deleting the edge $V_i^-V_i$.) We say that $SW(A,B)$ *traverses* C t *times*. We call the edges $V_i^-V_{i+1}$ the *chord edges* of $SW(A,B)$. If $A = B$ then A is also a shifted walk from A to B. Without loss of generality, we may assume that an edge of C is not a chord edge according to the above definition. (Indeed, suppose that $V_i^-V_{i+1}$ is an edge of C. Then $V_{i+1} = V_i$ and so we can obtain a shorter shifted walk from A to B.)

For our purposes, it turns out that shifted walks contain too many edges. So we will only use their chord edges. So given a shifted walk

$$SW(A,B) = V_1CV_1^-V_2CV_2^- \dots V_tCV_t^-V_{t+1},$$

the corresponding *chord sequence* $CS(A,B)$ from A to B consists of all chord edges in $SW(A,B)$ in the same order as they appear in $SW(A,B)$. (In [38], this was called a skeleton walk, but we prefer not to use this name here as the chord edges do not actually form a walk.) We say that V lies in the *interior of* $CS(A,B)$ if $V \in \{V_2, V_2^-, \dots, V_t, V_t^-\}$.

The next result guarantees a short chord sequence between any two vertices in a robust outexpander. Moreover, this chord sequence can be chosen so that its interior avoids a given small set. The proof does not require the outexpansion property to be robust.

Lemma 7.2. *Let R be a robust (ν,τ)-outexpander with $\delta^0(R) \geq 2\tau|R|$ and $\nu \leq \tau \leq 1/3$. Let C be Hamilton cycle in R. Given vertices $A, B \in V(R)$ and a set of vertices $\mathcal{V}' \subseteq V(R)$ with $|\mathcal{V}'| \leq \nu|R|/4$, there is a chord sequence $CS(A,B)$ in R containing at most $3/\nu$ edges whose interior avoids $\mathcal{V}'$.*

Proof. Let $\mathcal{V}''$ be the union of $\mathcal{V}'$ and the set of all those clusters V for which $V^- \in \mathcal{V}'$ (where V^- is the predecessor of V on C). Thus $|\mathcal{V}''| \leq \nu|R|/2$. Let $\mathcal{V}^* := V(R) \setminus \mathcal{V}''$. Pick any outneighbour A_0 of A^- in $\mathcal{V}^*$. Let $S_1 := N_R^+(A_0^-) \cap \mathcal{V}^*$. So $|S_1| \geq 2\tau|R| - |\mathcal{V}''| \geq \nu|R|/2$. For each $i \geq 2$ let $S_i := N_R^+(N_C^-(S_{i-1})) \cap \mathcal{V}^*$. Thus $S_{i-1} \subseteq S_i$ and each cluster in S_i can be reached by a shifted walk from A_0 that traverses C at most i times and avoids $\mathcal{V}'$. Moreover for each $i \geq 2$ either $|S_{i-1}| < (1-\tau)|R|$ and

$$|S_i| \geq |RN_{\nu,R}^+(N_C^-(S_{i-1}))| - |\mathcal{V}''| \geq |N_C^-(S_{i-1})| + \nu|R| - |\mathcal{V}''| \geq |S_{i-1}| + \nu|R|/2$$

or $|S_{i-1}| \geq (1-\tau)|R|$ and $|S_i| \geq |S_{i-1}| \geq (1-\tau)|R|$. But this implies that $|S_{\lceil 2/\nu \rceil}| \geq (1-\tau)|R|$. Thus B has a neighbour B_0 such that its successor B_0^+ on C lies in $S_{\lceil 2/\nu \rceil}$. Since $B_0^+ \in S_{\lceil 2/\nu \rceil}$ there is a shifted walk $SW(A_0, B_0^+)$ which traverses C at most $\lceil 2/\nu \rceil$ times and which avoids $\mathcal{V}'$. Thus $ACA^-A_0 \cup SW(A_0, B_0^+) \cup B_0^+CB_0B$ is a shifted walk from A to B which traverses C at most $\lceil 2/\nu \rceil + 2 \leq 3/\nu$ times and which meets $\mathcal{V}'$ at most in the clusters A^- and B. So the chord sequence corresponding to this shifted walk is as required. □

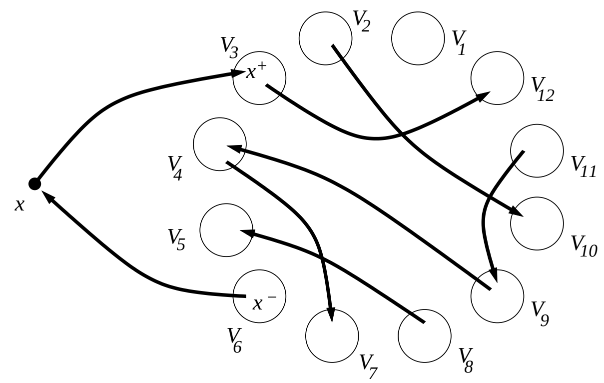

FIGURE 3. Balancing out a path x^-xx^+ using a sequence CS' obtained from the chord sequence $CS(U(x^+), U(x^-)^+) = CS(V_3, V_7) = (V_2V_{10}, V_9V_4, V_3V_{12}, V_{11}V_9, V_8V_5, V_4V_7)$. Here $C = V_1 \dots V_{12}$.

The following proposition records the crucial property of chord sequences for later use. As indicated earlier, it means we can use these sequences to 'balance out' an arbitrary edge x^-x^+ (or a path x^-xx^+ with $x \in V_0$) to obtain a 'locally balanced' set of edges.

Proposition 7.3. *Let $(G, \mathcal{P}, R, C)$ be a (k, m, ε, d)-scheme. Given vertices x^-, x^+ which are contained in clusters $U(x^-)$ and $U(x^+)$ in $\mathcal{P}$ respectively, consider any chord sequence $CS(U(x^+), U(x^-)^+) =: CS$ in R. Let CS' be obtained from CS by replacing each edge UW of CS with an edge of $G[U, W]$. Suppose that CS' is a matching which avoids both x^- and x^+. Let CS^* be obtained from CS' by adding the edge x^-x^+. For each cluster U in $\mathcal{P}$, let U^1 be the set of vertices of U which are not an initial vertex of an edge in CS^* and let U^2 be the set of vertices of U which are not a final vertex of an edge in CS^*. Then for each edge UW on C, we have $|U^1| = |W^2|$.*

Proof. This follows immediately from the fact that for every edge W^*W of CS (apart from the final edge of CS), the next edge of CS will be of the form UU^*, where U is the predecessor of W on C. If W^*W is the final edge of CS then $W = U(x^-)^+$ and so the edge $x^-x^+ \in CS^*$ has its initial vertex x^- in the predecessor of W on C (see Figure 3). □

In a typical application of this observation, the assertion that $|U^1| = |W^2|$ means that it will be possible to choose a perfect matching in $G[U^1, W^2]$. If we do this for each pair of consecutive clusters on C, then the union of all these matchings and the edges in CS^* forms a 1-regular digraph F covering all vertices in all clusters. We will then be able to transform F into a Hamilton cycle, e.g. using Lemma 6.4 or 6.5.

7.3. **Complete exceptional sequences and exceptional factors.** Suppose that $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme. Let $V_1, \dots, V_k$ denote the clusters in $\mathcal{P}$ such that $C = V_1 \dots V_k$. Recall that V_0 denotes the exceptional set. An *original exceptional edge* in G is an edge with one endvertex in V_0. (So by (Sch4) the other endvertex lies in $V(G) \setminus V_0$.) An *exceptional cover* EC consists of precisely one outedge and precisely one inedge incident to every vertex in V_0. Thus $|EC| = 2|V_0|$.

Let $G^{\rm basic}$ be obtained from G by adding all edges from $N^-(x)$ to $N^+(x)$ for every $x \in V_0$ and by deleting V_0. We call each edge yz from $N^-(x)$ to $N^+(x)$ that was added to G in order to obtain $G^{\rm basic}$ an *exceptional edge* and x the *vertex associated with* yz. Note that we might have $y = z$ in which case we add a loop. Moreover, we allow $G^{\rm basic}$ to have multiple edges: if e.g. $yz \in E(G)$ and yz is an edge from $N^-(x)$ to $N^+(x)$ for precisely two exceptional vertices then yz has multiplicity 3 in $G^{\rm basic}$ and precisely two of the edges from y to z in $G^{\rm basic}$ are exceptional edges. We sometimes write $G^{\rm orig}$ for G.

Given a spanning subdigraph H of G, we define $H^{\rm basic}$ in a similar way. Conversely, if H is a subdigraph of $G^{\rm basic}$, then $H^{\rm orig}$ is the subdigraph of G obtained from H by replacing each exceptional edge yz of H with the path yxz, where $x \in V_0$ is the exceptional vertex associated with yz. Note that if F is a 1-factor of G, then $F^{\rm basic}$ is a 1-factor of $G^{\rm basic}$. Conversely, a 1-factor F of $G^{\rm basic}$ which contains exactly one exceptional edge associated with every exceptional vertex corresponds to a 1-factor $F^{\rm orig}$ in $G^{\rm orig}$. Moreover, in this case F is a Hamilton cycle if and only if $F^{\rm orig}$ is a Hamilton cycle.

A *complete exceptional sequence* CES is a matching in $G^{\rm basic}$ which consists of precisely one exceptional edge associated with every exceptional vertex. So in particular CES does not contain loops. Note that the original version $CES^{\rm orig}$ of CES forms an exceptional cover. However, if EC is an exceptional cover, then $EC^{\rm basic}$ might contain paths, cycles (and loops) and so it might not form a complete exceptional sequence.

When constructing Hamilton cycles, we will usually do this by constructing a Hamilton cycle in $G^{\rm basic}$ which contains exactly one complete exceptional sequence and no other exceptional edges. In other words, we will use the following observation, which is an immediate consequence of the above discussion.

Observation 7.4. *Suppose that D is a Hamilton cycle in $G^{\rm basic}$ which contains exactly one complete exceptional sequence CES and no other exceptional edges. Then $D^{\rm orig}$ is a Hamilton cycle in G.*

For our arguments it is convenient to be able to define and use a digraph which is regular (when viewed as a subdigraph of $G^{\rm basic}$) and contains many different complete exceptional sequences – each of these will be part of a different Hamilton cycle in our decomposition. The exceptional factors EF defined below have the required properties.

Given $L \in \mathbb{N}$ which divides k, the *canonical interval partition of C into L intervals* consists of the intervals $V_{(i-1)L+1} \dots V_{iL+1}$ for all $i = 1, \dots, k/L$ (where $V_{k+1} := V_1$). Given an interval I on C, we write I° for the interior of I and $I^{\circ\circ}$ for the interior of I°. Moreover, we write $U \in I$ if U is a cluster on I.

Suppose that $k/L, m/K \in \mathbb{N}$ and let $\mathcal{I}$ be the canonical interval partition of C into L intervals of equal length. A *complete exceptional path system* $CEPS$ (with respect to C) with parameters (K, L) spanning an interval $I = U_jU_{j+1} \dots U_{j'}$ with $I \in \mathcal{I}$ consists of m/K vertex-disjoint paths $P_1, \dots, P_{m/K}$ in $G^{\rm basic}$ such that the following conditions hold.

(CEPS1) Every P_s has its initial vertex in U_j and its final vertex in $U_{j'}$.

(CEPS2) $CEPS$ contains a complete exceptional sequence CES (but no other exceptional edges) and all the edges in CES avoid the endclusters U_j and $U_{j'}$ of I.

(CEPS3) $CEPS$ contains precisely m/K vertices from every cluster in I and no other vertices.

Note that the above implies that $CEPS^{\rm orig}$ consists of m/K vertex-disjoint paths which cover all vertices in V_0 as well as m/K vertices in each cluster in I. Moreover, every path of $CEPS^{\rm orig}$ has its initial vertex in U_j and its final vertex in $U_{j'}$. If $K = 1$ (and thus the vertex set of $CEPS$ is the union of all the clusters in I), then we say that $CEPS$ *completely spans* I.

Suppose that $\mathcal{P}'$ is a K-refinement of $\mathcal{P}$. For each cluster $U \in \mathcal{P}$, let $U(1), \dots, U(K)$ denote the subclusters of U in $\mathcal{P}'$. Consider a complete exceptional path system $CEPS$ as above. We say that $CEPS$ has *style* b if its vertex set is $U_j(b) \cup \dots \cup U_{j'}(b)$. An *exceptional factor* EF *with parameters* (K, L) (with respect to C, $\mathcal{P}'$) is a 1-factor of $G^{\rm basic}$ which satisfies the following properties:

(EF1) On each of the L intervals $I \in \mathcal{I}$, EF induces the vertex-disjoint union of K complete exceptional path systems.

(EF2) Moreover, for each $I \in \mathcal{I}$ and each $b = 1, \dots, K$, exactly one of the exceptional path systems in EF spanning I has style b.

Note that EF consists of KL edge-disjoint complete exceptional path systems. Moreover, the second part of (CEPS2) implies that the union of all the KL complete exceptional sequences contained in these complete exceptional path systems forms a matching. This will be used in the proof of Lemma 8.1.

The reason that the definition of a consistent system involves not only the reduced digraph R_0 but also its refinement R is that this enables us to find complete exceptional path systems within an interval I of C and thus we will be able to find exceptional factors (see the proof of Lemma 7.5 for more details).

7.4. **Finding exceptional factors in a consistent system.** The following lemma will be used to construct the exceptional factors defined in the previous subsection. For (a), recall that an ε-uniform K-refinement of a partition $\mathcal{P}$ was defined before Lemma 4.7. Assertion (i) guarantees a 'localized' exceptional cover for V_0, assertion (ii) finds chord sequences in the reduced graph which 'balance out' this exceptional cover and (iii) finds edges in G which correspond to these chord sequences.

Lemma 7.5. *Suppose that* $0 < 1/n \ll 1/k \ll \varepsilon \ll \varepsilon' \ll d \ll \nu \ll \tau \ll \alpha, \theta \leq 1$, *that* $\ell^*/L, m/K \in \mathbb{N}$, *that* $L/\ell^* \ll 1$ *and* $\varepsilon \ll 1/K, 1/L$. *Let* $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ *be a consistent* $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$*-system with* $|G| = n$. *Suppose that* G' *is a spanning subdigraph of* G *and that* $\mathcal{P}'$ *is a partition of* $V(G)$ *such that the following conditions are satisfied:*

(a) $\mathcal{P}'$ *is an* ε*-uniform* K*-refinement of* $\mathcal{P}$*. (So in particular* V_0 *is the exceptional set in* $\mathcal{P}'$*.)*

(b) *Every vertex* $x \in V_0$ *satisfies* $d^\pm_G(x) - d^\pm_{G'}(x) \leq \varepsilon n$.

(c) *Every vertex* $x \in V(G) \setminus V_0$ *satisfies* $d^\pm_G(x) - d^\pm_{G'}(x) \leq (\varepsilon')^3 m/K$.

For every cluster U in $\mathcal{P}$ let $U(1),\dots,U(K)$ denote all those clusters in $\mathcal{P}'$ which are contained in U. Let $\mathcal{I}$ be the canonical interval partition of C into L intervals of equal length. Consider any $I\in\mathcal{I}$ and any j with $1\le j\le K$. Then the following properties hold:

(i) *For every exceptional vertex $x\in V_0$ there is a pair $x^-\neq x^+$ such that $x^-\in N^-_{G'}(x)\cap\bigcup_{U\in I^{\circ\circ}}U(j)$, $x^+\in N^+_{G'}(x)\cap\bigcup_{U\in I^{\circ\circ}}U(j)$, such that the vertices in all these pairs x^-,x^+ are distinct for different exceptional vertices and where each cluster in $\mathcal{P}$ contains at most $\varepsilon^{1/4}m$ vertices in these pairs.*

(ii) *Given a vertex $z\in V(G)\setminus V_0$, let $U(z)$ denote the cluster in $\mathcal{P}$ which contains z and let $U(z)^+$ be the successor of that cluster on C. Then for each of the pairs x^-,x^+ guaranteed by (i) there is a chord sequence CS_x from $U(x^+)$ to $U(x^-)^+$ in R which contains at most $3/\nu^3$ edges such that each of these edges has both endvertices in I° and such that in total no cluster in $\mathcal{P}$ is used more than $4\varepsilon^{1/4}m$ times by all these CS_x.*

(iii) *For each $x\in V_0$ there is a sequence CS'_x of edges in G' obtained by replacing every edge UU' in CS_x by an edge in $G'[U(j),U'(j)]$ such that CS'_x forms a matching which avoids all the pairs y^-,y^+ guaranteed by (i) (for all $y\in V_0$) and such that the CS'_x are pairwise vertex-disjoint. (Thus $\bigcup_{x\in V_0}CS'_x$ is a matching too.)*

(iv) *For every edge UW of C the pair $G'[U(j),W(j)]$ is $[\varepsilon',\ge d]$-superregular.*

Proof. Recall that $k=|C|$ denotes the number of clusters in $\mathcal{P}$. We will first prove (i). We will choose the pairs x^-,x^+ for every exceptional vertex x in turn. So suppose that for some exceptional vertices we have already chosen such pairs and that we now wish to choose such a pair for $x\in V_0$. Let S denote the set of all vertices lying in the pairs that we have chosen already. So $|S|<2|V_0|\le 2\varepsilon n$. We say that a cluster of $\mathcal{P}$ is *full* if it contains at least $\varepsilon^{1/4}m$ vertices from S. Then the number of full clusters is at most

$$\frac{|S|}{\varepsilon^{1/4}m}\le\frac{2\varepsilon n}{\varepsilon^{1/4}m}\le\sqrt{\varepsilon}k.$$

So for any vertex $x\in V_0$, the number of outneighbours of x in the full clusters is at most $\sqrt{\varepsilon}km\le\sqrt{\varepsilon}n$.

Set $k_0:=k/\ell^*$ and recall that $k_0=|R_0|=|C_0|$. Note that the length of I is $k/L=k_0\ell^*/L$ and thus the length of I is a multiple of k_0 (as ℓ^* is a multiple of L by assumption). But by (CSys7) C was obtained from the Hamilton cycle C_0 on the clusters in $\mathcal{P}_0$ by winding ℓ^* times around C_0. Thus for each cluster V in $\mathcal{P}_0$, I contains at least

$$\frac{\text{length}(I)}{k_0}=\frac{k/L}{k/\ell^*}=\frac{\ell^*}{L}\tag{7.1}$$

clusters of the current partition $\mathcal{P}$ which are contained in V.

We say that a cluster $V\in\mathcal{P}_0$ is *friendly* if $|N^+_G(x)\cap V|\ge\alpha|V|/3$. (CSys1) implies that $\delta^0(G)\ge\alpha n$ and so at least $\alpha k_0/3$ clusters in $\mathcal{P}_0$ are friendly. But (CSys8) now implies that for each such cluster V, x has at least $\theta|N^+_G(x)\cap V|/\ell^*$ outneighbours in each cluster U of $\mathcal{P}$ with $U\subseteq V$ (in the digraph G). Then by (a) and the condition

(URef) in the definition of an ε-uniform refinement, we have that x has at least

$$\frac{\theta|N_G^+(x)\cap V|}{2\ell^* K}\geq \frac{\theta\alpha|V|}{6\ell^* K}\geq \frac{\theta\alpha n}{7\ell^* K k_0}$$

outneighbours in each subcluster $U(j)$ of $\mathcal{P}'$ with $U(j)\subseteq V$ (again in the digraph G). Together with (7.1) this implies that x has at least

$$\frac{\alpha k_0}{3}\cdot\frac{\theta\alpha n}{7\ell^* K k_0}\cdot\frac{\ell^*}{2L}\geq \varepsilon^{1/8}n$$

outneighbours in G which lie in $\bigcup_{U\in I^{\circ\circ}}U(j)$. (Here we multiply with $\ell^*/2L$ instead of ℓ^*/L since counting outneighbours in $\bigcup_{U\in I^{\circ\circ}}U(j)$ means that we lose 4 clusters when considering $I^{\circ\circ}$ instead of I.) Together with (b) and the fact that at most $\sqrt{\varepsilon}n$ outneighbours of x lie in full clusters of $\mathcal{P}$ this shows that we can always find an outneighbour $x^+\in\bigcup_{U\in I^{\circ\circ}}U(j)\setminus S$ of x in G' such that x^+ lies in a cluster of $\mathcal{P}$ which is not full. We can argue similarly to find x^-. This proves (i).

Our next aim is to prove (ii). We will choose the sequence CS_x for every exceptional vertex x in turn. So suppose that for some exceptional vertices we have already chosen such sequences and that we now wish to choose CS_x for $x\in V_0$. Call a cluster in $\mathcal{P}$ *crowded* if it is used at least $\varepsilon^{1/4}m$ times by the sequences $CS_{x'}$ found so far. Thus the total number of crowded clusters in $\mathcal{P}$ is at most $|V_0|\cdot\frac{6/\nu^3}{\varepsilon^{1/4}m}\leq\varepsilon^{2/3}k\leq\sqrt{\varepsilon}|I|$.

Let $U_{\rm first}$ and $U_{\rm final}$ denote the first and the final cluster of I. Let I^* denote the interval obtained from I by deleting $U_{\rm final}$. As already observed in our proof of (i), the length of I is a multiple of k_0. So for every cluster V in $\mathcal{P}_0$ there are precisely length$(I)/k_0 =: D$ clusters in I^* which are contained in V. Moreover, $U_{\rm first}$ and $U_{\rm final}$ are contained in the same cluster of $\mathcal{P}_0$. Consider the subgraph R^* of R induced by all the clusters in I^*. Note that R^* contains an edge E^* from the final cluster $U^*_{\rm final}$ of I^* to $U_{\rm first}$. (This follows since (CSys6) implies that R_0 has an edge from the cluster in $\mathcal{P}_0$ containing $U^*_{\rm final}$ to the cluster in $\mathcal{P}_0$ containing $U_{\rm final}$. But $U_{\rm first}$ and $U_{\rm final}$ are contained in the same cluster of $\mathcal{P}_0$. So (CSys6) now implies that R (and thus also R^*) has an edge from $U^*_{\rm final}$ to $U_{\rm first}$.) Let C^* be the Hamilton cycle of R^* which consists of I^* together with this edge E^*. (CSys6) also implies that R^* can be viewed as being obtained from R_0 by replacing each vertex of R_0 by D vertices and replacing each edge WW' of R_0 by a complete bipartite graph between the two corresponding sets of D vertices (where all the edges are directed from the D vertices corresponding to W to the D vertices corresponding to W'). In other words, R^* is a D-fold blow-up of R_0. Together with (CSys1) and Lemma 5.3 this in turn implies that R^* is a robust $(\nu^3,2\tau)$-outexpander. Moreover, it is easy to check that $\delta^0(R^*)\geq\alpha|R^*|$.

Apply Lemma 7.2 to R^* (with $\mathcal{V}'$ consisting of $U_{\rm first}$ and all the crowded clusters in I^* and with C^* playing the role of C) to find a chord sequence CS_x from $U(x^+)$ and $U(x^-)^+$ in R^* which contains most $3/\nu^3$ edges and whose interior avoids $U_{\rm first}$ as well as all the crowded clusters. Since $x^-,x^+\in\bigcup_{U\in I^{\circ\circ}}U(j)$ by (i), this implies that CS_x uses only edges whose endclusters both lie in I°. But $R^*\subseteq R$. Thus CS_x is actually a chord sequence in R.

We proceed in this way to choose CS_x for all $x\in V_0$. We claim that in total no cluster in $\mathcal{P}$ is used more than $4\varepsilon^{1/4}m$ times by the sequences CS_x. To see this, note that no cluster in $\mathcal{P}$ is used more than $\varepsilon^{1/4}m + 3/\nu^3 \leq 2\varepsilon^{1/4}m$ times by the interiors of the all CS_x (as all these interiors avoid the crowded clusters). Moreover, since by (i) no cluster in $\mathcal{P}$ contains more than $\varepsilon^{1/4}m$ vertices belonging to pairs x^-, x^+ guaranteed by (i), it follows that no cluster in $\mathcal{P}$ is used more than $2\varepsilon^{1/4}m$ times as one of the two remaining clusters $U(x^+)^-$, $U(x^-)^+$ of CS_x (for all $x \in V_0$ together). This completes the proof of (ii).

It remains to check (iii) and (iv). To do this, we will first prove the following claim.

Claim 1. *For every edge UU' of R the pair $G'[U(j), U'(j)]$ is ε'-regular of density at least $d/2$. Moreover, for every edge UW of C the pair $G'[U(j), W(j)]$ is $[\varepsilon', \geq d]$-superregular.*

To prove the first part of Claim 1, consider any edge UU' of R. (CSys4) implies that $G[U(j), U'(j)]$ is $K\varepsilon$-regular of density at least $d - 2\varepsilon$. Together with (c) and Proposition 4.3(i) this implies that $G'[U(j), U'(j)]$ is still ε'-regular of density at least $d - 2\varepsilon' \geq d/2$.

Now consider any edge UW of C. Then (CSys5), (a) and Lemma 4.7(i) together imply that $G[U(j), W(j)]$ is $[(\varepsilon')^3, \geq d]$-superregular. As before, together with (c) and Proposition 4.3(iii) this implies that $G'[U(j), W(j)]$ is still $[\varepsilon', \geq d]$-superregular. This proves Claim 1 and thus in particular (iv).

In order to replace the edges of CS_x to obtain CS'_x, we again consider each exceptional vertex $x \in V_0$ in turn. By making CS_x shorter if necessary, we may assume that every edge of R occurs at most once in each CS_x. Suppose that we have already chosen $CS'_{x'}$ for some vertices $x' \in V_0$ and that we now wish to replace the edges of CS_x in order to choose CS'_x. Moreover, suppose that we have already replaced some edges $U''U'''$ of CS_x and we now wish to replace the edge UU' of CS_x.

For this, we let $U_*(j)$ denote the set of all those vertices $u \in U(j)$ that satisfy the following three conditions:

- There is no $y \in V_0$ such that u lies in the pair y^-, y^+ guaranteed by (i).
- There exists no $x' \in V_0$ for which we have already defined $CS'_{x'}$ and for which u is an endvertex of some edge in $CS'_{x'}$.
- u is not an endvertex of an edge of $G'[U''(j), U'''(j)]$ which was used to replace some edge $U''U'''$ of CS_x.

(i) implies that U contains at most $\varepsilon^{1/4}m$ vertices violating the first condition and (ii) implies that U contains at most $4\varepsilon^{1/4}m$ vertices violating the second or third condition. Thus

$$|U_*(j)| \geq |U(j)| - 5\varepsilon^{1/4}m = \frac{m}{K} - 5\varepsilon^{1/4}m \geq \frac{|U(j)|}{2}. \tag{7.2}$$

Our aim is to replace UU' in CS_x by an edge in $G'[U_*(j), U'_*(j)]$. But by Claim 1, $G'[U(j), U'(j)]$ is ε'-regular of density at least $d/2$. Together with (7.2) this implies that $G'[U_*(j), U'_*(j)]$ contains an edge. We do this for every edge of CS_x in turn and let CS'_x denote the sequence obtained in this way. Then these sequences CS'_x are as required in (iii). □

Based on the previous lemma, it is now straightforward to construct many edge-disjoint exceptional factors.

Lemma 7.6. *Suppose that*

$$0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll \nu \ll \tau \ll \alpha, \theta \leq 1; \quad \varepsilon \ll 1/K, 1/L; \quad Kr_0/m \ll d,$$

that $\ell^*/L, m/K \in \mathbb{N}$ and $L/\ell^* \ll 1$. Let $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ be a consistent $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$-system with $|G| = n$. Let $\mathcal{P}'$ be an ε-uniform K-refinement of $\mathcal{P}$. Then there is a set $\mathcal{EF}$ of r_0 exceptional factors with parameters (K, L) (with respect to C, $\mathcal{P}'$) such that the original versions of all these r_0 exceptional factors are pairwise edge-disjoint subdigraphs of G.

Proof. Choose a new constant ε' with $\varepsilon, Kr_0/m \ll \varepsilon' \ll d$. We will choose the r_0 exceptional factors for $\mathcal{EF}$ with respect to $C, \mathcal{P}'$ one by one. So suppose that for some $0 \leq s < r_0$ we have already chosen exceptional factors $EF_1, \dots, EF_s$ with parameters (K, L) such that the original versions of these factors are pairwise edge-disjoint from each other. So our aim is to show that we can choose EF_{s+1} such that all these properties still hold.

Let G' be the digraph obtained from G by deleting all the edges contained in the original versions of all the exceptional factors chosen so far. Note that

$$d^{\pm}_G(x) - d^{\pm}_{G'}(x) \leq KLr_0 \leq Ldm \leq Ldn/k \leq \varepsilon n.$$

for every vertex $x \in V_0$ and

$$d^{\pm}_G(x) - d^{\pm}_{G'}(x) \leq r_0 \leq (\varepsilon')^3 m/K$$

for every vertex $x \in V(G) \setminus V_0$. Thus G' satisfies conditions (b) and (c) of Lemma 7.5.

Let $\mathcal{I}$ denote the canonical interval partition of C into L intervals of equal length. For each cluster U of $\mathcal{P}$ let $U(1), \dots, U(K)$ denote the subclusters in $\mathcal{P}'$ which are contained in U. Consider any interval $I \in \mathcal{I}$ and any j with $1 \leq j \leq K$. In order to construct EF_{s+1}, it suffices to show that there is a complete exceptional path system $CEPS$ spanning the interval I whose vertex set is $\bigcup_{U \in I} U(j)$ and such that $CEPS^{\mathrm{orig}} \subseteq G'$. To do this, we apply Lemma 7.5 to find pairs x^-, x^+, chord sequences CS_x in R and sequences CS'_x (for every vertex $x \in V_0$) satisfying (i)–(iii).

Now let CES be the union of all the edges x^-x^+ over all exceptional vertices x. (Note that x^-x^+ is an exceptional edge in $(G')^{\mathrm{basic}}$ and it is irrelevant whether it lies in G' or not.) It remains to enlarge CES into a complete exceptional path system $CEPS$. For this, we first add all edges in the sequences CS'_x guaranteed by (iii) (for all vertices $x \in V_0$). Together with CES, this gives a matching M which meets every cluster in $\mathcal{P}$ in at most $5\varepsilon^{1/4}m$ vertices.

Suppose that $U, W \in \mathcal{P}$ are consecutive clusters on I. Let $U^1(j)$ be the set of all those vertices in $U(j)$ which are not the initial vertex of an edge in M and let $W^2(j)$ be the set of all those vertices in $W(j)$ which are not the final vertex of an edge in M. Then Proposition 7.3 implies that $|U^1(j)| = |W^2(j)|$. Also note that $|U(j) \setminus U^1(j)| \leq 5\varepsilon^{1/4}m \leq \varepsilon'|U(j)|$. Together with Lemma 7.5(iv) and Proposition 4.3(iii) this implies that $G'[U^1(j), W^2(j)]$ is still $[2\sqrt{\varepsilon'}, \geq d]$-superregular and thus it contains a perfect matching M_{UW} by Proposition 4.14.

The union $CEPS'$ of M with these matchings M_{UW} for all pairs U, W of consecutive clusters on I contains paths satisfying (CEPS1) and (CEPS2), but in addition $CEPS'$ might contain cycles. (Note each such cycle will contain at least one edge from M.) So our aim is to apply Lemma 6.5 in order to transform $CEPS'$ into a path system. To do this, we let $U_{\rm first} \in \mathcal{P}$ denote the first cluster in I and let $U_{\rm final} \in \mathcal{P}$ denote the last cluster in I. Let C_I be the cycle obtained from I by identifying $U_{\rm first}$ and $U_{\rm final}$. Let G'_{Ij} be the digraph obtained from $\bigcup_{UW \in E(I)} G'[U(j), W(j)]$ by identifying each vertex in $U_{\rm first}(j)$ with a different vertex of $U_{\rm final}(j)$. Let $CEPS''$ be obtained from $CEPS'$ by identifying these vertices. Note that G'_{Ij} can be viewed as a blow-up of C_I in which every edge of C_I corresponds to an $[\varepsilon', \geq d]$-superregular pair (the latter holds by Lemma 7.5(iv)). Moreover, $CEPS''$ is a 1-regular digraph on $V(G'_{Ij})$ which has the property that every cycle D in $CEPS''$ contains at least one edge in some matching M_{UW} for some pair U, W of consecutive clusters on C_I. (To see the latter, recall that M was a matching which avoids both $U_{\rm first}$ and $U_{\rm final}$. So every vertex in $V(G'_{Ij})$ is incident to at most one edge in M.) So D contains a vertex in $U^1(j)$. Thus we can apply Lemma 6.5 with G'_{Ij}, C_I, $U^1(j)$, $U^2(j)$ and $E(C_I)$ playing the roles of G, C, V_i^1, V_i^2 and J. This shows that we can replace each matching M_{UW} by a different matching in $G'[U^1(j), W^2(j)]$ in order to transform $CEPS''$ into a Hamilton cycle of G'_{Ij}. But this Hamilton cycle corresponds to a complete exceptional path system $CEPS$ spanning the interval I whose vertex set is $\bigcup_{U \in I} U(j)$ and such that $CEPS^{\rm orig} \subseteq G'$. (Note that $CEPS$ still contains M and thus CES.)

We do this for each $j = 1, \dots, K$ and each interval I in the canonical interval partition $\mathcal{I}$ of C in turn. Then the union EF_{s+1} of all these complete exceptional path systems is an exceptional factor with the desired properties. This completes the proof of the existence of $\mathcal{EF}$. $\square$

8. The preprocessing step

Let G' be the leftover of G obtained from an application of Theorem 1.3. So G' is regular and very sparse. Roughly speaking, the aim of this section is to find a sparse 'preprocessing graph' PG so that $G' \cup PG$ has a set of edge-disjoint Hamilton cycles covering all edges of G'. We need to do this because the edges at the exceptional vertices in the leftover G' might be distributed very badly. So the idea is to choose PG at the beginning of the proof of Theorem 1.2, to apply Theorem 1.3 to $G \setminus E(PG)$ and then to cover the leftover G' by edge-disjoint Hamilton cycles in $G' \cup PG$. In this way we replace G' by the resulting leftover PG' of PG. Moreover, the goal is that PG' will have no edges incident to V_0. So PG' will not be regular. But the chord absorber, which we will use in Section 9 to absorb the edges of PG', will contain additional edges at the exceptional vertices to compensate for this, and these edges will be nicely distributed. (More precisely, this chord absorber will contain some additional exceptional factors.)

In order to find the Hamilton cycles covering G', we decompose G' into 1-factors H, split each 1-factor H into small path systems H_i and extend H_i into a Hamilton cycle using the edges of PG. As we shall see, when finding these Hamilton cycles, the most difficult part is to find suitable edges joining V_0 to the (non-exceptional) clusters. For this step we use the complete exceptional sequences defined in Section 7.3.

As we shall explain in Section 8.3, the preprocessing graph PG will consist (mainly) of the edge-disjoint union of exceptional factors and a 'path system extender' PE, which we define and use in Section 8.2. Once we have found suitable edges joining V_0 to the clusters, the path system extender will be used to extend these into Hamilton cycles.

8.1. **Cycle breaking.** Suppose that H is a 1-factor of the leftover G'. As discussed above, our aim is to split H into small path systems H_i which we then extend into Hamilton cycles C_i of G. Note that H contains exactly one exceptional cover $H^{\rm exc}$ and each C_i must contain exactly one as well. The most obvious way to achieve the latter might be to let H_1 consist of $H^{\rm exc}$, and add a new exceptional cover (from the preprocessing graph PG) to each of the other H_i when forming the C_i. However, $H^{\rm exc}$ might contain cycles, in which case we cannot extend $H_1 = H^{\rm exc}$ to a Hamilton cycle. So when splitting H into the H_i, we also need to split $H^{\rm exc}$ to ensure that such cycles are 'broken'. As an intermediate step towards extending the H_i into Hamilton cycles, we then add edges from (the original version of) an exceptional factor CB to extend each H_i into a path system Q_i which contains exactly one exceptional cover. In other words, one can split $H \cup CB^{\rm orig}$ into small path systems Q_i such that each of them contains precisely one exceptional cover.

Let $(G, \mathcal{P}, R, C)$ be a (k, m, ε, d)-scheme and assume that $m/50 \in \mathbb{N}$. Consider an ε-uniform 50-refinement $\mathcal{P}'$ of $\mathcal{P}$ (recall again that these were defined before Lemma 4.7). For each cluster V_i of $\mathcal{P}$, let $V_i(1), \dots, V_i(50)$ denote all those clusters in $\mathcal{P}'$ which are contained in V_i. Suppose $J \subseteq \{1, \dots, 50\}$. Generalizing the notion of styles defined at the end of Section 7.3, we say that a vertex $x \in V(G) \setminus V_0$ has *style* J if $x \in \bigcup_{i=1}^k \bigcup_{j\in J} V_i(j)$. We then say that a set E of edges of $G^{\rm basic}$ or of G has *style* J if every endvertex x of an edge in E with $x \notin V_0$ has style J. If E has style J, we say that E has *style size* $|J|$ (with respect to $\mathcal{P}'$). Note that in this definition J need not have minimum size, i.e. if E has style size t, then it also has style size $t+1$.

Lemma 8.1. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll 1/s \ll 1$ and that $m/50, 50k/(s-1) \in \mathbb{N}$. Let $(G, \mathcal{P}, R, C)$ be a (k, m, ε, d)-scheme with $|G| = n$. Let $\mathcal{P}'$ be an ε-uniform 50-refinement of $\mathcal{P}$. Let H be a 1-factor of G. Let CB be an exceptional factor with parameters $(50, (s-1)/50)$ with respect to $C, \mathcal{P}'$. Suppose that H and $CB^{\rm orig}$ are edge-disjoint. Then the edges of $H \cup CB^{\rm orig}$ can be decomposed into edge-disjoint path systems $Q_1, \dots, Q_s$ so that*

- (i) *each Q_i contains precisely one exceptional cover (and no other original exceptional edges);*
- (ii) *each Q_i has style size 5 (with respect to $\mathcal{P}'$);*
- (iii) $|E(Q_i)| \leq 1230n/s$.

Proof. Let $H^{\rm exc}$ denote the set of original exceptional edges of H. Note that each edge in $H^{\rm exc}$ has precisely one of 50 styles. So we can partition $H^{\rm exc}$ into $H_1, \dots, H_{50}$ such that each H_i has style i. Now we split each H_i into H_i^- and H_i^+ by placing one edge from each cycle of H_i into H_i^- and all the others into H_i^+. Enumerate the path systems obtained in this way as $H_1^{\rm exc},\dots,H_{100}^{\rm exc}$. Thus $H^{\rm exc}$ is the union of all these path systems. Denote the complete exceptional path systems of CB by $CEPS_i$, for $i=1,\dots,s-1$, and let CES_i denote the complete exceptional sequence contained in $CEPS_i$. Without loss of generality, we may assume that $CEPS_1,\dots,CEPS_{100}$ are complete exceptional path systems of CB so that exactly 20 of them have style j for $j=1,\dots,5$ (as there are many more with these styles in CB). Moreover, by relabeling the $CEPS_i$ if necessary, we may assume that $H_i^{\rm exc}$ and $CEPS_i$ have different styles for all $i=1,\dots,100$.

For each $i=1,\dots,100$, let $V_{0,i}^+\subseteq V_0$ be the set of exceptional vertices x so that x is the initial vertex of an edge in $H_i^{\rm exc}$. Similarly, let $V_{0,i}^-\subseteq V_0$ be the set of exceptional vertices x so that x is the final vertex of an edge in $H_i^{\rm exc}$. Let CES_i' be the set of all those (original exceptional) edges xy in $CES_i^{\rm orig}$ for which $x\in V_0\setminus V_{0,i}^+$ as well as all those (original exceptional) edges yx in $CES_i^{\rm orig}$ for which $x\in V_0\setminus V_{0,i}^-$. Let $Q_i:=H_i^{\rm exc}\cup CES_i'\cup(CEPS_i\setminus CES_i)$. Note that Q_i contains the exceptional cover $H_i^{\rm exc}\cup CES_i'$. Moreover, since $H_i^{\rm exc}$ and $CEPS_i$ have different styles it follows that Q_i is a path system, i.e. it does not contain any cycles. Furthermore, Q_i has style size 2 and $|E(Q_i)|\le|V_0|+n/(s-1)\le 2n/s$.

Now let Q_{101} consist of those edges of $CES_1^{\rm orig},\dots,CES_{100}^{\rm orig}$ which are not contained in any of $Q_1,\dots,Q_{100}$. Using the fact that $H^{\rm exc}=H_1^{\rm exc}\cup\dots\cup H_{100}^{\rm exc}$ is an exceptional cover, it is easy to see that Q_{101} also forms an exceptional cover. Moreover, as remarked after the definition of an exceptional factor, $CES_1\cup\dots\cup CES_{100}$ is a matching. In particular, every vertex in $V(G)\setminus V_0$ is incident to at most one edge in $CES_1^{\rm orig}\cup\dots\cup CES_{100}^{\rm orig}$. Together with the fact that Q_{101} is an exceptional cover, it follows that Q_{101} is a path system. Also, Q_{101} has style size 5 and $|E(Q_{101})|=2|V_0|\le 2\varepsilon n\le n/s$.

Note that each edge of $H\setminus H^{\rm exc}$ has at least one of $\binom{50}{2}=1225$ styles ij with $i\neq j$. (Recall that if an edge has style i then it also has style ij for any j.) Partition $H\setminus H^{\rm exc}$ into 1225 sets $H_{i,j}$ such that all edges in $H_{i,j}$ have style ij. Greedily split each set $H_{i,j}$ further into $(s-1)/1226$ sets such that each of them has size at most $\lceil 1226n/(s-1)\rceil$ and none of them contains a cycle. Let $s':=1225(s-1)/1226$ and let $H_1^*,\dots,H_{s'}^*$ denote the resulting sets of edges.

Claim. *We may assume that $CEPS_{101},\dots,CEPS_{s-1}$ are enumerated in such a way that for all $t=1,\dots,s'$ the following property holds: if H_t^* has style ij and $CEPS_{100+t}$ has style j' then $j'\notin\{i,j\}$.*

To prove the claim, we consider the auxiliary bipartite graph Y with vertex classes A and B, where A consists of all the H_t^* and B consists of $CEPS_{101},\dots,CEPS_{s-1}$. (So $|A|=s'$ and $|B|=(s-1)-100>s'$.) Suppose that H_t^* has style ij and $CEPS_\ell$ has style j', where $j'\notin\{i,j\}$. Then H_t^* and $CEPS_\ell$ are connected by an edge in Y. Note that every $CEPS_\ell$ has degree at least $|A|-49\cdot(s-1)/1226\ge|A|/2$ (since if $j'=i$ then there are only 49 possibilities for j and since only $(s-1)/1226$ of the H_t^* have the same style). Similarly every H_t^* has degree at least $(s-1-100)-2\cdot(s-1)/50\ge |B|/2$. So Y has a matching covering A. (Indeed, this follows from Hall's theorem: Consider any $A'\subseteq A$. If $|A'|>|A|/2$ then $N_Y(A')=B$ and if $|A'|\le|A|/2$ then $|N_Y(A')|\ge|B|/2\ge|A|/2\ge|A'|$.) This completes the proof of the claim.

For $i = 102, \dots, 101 + s'$, let $Q_i := H^*_{i-101} \cup CEPS^{\rm orig}_{i-1}$. Then the claim implies that Q_i is a path system. Moreover, Q_i has style size 3 and $|E(Q_i)| \le \lceil 1226n/(s-1)\rceil + n/(s-1) + \varepsilon n \le 1230n/s$. Finally, for $i = 102+s', \dots, s$ we let $Q_i := CEPS^{\rm orig}_{i-1}$. Then $Q_1, \dots, Q_s$ have the desired properties. □

8.2. **Extending path systems into Hamilton cycles.** Suppose that $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme. Recall that V_0 denotes the exceptional set in $\mathcal{P}$. The purpose of this subsection is to define and use the 'path system extender' PE. To motivate its purpose, let G' be the leftover of G obtained by an application of Theorem 1.3 and let H be a 1-factor in a 1-factorization of G'. Recall that Lemma 8.1 assigns edges of H to edge-disjoint path systems Q_i. PE will be used to extend each path system Q_i into a Hamilton cycle. Altogether this means that we will find edge-disjoint Hamilton cycles in the union of $G' \cup PE$ with some exceptional factors CB_i which cover both the edges of G' and the edges of CB_i. Since PE will be a spanning subdigraph of $G - V_0$, this in turn implies that we have replaced G' by a digraph (namely the digraph obtained from the leftover of PE by adding V_0) in which every exceptional vertex is isolated.

A *path system extender PE (for C, R) with parameters* $(\varepsilon, d, d', \zeta)$ is a spanning subgraph of $G - V_0$ consisting of an edge-disjoint union of two graphs $\mathcal{B}(C)_{PE}$ and $\mathcal{B}(R)_{PE}$ on $V(G) \setminus V_0$ which are defined as follows:

- (PE1) $\mathcal{B}(C)_{PE}$ is a blow-up of C in which every edge UW of C corresponds to an $(\varepsilon, d', \zeta d', 2d'/d)$-superregular pair $\mathcal{B}(C)_{PE}[U, W]$.
- (PE2) $\mathcal{B}(R)_{PE}$ is a blow-up of R in which every edge UW of R corresponds to an $(\varepsilon, d'/k, 2d'/dk)$-regular pair $\mathcal{B}(R)_{PE}[U, W]$.

Note that the path system extender PE is not necessarily a regular digraph. (Reg3) and the fact that $\Delta(R) \le 2k$ imply that

$$\Delta(PE) \le 2 \cdot \frac{2d'm}{d} + \Delta(R) \cdot \frac{2d'm}{dk} \le \frac{8d'm}{d}. \tag{8.1}$$

In order to cover the edges of the leftover of the path system extender PE with Hamilton cycles in the chord-absorbing step (Section 9), we will need that this maximum degree is small compared to εm (so PE is a rather sparse graph). This is what forces us to use the above more technical notion of regularity when defining PE, rather than the usual ε-regularity.

Lemma 8.2. *Suppose that* $0 < 1/n \ll d' \ll 1/k \ll \varepsilon \ll d \ll \zeta \le 1/2$*. Let* $(G, \mathcal{P}, R, C)$ *be a* (k, m, ε, d)*-scheme with* $|G| = n$*. Then G contains a path system extender for C, R having parameters* $(\varepsilon^{1/25}, d, d', \zeta)$*.*

Proof. Recall that by (Sch3) every edge UW of C corresponds to an $[\varepsilon, d_{UW}]$-superregular pair $G[U, W]$ for some $d_{UW} \ge d$. Thus we can apply Lemma 4.10(ii) to each such $G[U, W]$ to obtain a blow-up $\mathcal{B}(C)_{PE}$ of C in which every edge of C corresponds to an $(\varepsilon^{1/12}, d', d'/2, 3d'/2d_{UW})$-superregular pair (which is therefore also $(\varepsilon^{1/25}, d', \zeta d', 2d'/d)$-superregular). Let G^* be the digraph obtained from G by deleting all the edges in $\mathcal{B}(C)_{PE}$.

Now consider any edge UW of R and recall from (Sch2) that $G[U,W]$ is ε-regular of density $d_{UW} \geq d - \varepsilon$. Since $G^*[U,W]$ is obtained from $G[U,W]$ by deleting at most $2d'm/d \leq \varepsilon m$ edges at every vertex, Proposition 4.3(i) with $d' := \varepsilon$ implies that $G^*[U,W]$ is still $2\sqrt{\varepsilon}$-regular of density at least $3d_{UW}/4$. Thus we can apply Lemma 4.10(i) to find a $((4\varepsilon)^{1/24}, d'/k, 2d'/d_{UW}k)$-regular spanning subgraph $G'[U,W]$ of $G^*[U,W]$ (which is therefore also $(\varepsilon^{1/25}, d'/k, 2d'/dk)$-regular). Let $\mathcal{B}(R)_{PE}$ be the union of all the $G'[U,W]$ over all edges UW of R. Then $\mathcal{B}(C)_{PE} \cup \mathcal{B}(R)_{PE}$ is a path system extender for C, R with parameters $(\varepsilon^{1/25}, d, d', \zeta)$. □

The following lemma implies that we can use the edges of a path system extender to extend the path systems Q_i obtained from Lemma 8.1 into Hamilton cycles.

Lemma 8.3. *Suppose that $0 < 1/n \ll d' \ll 1/k \ll \varepsilon \ll 1/\ell^* \ll d \ll \nu \ll \tau \ll \alpha, \theta \leq 1$, that $d \ll \zeta \leq 1/2$ and that $m/50 \in \mathbb{N}$. Let $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ be a consistent $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$-system with $|G| = n$. Let $\mathcal{P}'$ be a $(d')^2$-uniform 50-refinement of $\mathcal{P}$. Let PE be a path system extender with parameters $(\varepsilon, d, d', \zeta)$ for C, R. Let $s := 10^7/\nu^2$ and suppose that Q is a path system in G such that*

- *Q and PE are edge-disjoint;*
- *Q contains precisely one exceptional cover and no other original exceptional edges;*
- *Q has style size 5 (with respect to $\mathcal{P}'$);*
- *$|E(Q)| \leq 1230n/s$.*

Then G contains a Hamilton cycle D such that $Q \subseteq D \subseteq PE \cup Q$.

For the proof of Lemma 8.3, we add additional edges to Q to obtain a 'locally balanced' path system. (To ensure the existence of these additional edges we need to work with a consistent system rather than the simpler notion of a scheme.) This path system can then be extended into a 1-factor using matchings between consecutive clusters of C. Finally, we apply Lemma 6.4 to transform this 1-factor into a Hamilton cycle.

Proof. Recall that Q^{basic} is obtained from Q by replacing each path of the form x^-xx^+ (where $x \in V_0$) by the exceptional edge $x^-x^+ \in E(G^{\mathrm{basic}})$. For each edge $e = ab$ of Q^{basic} in turn, our aim is to apply Lemma 7.2 to find a chord sequence $CS(U(b), U(a)^+)$ in R which consists of at most $3/\nu$ edges. (Here $U(b)$ denotes the cluster in $\mathcal{P}$ containing b and $U(a)^+$ denotes the successor on C of the cluster in $\mathcal{P}$ containing a.) We say that a cluster in $\mathcal{P}$ is *full* if it is visited at least $m/110$ times by the interiors of the chord sequences chosen so far. (Recall that we disregard the first cluster $U(b)^-$ and the final cluster $U(a)^+$ of $CS(U(b), U(a)^+)$ when considering its interior.) Then the number of full clusters is at most

$$\frac{2|Q^{\mathrm{basic}}|(3/\nu)}{m/110} \leq \frac{811800n}{\nu s m} \leq \frac{\nu k}{4}.$$

After each application of Lemma 7.2, let $\mathcal{V}' \subseteq V(R)$ be the set of all those clusters which are now full. So we can apply Lemma 7.2 to R and $\mathcal{V}'$ as above to find a chord sequence $CS(U(b), U(a)^+)$ whose interior avoids the full clusters. Thus altogether the interiors of all the $CS(U(b),U(a)^+)$ (for all edges ab of $Q^{\rm basic}$) visit each cluster in $\mathcal{P}$ at most $m/110+3/\nu \le m/100$ times.

Note that since Q has style size 5, every cluster in $\mathcal{P}$ meets $Q^{\rm basic}$ in at most $5m/50=m/10$ vertices. Thus every cluster plays the role of $U(b)^-$ or $U(a)^+$ for at most $2m/10$ edges ab of $Q^{\rm basic}$. This implies that the union $S(Q)$ of all the chord sequences $CS(U(b),U(a)^+)$ (over all edges ab of $Q^{\rm basic}$) visits each cluster at most

$$m/100+2m/10=21m/100 \tag{8.2}$$

times (where we count the edges in $S(Q)$ with multiplicities).

Now for each edge E of R, let s_E be the number of times that E occurs in $S(Q)$. Note that if E occurs in $CS(U(b),U(a)^+)$ then at least one of the endclusters of E lies in the interior of $CS(U(b),U(a)^+)$. In particular, suppose that $E=UW$ occurs as an initial edge in $CS(U(b),U(a)^+)$. Then W lies in the interior of the sequence. Since altogether the interiors of all the $CS(U(b),U(a)^+)$ visit W at most $m/100$ times, this implies that altogether E can occur at most $m/100$ times as an initial edge of some $CS(U(b),U(a)^+)$. Similarly, suppose that $E=UW$ occurs in $CS(U(b),U(a)^+)$, but not as an initial edge. Then U lies in the interior of the sequence and altogether E can occur at most $m/100$ times in this way. It follows that $s_E\le 2\cdot m/100=m/50$.

Let $J\subseteq\{1,\dots,50\}$ be a set of size 5 so that Q has style J. Without loss of generality, we may assume that $J=\{1,\dots,5\}$. Let B_E be the bipartite subgraph of $\mathcal{B}(R)_{PE}$ which corresponds to E. Let B'_E be the induced bipartite subgraph of B_E of style $\{6,\dots,29\}$. (So if $E=UW$ then B'_E is the subgraph of B_E induced by $U(6)\cup\dots\cup U(29)$ and $W(6)\cup\dots\cup W(29)$, where $U(1),\dots,U(50)$ are the clusters in $\mathcal{P}'$ contained in U.)

For each edge E of R in turn, we choose a matching M_E of size s_E in B'_E such that M_E avoids all matchings chosen previously. To see that this can be done, recall from (PE2) that B_E is $(\varepsilon,d'/k,2d'/dk)$-regular. Since the size of the vertex classes of B'_E is precisely $24/50$ times the size of the vertex classes of B_E, this implies that B'_E is still $(3\varepsilon,d'/k,6d'/dk)$-regular. Thus Lemma 4.12(i) implies that B'_E contains a matching M'_E of size $(1-3\varepsilon)\cdot 24m/50\ge 23m/50$. But (8.2) implies that at most $2\cdot 21m/100=21m/50$ edges of M'_E are incident to an edge of a previously chosen matching (the extra factor 2 comes from the fact that there is a contribution from both endclusters of E). So we can indeed find the required matchings M_E.

Let Q_* consist of the edges of $Q^{\rm basic}$ together with the edges in the matchings M_E (for all edges E of R). So Q_* is a path system whose style is $\{1,\dots,29\}$.

Let Q^{in} denote the set of final vertices of the edges in Q_* and let Q^{out} denote the set of their initial vertices. For each cluster U in $\mathcal{P}$, let $U^2:=U\setminus Q^{in}$ and $U^1:=U\setminus Q^{out}$. Consider any edge UW on C. Then Proposition 7.3 implies that $|U^1|=|W^2|$. Let B^*_{UW} be the bipartite subgraph of $\mathcal{B}(C)_{PE}[U,W]$ induced by U^1 and W^2. Recall from (PE1) that $\mathcal{B}(C)_{PE}[U,W]$ is $(\varepsilon,d',\zeta d',2d'/d)$-superregular. Together with the facts that B^*_{UW} contains all vertices of $U\cup W$ of style $30,\dots,50$ and that $\mathcal{P}'$ is an $(d')^2$-uniform refinement of $\mathcal{P}$ this implies that B^*_{UW} is still $(3\varepsilon,d',\zeta d'/3,6d'/d)$-superregular. So Lemma 4.12(ii) implies that B^*_{UW} contains a perfect matching M_{UW}. The union of all the M_{UW} (over all edges UW of C) together with all the edges of Q_* forms a 1-factor F of $Q^{\rm basic}\cup PE$.

Since Q_* is a path system, each cycle of F has a vertex in U^1 for some cluster U in $\mathcal{P}$. Now apply Lemma 6.4 with $\mathcal{B}(C)_{PE}$, F, C, $E(C)$, U^1 and U^2 playing the roles of G, F, C, J, V_i^1 and V_i^2 in order to modify F into a Hamilton cycle D' of $G^{\rm basic}$. (Note that $U^1 \cap U^2 \neq \emptyset$ as they both contain all vertices of style $\{30, \dots, 50\}$.) Then Observation 7.4 implies that $D := (D')^{\rm orig}$ is a Hamilton cycle of $G = G^{\rm orig}$, as required. □

We obtain the following lemma by repeated applications of Lemma 8.1 and 8.3.

Lemma 8.4. *Suppose that $0 < 1/n \ll r/m \ll d' \ll 1/k \ll \varepsilon \ll 1/\ell^* \ll d \ll \nu \ll \tau \ll \alpha, \theta \leq 1$ and that $d \ll \zeta \leq 1/2$. Let $s := 10^7/\nu^2$. Suppose that $m/50, 50k/(s-1) \in \mathbb{N}$. Let $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ be a consistent $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$-system with $|G| = n$. Let $\mathcal{P}'$ be a $(d')^2$-uniform 50-refinement of $\mathcal{P}$. Let H be an r-factor of G. Let $CB_1, \dots, CB_r$ be r exceptional factors with parameters $(50, (s-1)/50)$ with respect to $C, \mathcal{P}'$. Let PE be a path system extender with parameters $(\varepsilon, d, d', \zeta)$ for C, R. Suppose that H, PE and the original versions of $CB_1, \dots, CB_r$ are pairwise edge-disjoint. Then G contains edge-disjoint Hamilton cycles $C_1, \dots, C_{sr}$ such that*

- (i) *altogether $C_1, \dots, C_{sr}$ contain all edges of $H \cup CB_1^{\rm orig} \cup \dots \cup CB_r^{\rm orig}$;*
- (ii) *each C_i lies in $PE \cup H \cup CB_1^{\rm orig} \cup \dots \cup CB_r^{\rm orig}$.*

Proof. Consider a 1-factorization $F_1, \dots, F_r$ of H (which exists by Proposition 6.1). So each F_j contains precisely one exceptional cover (and no other original exceptional edges). For each $j = 1, \dots, r$ apply Lemma 8.1 with F_j and CB_j playing the roles of H and CB to obtain path systems $Q_{j,1}, \dots, Q_{j,s}$ as described there. Relabel all these path systems as $Q_1, \dots, Q_{sr}$. Note that the Q_i form a decomposition of the edges of $H \cup CB_1^{\rm orig} \cup \dots \cup CB_r^{\rm orig}$.

Apply Lemma 8.3 to obtain a Hamilton cycle C_1 in G with $Q_1 \subseteq C_1 \subseteq PE \cup Q_1$. Repeat this for each of $Q_2, \dots, Q_{sr}$ in turn, with the subdigraph PE' of PE obtained by deleting all the edges in the Hamilton cycles found so far playing the role of PE. Note that at each stage we have removed at most $sr = 10^7 r/\nu^2 \leq \varepsilon^2 d'm/k$ outedges and at most $sr \leq \varepsilon^2 d'm/k$ inedges at each vertex of PE. So using Proposition 4.8 it is easy to check that PE' is still a path system extender with parameters $(2\varepsilon, d, d', \zeta/2)$. Thus we can indeed apply Lemma 8.3 in each step. □

8.3. **The preprocessing graph.** Let G be the digraph given in Theorem 1.2. As described at the beginning of Section 8.2, the purpose of Lemma 8.4 was to replace the leftover H of an almost decomposition of G by a digraph H' on $V(G) \setminus V_0$, i.e. by a digraph in which every exceptional vertex can be thought of as being isolated. This digraph H' is obtained from PE by deleting all the edges in the Hamilton cycles guaranteed by Lemma 8.4. But this means that H' will not be regular since PE was not regular. However, for later purposes we want to assume that this leftover H' is regular. So we will need to add another digraph $PG^\diamond$ to H' whose degree sequence complements that of PE. Then the union of H' and $PG^\diamond$ will be regular. So the preprocessing graph defined below contains all the digraphs which we need in order to replace H by a regular digraph on $V(G) \setminus V_0$ not containing any exceptional edges.

Suppose that $(G,\mathcal{P},R,C)$ is a (k,m,ε,d)-scheme and that $m/50, 50k/(s-1)\in\mathbb{N}$. Let $\mathcal{P}'$ be an ε-uniform 50-refinement of $\mathcal{P}$. A *preprocessing graph PG* with parameters $(s,\varepsilon,d,r',r'',r,\zeta)$ (with respect to $C,R,\mathcal{P}'$) is the edge-disjoint union of two graphs PG^* and $PG^\diamond$ satisfying the following conditions:

(PG1) PG^* is the edge-disjoint union of a path system extender PG_{PE} with parameters $(\varepsilon,d,r'/m,\zeta)$ (for C,R) and of r exceptional factors $CB_1,\dots,CB_r$ with parameters $(50,(s-1)/50)$ (with respect to $C,\mathcal{P}'$). Moreover, the original versions of these r exceptional factors are pairwise edge-disjoint.

(PG2) $PG^\diamond$ is a spanning subgraph of $G-V_0$ which satisfies $d^+_{PG^\diamond}(x)=r''-d^+_{PG^*}(x)$ and $d^-_{PG^\diamond}(x)=r''-d^-_{PG^*}(x)$.

Note that PG is a spanning r''-regular subgraph of $G^{\rm basic}$. Moreover, in $PG^{\rm orig}$ we have

$$d^\pm(x)=r(s-1)\ \ \forall x\in V_0 \qquad\text{and}\qquad d^\pm(y)=r''\ \ \forall y\in V(G)\setminus V_0. \tag{8.3}$$

Note also that (8.1) implies that

$$\Delta(PG^*)\le \frac{8r'}{d}+2r. \tag{8.4}$$

The following corollary is an immediate consequence of Lemma 8.4.

Corollary 8.5. *Suppose that $0<1/n\ll r/m\ll r'/m\ll r''/m\ll 1/k\ll\varepsilon\ll 1/\ell^*\ll d\ll\nu\ll\tau\ll\alpha,\theta\le 1$ and that $d\ll\zeta\le 1/2$. Let $s:=10^7/\nu^2$. Suppose that $m/50, 50k/(s-1)\in\mathbb{N}$. Let $(G,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ be a consistent $(\ell^*,k,m,\varepsilon,d,\nu,\tau,\alpha,\theta)$-system with $|G|=n$. Let $\mathcal{P}'$ be a $(r'/m)^2$-uniform 50-refinement of $\mathcal{P}$. Let H be an r-factor of G. Let PG be a preprocessing graph with respect to $C,R,\mathcal{P}'$ with parameters $(s,\varepsilon,d,r',r'',r,\zeta)$. Suppose that H and $PG^{\rm orig}$ are edge-disjoint. Then G contains edge-disjoint Hamilton cycles $C_1,\dots,C_{rs}$ such that the following conditions hold:*

(i) *Altogether $C_1,\dots,C_{rs}$ cover all edges of H and $C_i\subseteq H\cup PG^{\rm orig}$ for each $i=1,\dots,rs$.*
(ii) *Every vertex $x\in V_0$ is isolated in $PG':=PG^{\rm orig}\setminus E(C_1\cup\dots\cup C_{rs})$ and every vertex $x\in V(G)\setminus V_0$ has in- and outdegree $r''-(s-1)r$ in PG'.*

Proof. Apply Lemma 8.4 with PG_{PE} playing the role of PE. This gives us edge-disjoint Hamilton cycles $C_1,\dots,C_{rs}$ of G as described there. So in particular condition (i) of Corollary 8.5 holds. By (i) of Lemma 8.4, all the $C_1,\dots,C_{rs}$ together cover all the original exceptional edges of $PG^{\rm orig}$ (as they cover $CB_1^{\rm orig},\dots,CB_r^{\rm orig}$ and each original exceptional edge in $PG^{\rm orig}$ is contained in one of $CB_1^{\rm orig},\dots,CB_r^{\rm orig}$). Since V_0 is an independent set in G by (CSys9), it follows that every vertex $x\in V_0$ is isolated in PG'. Moreover, every vertex $x\in V(G)\setminus V_0$ has indegree $r+r''$ in $H\cup PG^{\rm orig}$ and indegree rs in $C_1\cup\dots\cup C_{rs}$. So the indegree of x in PG' is $r+r''-rs=r''-(s-1)r$. Similarly, every vertex in $V(G)\setminus V_0$ has outdegree $r''-(s-1)r$ in PG'. Altogether, this implies condition (ii) of Corollary 8.5. □

The next lemma shows that one can find a preprocessing graph within a consistent system.

Lemma 8.6. *Suppose that* $0 < 1/n \ll r/m \ll d' \ll d'' \ll 1/k \ll \varepsilon \ll \varepsilon' \ll 1/\ell^* \ll d \ll \nu \ll \tau \ll \alpha, \theta \leq 1$ *and that* $d \ll \zeta \leq 1/2$. *Let* $s := 10^7/\nu^2$. *Suppose that* $m/50, 50\ell^*/(s-1) \in \mathbb{N}$. *Let* $(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ *be a consistent* $(\ell^*, k, m, \varepsilon, d, \nu, \tau, \alpha, \theta)$*-system with* $|G| = n$. *Let* $\mathcal{P}'$ *be an* ε*-uniform 50-refinement of* $\mathcal{P}$. *Then* G^{basic} *contains a preprocessing graph with respect to* $C, R, \mathcal{P}'$ *with parameters* $(s, \varepsilon', d, d'm, d''m, r, \zeta)$.

Proof. Apply Lemma 7.6 to find r exceptional factors $CB_1, \dots, CB_r$ with parameters $(50, (s-1)/50)$ (with respect to $C, \mathcal{P}'$) whose original versions $CB_1^{\mathrm{orig}}, \dots, CB_r^{\mathrm{orig}}$ are pairwise edge-disjoint. (So we apply the lemma with $K := 50$, $L := (s-1)/50$ and $r_0 := r$.) Let G_1 be obtained from G by deleting all the edges in $CB_1^{\mathrm{orig}}, \dots, CB_r^{\mathrm{orig}}$. Thus G_1 is obtained from G by deleting at most $(s-1)r$ outedges and at most $(s-1)r$ inedges at each vertex. Since $(s-1)r \leq \varepsilon m$ we can apply Lemma 7.1 with ε playing the role of ε' in Lemma 7.1 to see that $(G_1, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ is still a consistent $(\ell^*, k, m, 3\sqrt{\varepsilon}, d, \nu/2, \tau, \alpha/2, \theta/2)$-system. Apply Lemma 8.2 to find a path system extender PG_{PE} in G_1 with parameters $(\varepsilon', d, d', \zeta)$ (with respect to C, R). Let PG^* be the union of PG_{PE} and $CB_1, \dots, CB_r$.

To find $PG^\diamond$, for every vertex $x \in V(G) \setminus V_0$ let $n_x^+ := d''m - d^+_{PG^*}(x)$ and $n_x^- := d''m - d^-_{PG^*}(x)$. Note that (8.4) implies that

$$\Delta(PG^*) \leq 8d'm/d + 2r \leq 9d'm/d. \tag{8.5}$$

So for every vertex $x \in V(G) \setminus V_0$ we have that

$$\left(1 - \frac{9d'}{dd''}\right) d''m = d''m - \frac{9d'm}{d} \leq n_x^+, n_x^- \leq d''m. \tag{8.6}$$

Let G_2 be the digraph obtained from $G - V_0$ by deleting all the edges in PG^*. Then (8.5) and the fact that G is a robust (ν, τ)-outexpander with $\delta^0(G) \geq \alpha n$ imply that G_2 is still a robust $(\nu/2, 2\tau)$-outexpander with $\delta^0(G_2) \geq \alpha n/2$. Together with (8.6) this shows that we can apply Lemma 5.2 with $q = 1$, with G_2 playing the role of $G = Q$, and with $d''m/|G_2|$, $9d'/dd''$ playing the roles of ξ, ε to obtain a spanning subgraph $PG^\diamond$ of G_2 satisfying $d^\pm_{PG^\diamond}(x) = n_x^\pm$ for every vertex x. Then $PG^* \cup PG^\diamond$ is a preprocessing graph as required. □

The following lemma decomposes the leftover PG' of the preprocessing graph obtained from Corollary 8.5 into path systems H_i. In Section 9, these path systems will be extended into Hamilton cycles using the 'chord absorber'.

Suppose that $k/g \in \mathbb{N}$ and that $C = V_1 \dots V_k$. Consider the canonical interval partition of C into g edge-disjoint intervals of equal length and for each $i = 1, \dots, g$ let X_i denote the union of all clusters in the ith interval. So $X_i = V_{(i-1)k/g+1} \cup \dots \cup V_{ik/g+1}$. We say that an edge of $G - V_0$ has *double-type* ij if its endvertices are contained in $X_i \cup X_j$. So the number of double-types is $\binom{g}{2}$. A digraph has *double-type* ij if all its edges have double-type ij.

Lemma 8.7. *Suppose that* $0 < 1/n \ll 1/k, \varepsilon, d, 1/q^*, 1/g \ll 1$ *and that* $2q^*/3g(g-1), k/g \in \mathbb{N}$. *Let* $(G, \mathcal{P}, R, C)$ *be a* (k, m, ε, d)*-scheme with* $|G| = n$ *and* $C = V_1 \dots V_k$.

Suppose that H is a 1-regular digraph on $V_1 \cup \dots \cup V_k$. Then we can decompose $E(H)$ into q^* (possibly empty) matchings $H_1, \dots, H_{q^*}$ such that the following conditions hold.

(i) *For all $i = 1, \dots, q^*$, H_i consists of at most $2g^2km/q^*$ edges.*
(ii) *If $|i-j| \leq 10$, then H_i and H_j are vertex-disjoint, with the indices considered modulo q^*.*
(iii) *Each H_i consists entirely of edges of the same double-type and for each $t \in \binom{g}{2}$ the number of H_i of double-type t is $q^*/\binom{g}{2}$.*

Proof. First split the edges of H into $2q^*/3g(g-1)$ sets whose sizes are as equal as possible. Now we arbitrarily split each of these sets into three matchings. Finally, we split each of these matchings further into submatchings consisting of edges of the same double-type. Since there are $\binom{g}{2}$ different double-types this gives $\binom{g}{2} \cdot 3 \cdot 2q^*/3g(g-1) = q^*$ matchings, which we denote by $H_1, \dots, H_{q^*}$. Moreover, each H_i consists of at most $2g^2km/q^*$ edges. Thus the H_i satisfy (i) and (iii).

Given a double-type ij, let $cl(ij)$ denote the set of all those numbers s for which at least one of $s-1, s, s+1$ belongs to $\{i, j\}$. (So $cl(23) = \{1, 2, 3, 4\}$ and $cl(27) = \{1, 2, 3, 6, 7, 8\}$.)

To obtain (ii), we first show the following claim: there is a cyclic ordering of the double-types so that if two double-types ab and cd have distance at most 10 in the ordering, then $cl(ab) \cap cl(cd) = \emptyset$.

To prove the claim, consider the following auxiliary graph A: the vertices are the double-types (i.e. the unordered pairs of numbers in $\{1, \dots, g\}$). We connect two double-types ab and cd by an edge if $cl(ab) \cap cl(cd) = \emptyset$. Then A has minimum degree at least $\binom{g}{2} - 6(g-1) \geq \frac{10}{11}\binom{g}{2}$. So A contains the 10th power of a Hamilton cycle by Theorem 6.3. The ordering of the double-types on the Hamilton cycle gives the required ordering of the double-types.

We now relabel the H_i as follows: first we take one H_i of each double-type in this ordering of the double-types, and then repeat with another H_i of each double-types and so on. □

9. The chord absorbing step

Recall from the previous section that we have a 'leftover' digraph G' which is a regular subdigraph of $G - V_0$. (G' is a subdigraph of the preprocessing graph.) The aim of this section is to define and use a 'chord absorber' CA in G^{basic} so that

(a) $G' \cup CA$ contains a collection of Hamilton cycles C_i which together cover all edges of G';
(b) removing the C_i from CA leaves a digraph which is a blow-up of the cycle $C = V_1 \dots V_k$;
(c) each C_i contains exactly one complete exceptional sequence and thus C_i^{orig} corresponds to a Hamilton cycle in G.

Similarly to the previous section, we first split G' into 1-factors. Then we split each such 1-factor H into small matchings H_i (as described in Lemma 8.7). Then each H_i is extended into a Hamilton cycle C_i using edges of the chord absorber CA. This is achieved mainly by Lemma 9.5. The main difficulty compared to the argument in the previous section is that we need to achieve (b).

The chord absorber consists of a blow-up of the cycle $C = V_1 \dots V_k$, some exceptional factors as well as some additional edges which are constructed via a 'universal walk'. These additional edges will be used to 'balance out' the edges in the H_i. The universal walk U will be constructed in the next subsection using the chord sequences defined in Section 7.2.

It turns out that a natural way to construct the universal walk U and CA would be the following (where we ignore requirement (c) for the moment): for each pair V_i, V_{i+1} of consecutive clusters on the cycle C, we fix a shifted walk SW_i from V_i to V_{i+1} (recall these were also defined in Section 7.2). We then let U be the concatenation of all the SW_i with $1 \le i \le k$. Then it is not hard to check that U is a closed walk which visits each cluster the same number of times. CA is then defined to be the union of a regular blow-up $\mathcal{B}(C)$ of C (which is also ε-regular) together with a regular blow-up $\mathcal{B}(U)$ of U. As a step towards extending each H_i into a Hamilton cycle, we balance out each edge $e = x_j x_{j'}$ of H_i by adding a suitable shifted walk $SW_{jj'}$ (see also the proof of Lemma 8.3 for a similar argument). We do this as follows: for any j, let $s(j)$ be such that $x_j \in V_{s(j)}$. Then we add the shifted walk $SW_{jj'}$ consisting of the concatenation $SW_{s(j')} SW_{s(j')+1} \dots SW_{s(j)-1} SW_{s(j)}$. If we replace each edge E of each $SW_{jj'}$ used for H_i with an edge from the bipartite subgraph of $\mathcal{B}(U)$ corresponding to E, then the union of these edges together with H_i satisfy a 'local balance property' (as described in Proposition 7.3) and can thus be extended into a Hamilton cycle using edges of $\mathcal{B}(C)$. The crucial fact now is that since H is a 1-factor, it turns out that altogether (i.e. when considering the union of the H_i), each SW_j is used the same number of times in this process. So for each edge E of U, overall we use the same number t of edges from the bipartite subgraph of $\mathcal{B}(U)$ corresponding to E. Thus t is independent of E. This means that we can indeed choose $\mathcal{B}(U)$ to be regular, which will enable us to satisfy (b).

Unfortunately, the above construction of U makes U too long – U would visit each cluster more than k times, which would create major technical difficulties in the proof of Lemma 9.5. So in Lemma 9.1 we present a 'compressed' construction (based on chord sequences) which has the same set of chord edges as the one given above, but which visits each cluster only ℓ' times, where $1/k \ll \varepsilon \ll 1/\ell'$.

9.1. **Universal walks, setups and chord absorbers.** Suppose that R is a digraph whose vertices are k clusters $V_1, \dots, V_k$ and that $C := V_1 \dots V_k$ is a Hamilton cycle in R. A closed walk U in R is a *universal walk for C with parameter ℓ'* if the following conditions hold:

- (U1) For every $i = 1, \dots, k$ there is a chord sequence $ECS(V_i, V_{i+1})$ from V_i to V_{i+1} such that U contains all edges of all these chord sequences (counted with multiplicities) and all remaining edges of U lie on C.
- (U2) Each $ECS(V_i, V_{i+1})$ consists of at most $\sqrt{\ell'}/2$ edges.
- (U3) U enters every cluster V_i exactly ℓ' times and it leaves every cluster V_i exactly ℓ' times.

We will often view U as a multidigraph. Whenever U is a universal walk for C with parameter ℓ', then $ECS(V_i, V_{i+1})$ will always refer to the chord sequence from V_i

to V_{i+1} which is contained in U. We will call $ECS(V_i, V_{i+1})$ an *elementary chord sequence from* V_i *to* V_{i+1} and the edges in $ECS(V_1, V_2) \cup \dots \cup ECS(V_k, V_1)$ the *chord edges of* U.

Note that condition (U1) means that if an edge $V_iV_j \in E(R) \setminus E(C)$ occurs in total 5 times (say) in $ECS(V_1, V_2), \dots, ECS(V_k, V_1)$ then it occurs precisely 5 times in U. We will identify each occurrence of V_iV_j in $ECS(V_1, V_2), \dots, ECS(V_k, V_1)$ with a (different) occurrence of V_iV_j in U. Note that the edges of $ECS(V_i, V_{i+1})$ are allowed to appear in a different order within $ECS(V_i, V_{i+1})$ and within U.

Suppose that F is a chord edge of U and that F' is the next chord edge of U. We let $P(F)$ denote the subwalk of U from F to F' (without the edges F and F'). So $P(F)$ contains no chord edges and if F_1 and F_2 are two occurrences of the same chord edge on U then $P(F_1)$ and $P(F_2)$ might be different from each other. We say that the edges in $P(F)$ are the *cyclic edges associated with* F. The *augmented elementary chord sequence* $AECS(V_i, V_{i+1})$ consists of all edges F in the elementary chord sequence $ECS(V_i, V_{i+1})$ together with all the edges in the corresponding subwalks $P(F)$. We order the edges of $AECS(V_i, V_{i+1})$ by taking the ordered sequence $ECS(V_i, V_{i+1})$ and by inserting the edges of $P(F)$ (in their order on U) after each edge $F \in ECS(V_i, V_{i+1})$. Thus the collection of all augmented elementary chord sequences $AECS(V_i, V_{i+1})$ forms a partition of the edges of U into k parts. But $AECS(V_i, V_{i+1})$ might not necessarily be connected (i.e. it might not form a walk in R).

Lemma 9.1. *Suppose that $0 < 1/k \ll \nu \ll \tau \ll \alpha < 1$. Suppose that R is a robust (ν, τ)-outexpander whose vertices are k clusters $V_1, \dots, V_k$ and $\delta^0(R) \geq \alpha k$. Let $C := V_1 \dots V_k$ be a Hamilton cycle in R. Then there exists a universal walk U for C with parameter $\ell' := 36/\nu^2$.*

Proof. Let $\mathcal{P} := \{V_1, \dots, V_k\}$. Our first aim is to choose the elementary chord sequences $ECS(V_j, V_{j+1})$ greedily in such a way that they satisfy (U2). Suppose that we have already chosen $ECS(V_1, V_2), \dots, ECS(V_{j-1}, V_j)$ such that each of them contains at most $3/\nu = \sqrt{\ell'}/2$ edges and together the interiors of these elementary chord sequences visit every cluster in $\mathcal{P}$ at most $2\ell'/3 + 3/\nu$ times. (Here the number of visits for a cluster V_i is the sum of the number of entries into V_i and the number of exits from V_i.) We say that a cluster in $\mathcal{P}$ is *full* if it is visited at least $2\ell'/3$ times by the interiors of the previously chosen chord sequences. Since each elementary chord sequence contains at most $3/\nu$ edges, it follows that the number of full clusters is at most $2 \cdot (3j/\nu)/(2\ell'/3) \leq 9k/\nu\ell' = \nu k/4$. Let $\mathcal{V}'$ denote the set of clusters in $\mathcal{P}$ which are full. Then we can apply Lemma 7.2 to obtain an (elementary) chord sequence $ECS(V_j, V_{j+1})$ which has at most $3/\nu$ edges and whose interior avoids $\mathcal{V}'$. We continue in this way to choose $ECS(V_1, V_2), \dots, ECS(V_k, V_1)$ such that together their interiors visit every cluster in $\mathcal{P}$ at most $2\ell'/3 + 3/\nu$ times and let U^* be the union of these chord sequences. Then U^* visits every cluster in $\mathcal{P}$ at most $2\ell'/3 + 3/\nu + 2 \leq 3\ell'/4$ times. Thus U^* satisfies (U2) and the first part of (U1).

Our next aim is to add further edges to U^* so that we will be able to satisfy (U3) and the second part of (U1). For each $j = 1, \dots, k$, let n_j^{out} denote the number of edges of U^* which leave the cluster V_j and let n_j^{in} denote the number of edges of U^*

which enter the cluster V_j. We claim that for each j, we have $n_j^{\rm out} = n_{j+1}^{\rm in}$. To see this, suppose that VW is an edge of an elementary chord sequence which is not its final edge. Then the next edge of this elementary chord sequence leaves the cluster W^- preceding W on C. If VW is the final edge of the elementary chord sequence, then the first edge of the next elementary chord sequence will leave W^-. This proves the claim.

Now let $\ell_j := \ell' - 1 - n_j^{\rm in}$ for each $j = 1, \dots, k$. Note that $\ell_j > 0$. Let $U^\diamond$ be obtained from U^* by adding exactly ℓ_j copies of the edge $V_{j-1}V_j$ for all j. The above claim implies that $U^\diamond$ is $(\ell' - 1)$-regular.

Finally we add another copy of each edge of C to $U^\diamond$ and denote the resulting multidigraph by U. So now U satisfies (U1)–(U3). It remains to show that the edges in U can be ordered so that the resulting sequence forms a (connected) closed walk in R. To see this, note that since $U^\diamond$ is an $(\ell' - 1)$-regular multidigraph, it has a decomposition into 1-factors by Proposition 6.1. We order the edges of U as follows: We first traverse all cycles of the 1-factor decomposition of $U^\diamond$ which contain the cluster V_1. Next, we traverse the edge V_1V_2 of C. Next we traverse all those cycles of the 1-factor decomposition which contain V_2 and which have not been traversed so far. Next we traverse the edge V_2V_3 of C and so on until we reach V_1 again. This completes the construction of U. □

$(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is called an $(\ell', k, m, \varepsilon, d)$-*setup* if $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme and the following conditions hold:

(ST1) U is a universal walk for $C = V_1 \dots V_k$ with parameter ℓ' and $\mathcal{P}'$ is an ε-uniform ℓ'-refinement of $\mathcal{P}$.

(ST2) Let $V_j^1, \dots, V_j^{\ell'}$ denote the clusters in $\mathcal{P}'$ which are contained in V_j (for each $j = 1, \dots, k$). Then U' is a closed walk on the clusters in $\mathcal{P}'$ which is obtained from U as follows: When U visits V_j for the ath time, we let U' visit the subcluster V_j^a (for all $a = 1, \dots, \ell'$).

(ST3) Each edge of U' corresponds to an $[\varepsilon, \geq d]$-superregular pair in G.

Since U visits every cluster in $\mathcal{P}$ precisely ℓ' times, it follows that U' visits every cluster in $\mathcal{P}'$ exactly once. So U' can be viewed as a Hamilton cycle on the clusters in $\mathcal{P}'$. We call U' the *universal subcluster walk* (with respect to C, U and $\mathcal{P}'$).

Given a digraph T whose vertices are clusters, an (ε, r)-*blow-up* $\mathcal{B}(T)$ *of* T is obtained by replacing each vertex V of T with the vertices in the cluster V and replacing each edge VW of T with a bipartite graph $\mathcal{B}(T)[V, W]$ with vertex classes V and W which satisfies the following three properties:

- $\mathcal{B}(T)[V, W]$ is ε-regular.
- All the edges in $\mathcal{B}(T)[V, W]$ are oriented towards the vertices in W.
- The underlying undirected graph of $\mathcal{B}(T)[V, W]$ is r-regular.

An r-*blow-up of* T is defined similarly: we do not require the bipartite graphs to be ε-regular.

We say that a digraph CA on $V(G) \setminus V_0 = V_1 \cup \dots \cup V_k$ is a *chord absorber for* C, U' *with parameters* $(\varepsilon, r, r', r'', q, f)$ if CA is the union of two digraphs $\mathcal{B}(C)$ and $\mathcal{B}(U')$ on $V(G) \setminus V_0$ satisfying conditions (CA1)–(CA3) below. In (CA1), $\mathcal{P}^*$ is a

(q/f)-refinement of $\mathcal{P}$. For (CA2), recall from (ST2) that $\mathcal{P}'$ denotes the partition whose clusters correspond to the vertices of U'.

(CA1) $\mathcal{B}(C)$ is the union of $\mathcal{B}(C)^*$ and $CA^{\rm exc}$. $\mathcal{B}(C)^*$ is an (ε, r)-blow-up of C. $CA^{\rm exc}$ consists of r'' exceptional factors with parameters $(q/f, f)$ (with respect to C, $\mathcal{P}^*$) whose original versions are pairwise edge-disjoint.

(CA2) $\mathcal{B}(U')$ is an r'-blow-up of U'. Moreover, $\mathcal{B}(U')$ has the following stronger property: for every cluster A in $\mathcal{P}'$ there is a partition $A_1, \dots, A_4$ of A into sets of equal size such that for every edge AB of U' and each $j = 1, \dots, 4$ there are r' edge-disjoint perfect matchings between A_j and B_j such that $\mathcal{B}(U')[A, B]$ is the union of all these $4r'$ matchings.

(CA3) $\mathcal{B}(C)^*$, $\mathcal{B}(U')$ and the original version of $CA^{\rm exc}$ are pairwise edge-disjoint subdigraphs of G.

Thus CA is a $(r + r' + r'')$-regular subdigraph of $G^{\rm basic}$. However, in the original version $CA^{\rm orig} = \mathcal{B}(C)^* \cup \mathcal{B}(U') \cup (CA^{\rm exc})^{\rm orig}$ of CA we have

$$d^{\pm}(x) = r''q \ \ \forall x \in V_0 \qquad \text{and} \qquad d^{\pm}(y) = r + r' + r'' \ \ \forall y \in V(G) \setminus V_0. \tag{9.1}$$

9.2. **Bi-universal walks, bi-setups and chord absorbers.** Suppose that R is a digraph whose vertices are k clusters $V_1, \dots, V_k$, where k is even, and that $C := V_1 \dots V_k$ is a Hamilton cycle in R. Let $\mathcal{V}_{\rm even}$ denote the set of all those clusters V_i for which i is even and define $\mathcal{V}_{\rm odd}$ similarly. We will now define a bi-universal walk, which is an analogue of a universal walk for a bipartite setting. The difference to Section 9.1 is that now we only assume the existence of a chord sequence from V to V' whenever $V, V' \in \mathcal{V}_{\rm even}$ or $V, V' \in \mathcal{V}_{\rm odd}$. Roughly speaking, if H is a bipartite graph whose vertex classes are $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$, U is a bi-universal walk and U' is a bi-universal subcluster walk, then a chord absorber for C, U' can still absorb all edges of H. (Note that C is also bipartite with vertex classes $\mathcal{V}_{\rm even}$ and $\mathcal{V}_{\rm odd}$.) This will only be used in [14] and not in this paper.

A closed walk U in R is a *bi-universal walk for C with parameter ℓ'* if the following conditions hold:

(BU1) The edge set of U has a partition into $U_{\rm odd}$ and $U_{\rm even}$. For every $i = 1, \dots, k$ there is a chord sequence $ECS^{bi}(V_i, V_{i+2})$ from V_i to V_{i+2} such that $U_{\rm even}$ contains all edges of all these chord sequences for even i (counted with multiplicities) and $U_{\rm odd}$ contains all edges of these chord sequences for odd i. All remaining edges of U lie on C.

(BU2) Each $ECS^{bi}(V_i, V_{i+2})$ consists of at most $\sqrt{\ell'}/2$ edges.

(BU3) $U_{\rm even}$ enters every cluster V_i exactly $\ell'/2$ times and it leaves every cluster V_i exactly $\ell'/2$ times. The same assertion holds for $U_{\rm odd}$.

Whenever U is a bi-universal walk for C with parameter ℓ', then $ECS^{bi}(V_i, V_{i+2})$ will always refer to the chord sequence from V_i to V_{i+2} which is contained in U. As before, we will call $ECS^{bi}(V_i, V_{i+2})$ an *elementary chord sequence from V_i to V_{i+2}* and the edges in $ECS^{bi}(V_1, V_3) \cup ECS^{bi}(V_2, V_4) \cup \dots \cup ECS^{bi}(V_k, V_2)$ the *chord edges of U*.

If F is a chord edge of U then $P(F)$ is defined as in Section 9.1 and we again call the edges in $P(F)$ are the *cyclic edges associated with F*. The *augmented elementary chord sequence* $AECS^{bi}(V_i, V_{i+2})$ consists of all edges F in the elementary chord sequence $ECS^{bi}(V_i, V_{i+2})$ together with all the edges in the corresponding subwalks $P(F)$. So similarly as in Section 9.1, the collection of all augmented elementary chord sequences $AECS^{bi}(V_i, V_{i+2})$ forms a partition of the edges of U into k parts.

We define an $(\ell', k, m, \varepsilon, d)$-*bi-setup* $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ similarly as an $(\ell', k, m, \varepsilon, d)$-*setup*, the only difference is that U is a bi-universal walk for C (rather than a universal walk). We call U' the *bi-universal subcluster walk* (with respect to C, U and $\mathcal{P}'$). A *chord absorber for C, U' with parameters* $(\varepsilon, r, r', r'', q, f)$ is defined analogously as before.

9.3. **Finding chord absorbers.** The first lemma of this subsection states that if one is given a setup and one deletes a few edges at every vertex, then one still has a setup with slightly worse parameters.

Lemma 9.2. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \le \varepsilon' \ll d \ll 1/\ell' \ll 1$. Let $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ be an $(\ell', k, m, \varepsilon, d)$-setup with $|G| = n$. Let G' be a digraph obtained from G by deleting at most $\varepsilon' m$ outedges and at most $\varepsilon' m$ inedges at every vertex of G. Then $(G', \mathcal{P}, \mathcal{P}', R, C, U, U')$ is still a $(\ell', k, m, (\varepsilon')^{1/3}, d)$-setup. The analogue holds if $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-bi-setup.*

Proof. We only consider the case when $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-setup. The argument for bi-setups is identical. By Lemma 7.1(ii) $(G', \mathcal{P}, R, C)$ is still a $(k, m, 3\sqrt{\varepsilon'}, d)$-scheme. Moreover, (ST1) and (ST2) clearly still hold. So we only need to check that (ST3) still holds with ε replaced by $(\varepsilon')^{1/3}$. But since the clusters in $\mathcal{P}'$ have size m/ℓ', Proposition 4.3(iii) implies that each edge of U' still corresponds to a $[2\sqrt{\varepsilon'\ell'}, \ge d]$-superregular pair in G' (and thus to an $[(\varepsilon')^{1/3}, \ge d]$-superregular pair). □

The next lemma asserts that we can find a blow-up of the cycle C and of the universal subcluster walk U' within a setup, so that each edge of C corresponds to a graph which is both regular and superregular and each edge of U' corresponds to a regular graph.

Lemma 9.3. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll \varepsilon' \ll d \ll 1/\ell' \ll 1$, that $r_0/m, r_0'/m \ll d$ and that $m/4\ell' \in \mathbb{N}$. If $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-setup with $|G| = n$, then $G - V_0$ contains edge-disjoint spanning subdigraphs $\mathcal{B}(C)$ and $\mathcal{B}(U')$ such that*

(i) *$\mathcal{B}(C)$ is an (ε', r_0)-blow-up of C;*
(ii) *$\mathcal{B}(U')$ is an r_0'-blow-up of U' which satisfies (CA2) with r_0' playing the role of r'.*

The analogue holds if $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-bi-setup.

Proof. Choose new constants ε^* and ε_1 with $\varepsilon \ll \varepsilon^* \ll \varepsilon'$ and $\varepsilon^*, r_0/m, r_0'/m \ll \varepsilon_1 \ll d$. For each edge VW of C let d_{VW} denote the density of $G[V, W]$. (So $d_{VW} \ge d - \varepsilon$.) Apply Lemma 4.10(iv) with

$$d' := \frac{r_0/m}{1 - 12 \cdot \varepsilon^*} + \varepsilon^*$$

to each edge VW of C to obtain a spanning subgraph $G'[V, W]$ of $G[V, W]$ which is $[\varepsilon^*, d']$-superregular. Now apply Lemma 4.6 to obtain a spanning r_0-regular subgraph

$G''[V,W]$ of $G'[V,W]$ which is also $[\varepsilon', r_0/m]$-superregular. Let $\mathcal{B}(C)$ be the union of all the $G''[V,W]$ over all edges VW of C. Then $\mathcal{B}(C)$ is an (ε', r_0)-blow-up of C.

To construct $\mathcal{B}(U')$ we now proceed as follows. First we apply Lemma 4.7(i) with $\ell = 4$ to obtain a partition of each cluster A in $\mathcal{P}'$ into subclusters $A_1 \dots A_4$ such that, for every edge AB of U', the pair $G[A_j, B_j]$ is $[\varepsilon^*, \geq d]$-superregular. Let G_1 be the digraph obtained from G by deleting every edge in $\mathcal{B}(C)$. Thus G_1 is obtained from G by deleting r_0 outedges and r_0 inedges at every vertex in $V(G)\setminus V_0$ (and no edges at the vertices in V_0). Since the subclusters A_j have size $m/4\ell'$ and since $r_0 \leq \varepsilon_1^3 m/4\ell'$, this means that $G_1[A_j, B_j]$ is still $[\varepsilon_1, \geq d]$-superregular by Proposition 4.3(iii). For every edge AB of U' and for all $j = 1, \dots, 4$, we choose r_0' edge-disjoint perfect matchings in $G_1[A_j, B_j]$ (recall that these are guaranteed by Proposition 4.14). At each stage we delete the edges in all the matchings chosen so far before we choose the next matching. Since $r_0' \leq \varepsilon_1 m/4\ell'$, this means that the leftover of $G_1[A_j, B_j]$ will always be $[2\sqrt{\varepsilon_1}, \geq d]$-superregular by Proposition 4.3(iii). So we can choose the next matching. The union $\mathcal{B}(U')$ of all these perfect matchings (over all edges of U') is an r_0'-blow-up of U' which satisfies (CA2).

The proof for bi-setups is identical. □

The next lemma is an immediate consequence of Lemma 9.3. Given suitable exceptional factors, it guarantees a chord absorber within a setup which contains these exceptional factors. Since Lemma 7.6 implies the existence of such exceptional factors in a consistent system, this will enable us to find a chord absorber.

Lemma 9.4. *Suppose that $0 < 1/n \ll 1/k, qr_0''/m \ll \varepsilon \ll \varepsilon' \ll d \ll 1/\ell' \ll 1$, that $r_0/m, r_0'/m \ll d$ and that $q/f, fm/q, m/4\ell' \in \mathbb{N}$. Let $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ be an $(\ell', k, m, \varepsilon, d)$-setup with $|G| = n$. Suppose that $\mathcal{P}^*$ is a (q/f)-refinement of $\mathcal{P}$ and $EF_1, \dots, EF_{r_0''}$ are exceptional factors with parameters $(q/f, f)$ with respect to C, $\mathcal{P}^*$ whose original versions are pairwise edge-disjoint. Then there is a chord absorber CA for C, U' in G having parameters $(\varepsilon', r_0, r_0', r_0'', q, f)$ such that $CA^{\rm exc} = EF_1 \cup \dots \cup EF_{r_0''}$. The analogue holds if $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-bi-setup.*

Proof. We only consider the case when $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-setup. The proof for bi-setups is identical. Let G_1 be the digraph obtained from G by deleting every edge in $(CA^{\rm exc})^{\rm orig}$. Thus G_1 is obtained from G by deleting r_0'' outedges and r_0'' inedges at every vertex in $V(G) \setminus V_0$ and by deleting qr_0'' outedges and qr_0'' inedges at every vertex in V_0. Since $qr_0''/m \leq \varepsilon$, Lemma 9.2 implies that $(G_1, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is still a $(\ell', k, m, \varepsilon^{1/3}, d)$-setup. So we can apply Lemma 9.3 to find edge-disjoint subgraphs $\mathcal{B}(C)^*$ and $\mathcal{B}(U')$ in G_1 as guaranteed by (i) and (ii). Let $CA^{\rm exc} := EF_1 \cup \dots \cup EF_{r_0''}$. We then take $CA := \mathcal{B}(C)^* \cup \mathcal{B}(U') \cup CA^{\rm exc}$. □

9.4. **Absorbing chords into a blown-up Hamilton cycle.** The following lemma contains the key statement of this section (and of the entire proof of Theorem 1.2). Let $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ be a $(\ell', k, m, \varepsilon, d)$-setup. Suppose we are given a 1-factor H of $G - V_0$ which is split into small matchings H_i and that we are given complete exceptional path systems H_i'' which avoid the H_i (see conditions (b)–(d)). The lemma states that we can extend each path system $H_i' := H_i \cup H_i''$ into a Hamilton cycle C_i using edges from the chord absorber CA. The crucial point is that the set of edges we are allowed to use for C_i from $\mathcal{B}(U')$ is predetermined, i.e. this set does not depend on H. More precisely, for each 1-factor H we will split off a regular digraph $\mathcal{B}'(U')$ from $\mathcal{B}(U')$. The Hamilton cycles guaranteed by Lemma 9.5 cover $H \cup \mathcal{B}'(U')$, but no other edges of $\mathcal{B}(U')$. In Lemma 9.7 we will use this property to ensure that when we cover the leftover of the preprocessing graph from the previous section with edge-disjoint Hamilton cycles, we use all edges of $\mathcal{B}(U')$ in the process. So, as mentioned earlier, the leftover from the chord absorbing step is a subdigraph of $\mathcal{B}(C)$.

The basic strategy is similar to that of Lemma 8.3: first we balance out the edges of each H_i by adding edges corresponding to (augmented) chord sequences (see Claims 1 and 4) in order to obtain path systems W_i'' containing H_i. As mentioned above, we will use all edges of $\mathcal{B}'(U')$ in this process. Claim 2 is a step towards this – it implies that the set of edges we added to the H_i in the above step are themselves 'globally balanced' in the sense that we use the same number from each bipartite graph corresponding to an edge of U'. (Note however that we achieve this 'global balance' property only when considering all the H_i together – it need not hold when we consider H_1 on its own say.) Then we extend the path system W_i'' into a 1-factor F_i. F_i is then transformed into a Hamilton cycle C_i using Lemma 6.5 (see Claim 5).

Moreover, Lemma 9.5 also works for $(\ell', k, m, \varepsilon, d)$-bi-setups, as long as H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$.

Lemma 9.5. *Suppose that* $0 < 1/n \ll 1/k, 1/q \ll \varepsilon \ll \phi, \varepsilon' \ll r_1/m \ll d \ll 1/\ell' \ll 1$ *and that* $m/4\ell' \in \mathbb{N}$. *Let*

$$r_2 := 12\phi\ell' q. \tag{9.2}$$

Suppose that $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ *is a* $(\ell', k, m, \varepsilon, d)$*-setup with* $|G| = n$ *and* $C = V_1 \dots V_k$. *Let* $\mathcal{B}(C)^*$ *be a blow-up of* C *such that every edge of* C *corresponds to an* $(\varepsilon', r_1/m)$*-superregular pair in* $\mathcal{B}(C)^*$. *Let* $\mathcal{B}'(U')$ *be an* r_2*-blow-up of* U' *which satisfies the following condition:*

(a) *For every cluster* A *in* $\mathcal{P}'$ *there is a partition* $A_1, \dots, A_4$ *of* A *into sets of equal size such that for every edge* AB *of* U' *and each* $j = 1, \dots, 4$ *there are* r_2 *edge-disjoint perfect matchings* $S_1^j(AB), \dots, S_{r_2}^j(AB)$ *between* A_j *and* B_j *such that* $\mathcal{B}'(U')[A, B]$ *is the union of all these* $4r_2$ *matchings.*

Suppose that H *is a* 1*-factor of* $G - V_0$ *and that* $H_1, \dots, H_q$ *is a partition of* H *into matchings which satisfy the following properties:*

(b) *For each* $i = 1, \dots, q$ *there is a complete exceptional path system* H_i'' *(with respect to* C*) which is vertex-disjoint from* H_i.
(c) *Write* $H_i' := H_i \cup H_i''$ *for* $i = 1, \dots, q$. *Then for all* $i = 1, \dots, q$ *the original versions* $(H_i')^{\rm orig} = H_i \cup (H_i'')^{\rm orig}$ *of* H_i' *are pairwise edge-disjoint, each* H_i' *consists of at most* ϕm *paths and* $|H_i' \cap V_j| \le \phi m$ *for every cluster* V_j *in* $\mathcal{P}$. *Moreover,* H_i' *and* H_j' *are pairwise vertex-disjoint whenever* $|i - j| \le 10$.
(d) $\mathcal{B}(C)^*$, $\mathcal{B}'(U')$ *and the original version of* $H' := H_1' \cup \dots \cup H_q'$ *are pairwise edge-disjoint subdigraphs of* G.

Then there are edge-disjoint Hamilton cycles $C_1, \dots, C_q$ *in* G *such that the following properties hold:*

- *$H_i' \subseteq C_i^{\rm basic}$ for all $i = 1, \dots, q$.*
- *All the $C_1, \dots, C_q$ together cover all the edges of $(H')^{\rm orig} \cup \mathcal{B}'(U')$ and all remaining edges in $C_1, \dots, C_q$ are contained in $\mathcal{B}(C)^*$.*

The analogue holds for an $(\ell', k, m, \varepsilon, d)$-bi-setup $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ if we assume in addition that H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$ (where $\mathcal{V}_{\rm even}$ is the set of all those V_i such that i is even and $\mathcal{V}_{\rm odd}$ is defined analogously).

Proof. Recall that $C = V_1 \dots V_k$. So $V_1, \dots, V_k$ are the clusters in $\mathcal{P}$, $|V_j| = m$ for each $j = 1, \dots, k$ and each cluster in $\mathcal{P}'$ has size

$$m' := m/\ell'. \tag{9.3}$$

Given a vertex x of $G - V_0$, we will write $V(x)$ for the cluster in $\mathcal{P}$ containing x and $V(x)^+$ for the successor of $V(x)$ on C. For each $i = 1, \dots, q$ in turn, we will find a Hamilton cycle C_i in G which contains the original version of H_i'. So consider any i. We will first add suitable edges of R to H_i to form a sequence W_i' of edges which is 'locally balanced'. So W_i' will consist both of edges of H and edges of R. When constructing the Hamilton cycle C_i, we will replace each occurrence of an edge from R in W_i' by an edge in the corresponding bipartite subgraph of G.

We will first consider the case when $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is a $(\ell', k, m, \varepsilon, d)$-setup. Recall that the augmented elementary chord sequence $AECS(V_a, V_{a+1})$ was defined in Section 9.1. Given clusters V_j and $V_{j'}$, the *augmented chord sequence $ACS(V_j, V_{j'})$ in U* from V_j to $V_{j'}$ is the (ordered) sequence defined as

$$ACS(V_j, V_{j'}) := AECS(V_j, V_{j+1}) \cup AECS(V_{j+1}, V_{j+2}) \cup \dots \cup AECS(V_{j'-1}, V_{j'});$$

where the indices are modulo k. If $j = j'$ we take $ACS(V_j, V_{j'}) := \emptyset$. Let W_i' be obtained from H_i by including the augmented chord sequence $ACS(V(y), V(x)^+)$ after each edge xy in H_i. Ordering the edges in H_i gives us an ordering of the edges of W_i'. Suppose for example that xy is an edge of H_i with x, y both contained in the cluster V. Then we include exactly $AECS(V, V^+)$. If $x \in V$ and $y \in V^+$, where V^+ is the successor of V on C, then we do not include any edge (apart from xy itself).

Claim 1. 'Sequences are locally balanced.' *For each cluster V_j in $\mathcal{P}$ and each $i = 1, \dots, q$, the number of edges of W_i' leaving V_j equals the number of edges of W_i' entering V_{j+1}.*

Here (and below) multiple occurrences of edges in W_i' are considered separately, i.e. if the edge AB appears u times in W_i', then we count it u times in Claim 1. Moreover, if xy is an edge of H_i whose endvertices lie in the same cluster V, then (here and below) we count xy both as an edge leaving V and an edge entering V.

To prove Claim 1, consider any edge xy of H_i. Recall that each augmented elementary chord sequence $AECS(V_j, V_{j+1})$ consists of chord edges (namely those in the elementary chord sequence $ECS(V_j, V_{j+1})$ corresponding to $AECS(V_j, V_{j+1})$) and of cyclic edges (namely those in the subwalks $P(F)$ which were added to $ECS(V_j, V_{j+1})$ in order to obtain $AECS(V_j, V_{j+1})$). But every cyclic edge, joining V_j to V_{j+1} say, contributes both to the number of edges leaving V_j and to the number of edges entering V_{j+1}. On the other hand, in the (cyclic) sequence

$$xy \cup ECS(V(y), V(y)^+) \cup ECS(V(y)^+, V(y)^{++}) \cup \dots \cup ECS(V(x), V(x)^+)$$

obtained from $xy \cup ACS(V(y), V(x)^+)$ by deleting all cyclic edges in the augmented elementary chord sequences, every edge entering some cluster V_{j+1} is followed by an edge leaving V_j. (This is essentially the same observation as Proposition 7.3.) Altogether this shows that for each cluster V_j in $\mathcal{P}$ the number of edges of $xy \cup ACS(V(y), V(x)^+)$ leaving V_j equals the number of edges of $xy \cup ACS(V(y), V(x)^+)$ entering V_{j+1}. Thus this is also true for the union W_i' of the $xy \cup ACS(V(y), V(x)^+)$ over all edges $xy \in H_i$. So the claim follows.

In what follows, the order of the edges in W_i' does not matter anymore. So we will view W_i' as a multiset consisting of edges in $E(U) \cup E(H)$.

Recall that each occurrence of an edge in U corresponds to an edge in U' (and that these edges in U' are different for different occurrences of the same edge from R in U). So we might also view each W_i' as a multiset consisting of edges in $E(U') \cup E(H)$.

Claim 2. 'Unions of sequences are globally balanced.' *Let W denote the union of $W_1', \dots, W_q'$. Then there is an integer t so that W contains each edge of U exactly t times. Thus if W is viewed as a multiset consisting of edges in $E(U') \cup E(H)$, then W also contains each edge of U' exactly t times.*

As before, here (and below) multiple occurrences of an edge in U are considered separately, i.e. if the edge AB appears u times in U, then it altogether appears ut times in W.

To prove Claim 2, consider the auxiliary multidigraph D whose vertices are $V_1, \dots, V_k$ and which contains an edge from V_i to V_j for every $ACS(V_i, V_j)$ included into W. So the multiplicity of the edge V_iV_j in D is the number of edges xy of H with $V(y) = V_i$ and $V(x)^+ = V_j$. Let H^c be obtained from H by first reversing the orientation of every edge and then contracting all the vertices lying in each cluster V_i into a new vertex v_i. So H^c is an m-regular multigraph (which might contain loops). Moreover, D can be obtained from H^c by replacing each edge v_iv_j with the edge v_iv_{j+1}. Thus D is m-regular too and so it can be decomposed into edge-disjoint 1-factors. Consider any cycle $D' = V_{i_1} \dots V_{i_r}$ in one of these 1-factors. Then W contains all edges in the multiset

$$S(D') := ACS(V(x_{i_1}), V(x_{i_2})) \cup ACS(V(x_{i_2}), V(x_{i_3})) \cup \dots \cup ACS(V(x_{i_r}), V(x_{i_1})).$$

But

$$ACS(V_{i_j}, V_{i_{j+1}}) = AECS(V_{i_j}, V_{i_j+1}) \cup \dots \cup AECS(V_{i_{j+1}-1}, V_{i_{j+1}}).$$

So it follows that $S(D')$ contains every $AECS(V_i, V_{i+1})$ the same number of times and thus $S(D')$ is a multiple of $E(U)$. (As an example, if $D' = V_9V_4V_8$, then $S(D')$ will contain each edge of $E(U)$ once. If however $D' = V_9V_8V_4$, then $S(D')$ will contain each edge of $E(U)$ twice.) Since W is the union of the $S(D')$ over all cycles D' in the 1-factor decomposition of D, this implies that W is a multiple of $E(U)$, i.e. there exists t such that W contains every edge in U exactly t times.

Since we consider multiple occurrences of an edge in U separately and each such occurrence corresponds to an edge of U' it follows immediately that W (viewed as a multiset consisting of edges in $E(U') \cup E(H)$) also contains each edge of U' exactly t times. This proves Claim 2.

Let

$$s' := \phi m. \tag{9.4}$$

Note that (9.3) implies that

$$s' = \phi \ell' m' \le m'/10^4. \tag{9.5}$$

We will need the following claim when replacing the edges of U by edges in G.

Claim 3. 'Sequences are well spread out.' *For all $i = 1, \dots, q$, any edge of U' (and of U) occurs in W_i' at most s' times.*

To prove Claim 3, note that for each $j = 1, \dots, k$ each edge in H_i contributes at most one occurrence of $AECS(V_j, V_{j+1})$ in W_i'. But H_i'' and H_i are vertex-disjoint by (b) and $H_i' = H_i \cup H_i''$ consists of at most ϕm paths by (c). So H_i consists of at most ϕm edges and thus the total number of occurrences of $AECS(V_j, V_{j+1})$ in W_i' is at most $\phi m = s'$. This proves Claim 3 since $AECS(V_1, V_2), \dots, AECS(V_k, V_1)$ forms a partition of $E(U)$.

By summing over all i with $1 \le i \le q$, it immediately follows that the constant t defined in Claim 2 satisfies

$$t \le s'q. \tag{9.6}$$

We will now add some further edges to W to obtain W' which contains all edges of H and which uses every edge AB of U' (and thus of U) *exactly* t' times, where

$$t' := r_2 m' \overset{(9.2)}{=} 12\phi\ell' q m' \overset{(9.3)}{=} 12\phi q m \overset{(9.4)}{=} 12 s' q. \tag{9.7}$$

Claim 4. *By adding some some further edges of U to each W_i' we can obtain multisets W_i'' which satisfy the following properties (as before, we also view W_i'' as a multiset consisting of edges in $E(U') \cup E(H)$):*

- *Each W_i'' is still locally balanced. That is, for every cluster V_j in $\mathcal{P}$ and each $i = 1, \dots, q$, the number of edges of W_i'' leaving V_j equals the number of edges of W_i'' entering V_{j+1}.*
- *For each $i = 1, \dots, q$ and each edge AB of U', let $s_i(AB)$ be the number of times that W_i'' uses AB. Then*

$$\sum_{i=1}^{q} s_i(AB) = t' \tag{9.8}$$

and

$$11s' \le s_i(AB) \le 13s' \tag{9.9}$$

for each $i = 1, \dots, q$.

To prove Claim 4, first note that

$$t + 11s'q \overset{(9.6)}{\le} 12s'q \overset{(9.7)}{=} t'.$$

Choose integers $s_1', \dots, s_q'$ so that

$$\sum_{i=1}^{q} s_i' = t' - t \qquad \text{and} \qquad 11s' \le s_i' \le 12s' \quad \text{for all } i = 1, \dots, q.$$

(The s'_i exist since $11s'q \le t'-t \le 12s'q$.) For each $i=1,\dots,q$, let W''_i be obtained from W'_i by adding s'_i copies of U. Then clearly (9.8) holds. Also, Claim 3 and the bounds on s'_i imply that

$$11s' \le s'_i \le s_i(AB) \le s'_i + s' \le 13s'$$

for each $i=1,\dots,q$. So (9.9) holds too.

The fact that each W''_i is still locally balanced immediately follows from Claim 1 and the fact that W''_i was obtained from W'_i by adding copies of U. (Note that (U3) implies that each copy of U contributes exactly ℓ' to the number of edges entering a cluster and ℓ' to the number of edges leaving a cluster.) This completes the proof of Claim 4.

Let W' be the union of the multisets W''_i over all $i=1,\dots,q$. Thus Claim 4 implies that in total W' contains each edge AB of U' exactly t' times. Since $\mathcal{B}'(U')$ is an r_2-blow-up of U', it follows that for any edge AB of U', t' equals the number of edges in the subgraph

$$S(AB) := \mathcal{B}'(U')[A,B] \tag{9.10}$$

of $\mathcal{B}'(U')$ spanned by the clusters A and B. (So here A and B are clusters in $\mathcal{P}'$.)

Next we will replace each occurrence of an edge AB of U' in $W''_i \setminus E(H_i)$ by an edge of $S(AB)$ to obtain a digraph W'''_i with $V(W'''_i) = V(H) = V(G)\setminus V_0$ which has the following properties:

- (α_1) W'''_i contains all edges in H_i.
- (α_2) For each edge AB of U', the bipartite subgraph $W'''_i[A,B]$ consisting of all edges in W'''_i from A to B contains exactly $s_i(AB)$ edges of $S(AB)$.
- (α_3) $W'''_i \cup H'_i$ is a path system.
- (α_4) For every pair V, V^+ of consecutive clusters on C and every $i=1,\dots,q$, there is an integer $w_i(V) \le \sqrt{\phi}m/2$ so that

$$w_i(V) = \sum_{v\in V} d^+_{W'''_i}(v) = \sum_{v\in V^+} d^-_{W'''_i}(v). \tag{9.11}$$

- (α_5) $W'''_1,\dots,W'''_q$ are pairwise edge-disjoint.

Note that (α_2) is equivalent to stating that each occurrence of AB in W''_i is replaced by an edge of $S(AB)$. Moreover (α_1), (α_2), (α_5), (9.7) and (9.8) together imply that the edge sets of the W'''_i form a partition $E(H\cup \mathcal{B}'(U'))$.

Before describing the construction of W'''_i, first note that (α_4) is an immediate consequence of (α_2) and Claim 4: any edge AB of W''_i (where $AB\in E(U')$ and so A and B are clusters in $\mathcal{P}'$) corresponds to an edge in W'''_i which goes from A to B. So (9.11) holds. To check that $w_i(V)\le \sqrt{\phi}m/2$ recall from (U3) and (ST2) that for each cluster V on C there are exactly ℓ' edges AB of U' leaving V. (9.9) implies that each such edge AB contributes at most $13s' = 13\phi m$ to $w_i(V)$. Together with (c) this implies that

$$w_i(V) \le |H_i\cap V| + 13\phi\ell' m \le \phi m + 13\phi \ell' m \le \sqrt{\phi}m/2.$$

To construct W'''_i for each i, we proceed as follows. Let u be the length of U' and label the edges of U' as $E_1,\dots,E_u$. Consider any edge $E_a = AB$ of U'. For each $j=1,\dots,4$ let $S^j(AB) := S^j_1(AB)\cup\dots\cup S^j_{r_2}(AB)$. (Recall the $S^j_i(AB)$ were defined in condition (a) of the lemma.) Then $S(AB) = S^1(AB) \cup \cdots \cup S^4(AB)$ (where $S(AB)$ is as defined in (9.10)) and (a) implies that

$$|S^j(AB)| = |S(AB)|/4 = r_2 m'/4. \tag{9.12}$$

Order the edges of $S(AB)$ in such a way that the following conditions hold:

(β_1) Every set of at most $m'/20$ consecutive edges in $S(AB)$ forms a matching.

(β_2) If $AB \neq E_u$ then for each $j = 1, 2, 3$ all the edges in $S^j(AB)$ precede all those in $S^{j+1}(AB)$.

(β_3) If $AB = E_u$ then all the edges in $S^3(AB)$ precede all those in $S^4(AB)$, which in turn precede those in $S^1(AB)$ and all the edges in $S^2(AB)$ are at the end of the ordering.

(β_3) will be used to ensure that (α_3) holds in the construction of the W_i'''.

To see that the above properties can be guaranteed, we use the properties of $S(AB)$ described in the assumption (a) of the lemma: for each edge $AB \neq E_u$ of U', order the edges in $S(AB)$ so that (β_2) is satisfied and so that within some $S^j(AB)$ all the edges of the matching $S_i^j(AB)$ come before all edges of the matching $S_{i+1}^j(AB)$ (for all $i = 1, \dots, r_2 - 1$). Order the edges of $S_1^j(AB)$ arbitrarily. Given an ordering of the edges in $S_i^j(AB)$, order the edges of $S_{i+1}^j(AB)$ in such a way that the first $m'/20$ edges of $S_{i+1}^j(AB)$ avoid the $m'/10$ endvertices of the final $m'/20$ edges of $S_i^j(AB)$. This ensures that (β_1) will be satisfied. If $AB = E_u$ then the argument is similar, but we start with an ordering of the edges in $S(AB)$ so that (β_3) is satisfied.

We now carry out the actual construction of the W_i''', where we consider the W_i''' in batches of 10. For each $a = 1, \dots, q/10$ and each edge E_j of U' we let

$$u_a(E_j) := s_{10(a-1)+1}(E_j) + \cdots + s_{10a}(E_j).$$

Thus (9.9) implies that for all $a = 1, \dots, q/10$,

$$110s' \leq u_a(E_j) \leq 130s'. \tag{9.13}$$

We let $S_a^*(E_j)$ denote the set of all those edges whose position in the ordering of the edges of $S(E_j)$ lies between $1 + \sum_{a'=1}^{a-1} u_{a'}(E_j)$ and $\sum_{a'=1}^{a} u_{a'}(E_j)$. So $u_a(E_j) = |S_a^*(E_j)|$. Together with (β_1) and the fact that $u_a(E_j) \leq 130s' \leq m'/20$ by (9.5), this implies that $S_a^*(E_j)$ forms a matching. Note that $u_a(E_j)$ is the total number of edges in $S(E_j)$ that we need to choose for $W_{10(a-1)+1}''', \dots, W_{10a}'''$. We will choose all these edges from $S_a^*(E_j)$.

To choose these edges, we consider an auxiliary bipartite graph B^* which is defined as follows. The first vertex class B_1 of B^* consists of $u_a(E_j)$ placeholders for the edges in $S(E_j)$ that we need to choose for $W_{10(a-1)+1}''', \dots, W_{10a}'''$, so for each $i = 10(a-1)+1, \dots, 10a$ there will be precisely $s_i(E_j)$ of these placeholders for (the edges to be chosen for) W_i'''. The second vertex class of B^* is $S_a^*(E_j)$. So

$$110s' \leq |S_a^*(E_j)| = |B_1| \leq 130s'$$

by (9.13). We join an edge $e \in S_a^*(E_j)$ to a placeholder for W_i''' if e is vertex-disjoint from H_i'. Since by condition (c) of the lemma $H_{10(a-1)+1}', \dots, H_{10a}'$ are pairwise vertex-disjoint, each edge $e \in S_a^*(E_j)$ can meet at most two H_i' with $10(a-1)+1 \leq i \leq 10a$ and so e will be joined to all placeholders apart from those for the corresponding two W_i'''. Since there are $s_i(E_j) \le 13s'$ placeholders for each W_i''', this means that e is joined in B^* to all but at most $2 \cdot 13s' \le |B_1|/2$ placeholders in B_1. Similarly, since $S_a^*(E_j)$ forms a matching and since by (c) every H_i' meets each of the two endclusters of E_j in at most ϕm vertices, each placeholder in B_1 is joined to all but at most $2\phi m = 2s' \le |S_a^*(E_j)|/2$ edges in $S_a^*(E_j)$. Thus B^* has a perfect matching. For each $i = 10(a-1)+1, \dots, 10a$ we add the $s_i(E_j)$ edges to W_i''' which are matched to the placeholders for W_i'''.

We carry out this procedure for every edge E_j of U' in turn. This completes the construction of the W_i'''. Clearly, (α_1) and (α_2) are satisfied. To check that (α_5) holds, note that the $W_i''' \setminus E(H_i)$ are pairwise edge-disjoint by construction and the H_i are pairwise edge-disjoint by definition (as they form a partition of the edges of H into matchings). Also $W_i''' \setminus E(H_i)$ is edge-disjoint from any H_j by (d).

So let us now check that (α_3) is satisfied. To do this, let $W_i'''[E_j]$ denote the bipartite subdigraph of W_i''' which consists of all edges from the first endcluster of E_j to the final endcluster of E_j. Note that by definition of B^*, the edges of $W_i''' \setminus E(H_i)$ and those of H_i' have no endvertices in common. Moreover, (b) implies that H_i' is a path system. So for (α_3), it suffices to show that $W_i''' \setminus E(H_i)$ is a path system. Now recall that the definition of B^* implies that $W_i'''[E_j] \setminus E(H_i)$ forms a matching for each edge E_j of U'. Thus the only possibility for a cycle C' in $W_i''' \setminus E(H_i)$ would be for C' to 'wind around' U'.

So in order to show that $W_i''' \setminus E(H_i)$ is a path system, it suffices to show that no vertex is incident to both an edge in $W_i'''[E_1] \setminus E(H_i)$ and an edge in $W_i'''[E_u] \setminus E(H_i)$. But this follows from (β_2) and (β_3). Indeed, recall that when choosing the edges in $W_i'''[E_1] \setminus E(H_i)$ we considered all the W_i''' in batches of 10. Let $a := \lceil i/10 \rceil$. So the edges in $W_i'''[E_1] \setminus E(H_i)$ were chosen in the ath batch. Let $p_1^{\rm first}$ and $p_1^{\rm final}$ denote the first and the final position of an edge from $W_i'''[E_1] \setminus E(H_i)$ in the ordering of all edges of $S(E_1)$. Define $p_u^{\rm first}$ and $p_u^{\rm final}$ similarly. Note that

$$110s'(a-1) \stackrel{(9.13)}{\le} p_1^{\rm first}, p_1^{\rm final} \stackrel{(9.13)}{\le} 130s'a. \tag{9.14}$$

But

$$130s'a - 110s'(a-1) = 20s'a + 110s' \le 20s'\frac{q}{10} + 110s' < 3s'q \stackrel{(9.7)}{=} \frac{r_2 m'}{4} \stackrel{(9.12)}{=} |S^j(E_1)|.$$

Together with (9.14) this implies that

$$110s'(a-1) \le p_1^{\rm first}, p_1^{\rm final} < 110s'(a-1) + |S^j(E_1)|. \tag{9.15}$$

Then (9.15), its analogue for $p_u^{\rm first}$ and $p_u^{\rm final}$, (β_2) and (β_3) together imply that there is some $j \le 3$ such that

$$W_i'''[E_1] \setminus E(H_i) \subseteq S^j(E_1) \cup S^{j+1}(E_1) \text{ and } W_i'''[E_u] \setminus E(H_i) \subseteq S^{j+2}(E_u) \cup S^{j+3}(E_u)$$

(where $S^5(E_u) := S^1(E_u)$ and $S^6(E_u) := S^2(E_u)$, see Figure 4). So no vertex is incident to both an edge in $W_i'''[E_1] \setminus E(H_i)$ and an edge in $W_i'''[E_u] \setminus E(H_i)$. Altogether this shows that (α_3) holds. So we have shown that (α_1)–(α_5) hold.

We now add all the edges in $H_i'' = H_i' \setminus E(H_i)$ to W_i''' and let W_i^* denote the graph on $V(H) = V(W_i''')$ obtained in this way. Recall that by (b), H_i'' consists of a complete exceptional path system with respect to C. Thus (CEPS1) and (CEPS3)

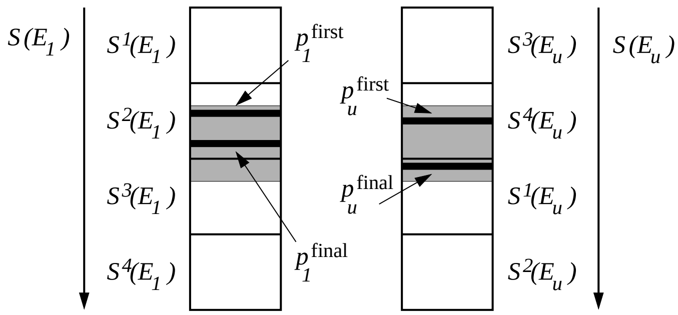

FIGURE 4. The shaded area corresponds to the possible positions of the edges in $W_i'''[E_1] \setminus E(H_i)$ within the set $S(E_1)$ and the edges in $W_i'''[E_u] \setminus E(H_i)$ within the set $S(E_u)$.

together imply that H_i'' is 'locally balanced', in the sense that for every cluster V in $\mathcal{P}$ the number of edges in H_i'' leaving V equals the number of edges in H_i'' entering V^+. (c) implies that the number of edges leaving V is at most ϕm. Together with (α_4) this implies that condition (γ_3) below holds. (γ_1) and (γ_2) follow from (α_1) and (α_3) respectively. (γ_4) follows from (α_5), the fact that the original versions of the H_i' (and thus of the H_i'') are pairwise edge-disjoint by (c) and the fact that $W_i^* \setminus E((H_i'')^{\rm orig})$ is edge-disjoint from any $(H_j'')^{\rm orig}$ by (d).

(γ_1) W_i^* contains all edges in H_i'.
(γ_2) W_i^* is a path system.
(γ_3) For every pair V, V^+ of consecutive clusters on C and every i, there is an integer $w_i(V) \le \sqrt{\phi} m$ so that

$$w_i(V) = \sum_{v \in V} d^+_{W_i^*}(v) = \sum_{v \in V^+} d^-_{W_i^*}(v).$$

(γ_4) The original versions of $W_1^*, \dots, W_q^*$ are pairwise edge-disjoint.

Note that for each $i = 1, \dots, q$ every exceptional edge in W_i^* lies in H_i''. Our next aim is to turn the original versions of $W_1^*, \dots, W_q^*$ into edge-disjoint Hamilton cycles by adding suitable edges from $\mathcal{B}(C)^*$.

Claim 5. *For all $i = 1, \dots, q$, there is a Hamilton cycle C_i in G which contains all edges in the original version $(W_i^*)^{\rm orig}$ of W_i^* and such that all the edges in $E(C_i) \setminus (W_i^*)^{\rm orig}$ lie in $\mathcal{B}(C)^*$. Moreover, all these Hamilton cycles C_i are pairwise edge-disjoint.*

Choose a new constant ε'' such that

$$\phi, \varepsilon' \ll \varepsilon'' \ll r_1/m.$$

Suppose that we have already transformed the original versions of $W_1^*, \dots, W_{i-1}^*$ into Hamilton cycles $C_1, \dots, C_{i-1}$. Let $\mathcal{B}(C)_i^*$ denote the subdigraph of $\mathcal{B}(C)^*$ obtained by removing all edges in $C_1, \dots, C_{i-1}$. For every cluster $V_j \in \mathcal{P}$ let V_j^1 be the set of all those vertices $v \in V_j$ for which $d^+_{W_j^*}(v) = 0$ and let V_j^2 be the set of all those vertices $v \in V_j$ for which $d^-_{W^*_j}(v) = 0$. Thus (γ_3) implies that

$$|V_j^1| \geq m - w_i(V_j) \geq (1 - \sqrt{\phi})m \geq (1 - (\varepsilon''/2)^2)m,$$

and similarly $|V_j^2| \geq (1 - (\varepsilon''/2)^2)m$. Proposition 4.3(ii) applied with $d' = (\varepsilon''/2)^2$ now implies that $\mathcal{B}(C)^*_i[V_j^1, V_{j+1}^2]$ is still $(\varepsilon'', r_1/m)$-superregular. (To see that Proposition 4.3(ii) can be applied we use that the removal of each C_j decreases the minimum out- and indegree of every vertex of $\mathcal{B}(C)^*[V_j, V_{j+1}]$ by at most 1. Thus $\mathcal{B}(C)^*_i[V_j^1, V_{j+1}^2]$ is obtained from $\mathcal{B}(C)^*[V_j, V_{j+1}]$ by deleting at most $q \leq (\varepsilon''/2)^2 m$ edges at every vertex and by removing at most $(\varepsilon''/2)^2 m$ vertices from each vertex class.)

On the other hand, (γ_2) and (γ_3) imply that $|V_j^1| = |V_{j+1}^2|$. So we can apply Proposition 4.14 to find a perfect matching M_j in $\mathcal{B}(C)^*_i[V_j^1, V_{j+1}^2]$. Then the union F_i of the M_j (for all $j = 1, \dots, k$) and of W_i^* is a 1-regular digraph on $V(G) \setminus V_0$. We can now apply Lemma 6.5 with F_i, $\mathcal{B}(C)^*_i$, $E(C)$, ε'', r_1/m playing the roles of F, G, J, ε, d to replace each M_j with a suitable other perfect matching in $\mathcal{B}(C)^*_i[V_j^1, V_{j+1}^2]$ to make F_i into a Hamilton cycle C_i' on $V(G) \setminus V_0$. To see that (ii) of Lemma 6.5 is satisfied, consider any cycle D in F_i. If D does not contain any edges from W_i^*, then it meets V_j^1 for every $j = 1, \dots, k$. So suppose that D contains some edges from W_i^* and let v be a final vertex on a subpath in $W_i^* \cap D$ (such a vertex exists by (γ_2)). Then $v \in V_j^1$, where V_j is the cluster containing v.

Let C_i be the original version $(C_i')^{\rm orig}$ of C_i'. Then Observation 7.4 implies that C_i is a Hamilton cycle of G. By (γ_4) all the Hamilton cycles $C_1, \dots, C_q$ will be pairwise edge-disjoint. This completes the proof of Claim 5.

Let us now consider the case when $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is a $(\ell', k, m, \varepsilon, d)$-bi-setup. The argument for this case is similar, so we only highlight the places where it differs. Given clusters V_j and $V_{j'}$ such that $|j' - j|$ is even, we now define

$$ACS(V_j, V_{j'}) := AECS^{bi}(V_j, V_{j+2}) \cup AECS^{bi}(V_{j+2}, V_{j+4}) \cup \dots \cup AECS^{bi}(V_{j'-2}, V_{j'}).$$

Since H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$, every edge xy of H either satisfies $V(x)^+, V(y) \in \mathcal{V}_{\rm even}$ or $V(x)^+, V(y) \in \mathcal{V}_{\rm odd}$. Thus we can define the W_i' as before and Claim 1 still holds.

Recall that $U_{\rm even}$ and $U_{\rm odd}$ form a partition of the edges of U. Since each occurrence of an edge in U corresponds to an edge in U', this also defines sets $U'_{\rm even}$ and $U'_{\rm odd}$ corresponding to $U_{\rm even}$ and $U_{\rm odd}$. Moreover, $U'_{\rm even}$ and $U'_{\rm odd}$ form a partition of the edges of U'. Instead of Claim 2 we now have the following claim.

Claim 2$'$. *Let W denote the union of $W_1', \dots, W_q'$. Then there are integers $t_{\rm even}$ and $t_{\rm odd}$ so that W contains each edge of $U_{\rm even}$ exactly $t_{\rm even}$ times and every edge of $U_{\rm odd}$ exactly $t_{\rm odd}$ times. Thus if W is viewed as a multiset consisting of edges in $E(U') \cup E(H)$, then W also contains each edge of $U'_{\rm even}$ exactly $t_{\rm even}$ times and every edge of $U'_{\rm odd}$ exactly $t_{\rm odd}$ times.*

To prove Claim 2$'$, define an auxiliary digraph D as before. Note that this time, if $V_iV_j \in E(D)$ then either both i and j are even or both i and j are odd. As before, D is regular and thus there exists a decomposition of D into edge-disjoint 1-factors. Consider any cycle $D' = V_{i_1} \dots V_{i_r}$ in one of these 1-factors. Then either all of $i_1,\dots,i_r$ are even or all of them odd. Suppose first that the former holds. Then W contains all edges in the multiset

$$S(D') := ACS(V(x_{i_1}),V(x_{i_2})) \cup ACS(V(x_{i_2}),V(x_{i_3})) \cup \dots \cup ACS(V(x_{i_r}),V(x_{i_1})).$$

But

$$ACS(V_{i_j},V_{i_{j+1}}) = AECS(V_{i_j},V_{i_j+2}) \cup \dots \cup AECS(V_{i_{j+1}-2},V_{i_{j+1}}).$$

So it follows that $S(D')$ contains every $AECS(V_i,V_{i+2})$ for which i is even the same number of times and thus $S(D')$ is a multiple of $E(U_{\rm even})$.

If all of $i_1,\dots,i_r$ are odd then it follows that $S(D')$ is a multiple of $E(U_{\rm odd})$. Since W is the union of the $S(D')$ over all cycles D' in the 1-factor decomposition of D, this implies Claim $2'$.

As before, one can show that Claim 3 holds and so instead of (9.6) we now have $t_{\rm even}, t_{\rm odd} \le s'q$ and so

$$t_{\rm even} + 11s'q,\ t_{\rm odd} + 11s'q \le 12s'q \overset{(9.7)}{=} t'.$$

Choose integers $s_1^{\rm even},\dots,s_q^{\rm even}$ so that

$$\sum_{i=1}^{q} s_i^{\rm even} = t' - t_{\rm even} \qquad \text{and} \qquad 11s' \le s_i^{\rm even} \le 12s' \quad \text{for all } i=1,\dots,q.$$

Define $s_1^{\rm odd},\dots,s_q^{\rm odd}$ similarly. For each $i=1,\dots,q$, let W_i'' be obtained from W_i' by adding $s_i^{\rm even}$ copies of $U_{\rm even}$ and $s_i^{\rm odd}$ copies of $U_{\rm odd}$. Then the W_i'' are as desired in Claim 4. The remainder of the proof is now identical. □

Suppose we are given a 1-factor H of $G - V_0$ which is split into suitable matchings H_i. We will apply the following lemma in the proof of Lemma 9.7 to assign a complete exceptional path system $CEPS_i$ to each H_i so that $CEPS_i$ can play the role of H_i'' in Lemma 9.5. We then use Lemma 9.5 to extend H_i into a Hamilton cycle.

Lemma 9.6. *Suppose that $0 < 1/n \ll 1/k, 1/q, \varepsilon, 1/f \ll 1$, that $t, k/f, fm/q \in \mathbb{N}$ and that $(G,\mathcal{P},R,C)$ is a (k,m,ε,d)-scheme on n vertices. Suppose that $\mathcal{P}^*$ is a (q/f)-refinement of $\mathcal{P}$ and that $EF_1,\dots,EF_t$ are exceptional factors with parameters $(q/f,f)$ with respect to $C,\mathcal{P}^*$. Let $\mathcal{I}$ denote the canonical interval partition of C into f intervals of equal length. Suppose that $H_1,\dots,H_{tq}$ are subdigraphs of G satisfying the following properties:*

- (a) *For each $i=1,\dots,tq$ there are at most $f/100$ intervals $I \in \mathcal{I}$ such that H_i contains a vertex lying in a cluster on I.*
- (b) *For each interval $I \in \mathcal{I}$ there are at most $tq/100$ indices i with $1 \le i \le tq$ and such that H_i contains a vertex lying in a cluster on I.*
- (c) *If $|i-j| \le 10$, then H_i and H_j are vertex-disjoint, with the indices considered modulo tq.*

Then the tq complete exceptional path systems contained in $EF_1,\dots,EF_t$ can be labelled $CEPS_1,\dots,CEPS_{tq}$ such that the following conditions hold:

- *$H_i \cup CEPS_i$ and $H_j \cup CEPS_j$ are pairwise vertex-disjoint whenever $|i-j| \leq 10$.*
- *H_i and $CEPS_i$ are vertex-disjoint.*

Proof. Let $\mathcal{CEPS}$ denote the set of the tq complete exceptional path systems contained in $EF_1, \dots, EF_t$. In order to label them, we consider an auxiliary bipartite graph B defined as follows. The first vertex class B_1 of B consists of all the H_i. The second vertex class B_2 is $\mathcal{CEPS}$. So $|B_1| = |B_2| = tq$. We join $H_i \in B_1$ to $CEPS \in B_2$ by an edge in B if $CEPS$ is vertex-disjoint from each of $H_{i-10}, H_{i-9}, \dots, H_{i+10}$. Our aim is to find a perfect matching in B which has the following additional property:

For all $1 \leq i < j \leq tq$ with $|i-j| \leq 10$ the two complete exceptional path systems which are matched to H_i and H_j are vertex-disjoint from each other. ($\heartsuit$)

To show that such a perfect matching exists, let M be a matching of maximum size satisfying ($\heartsuit$) and suppose that M is not perfect. Pick $H_i \in B_1$ and $CEPS^* \in B_2$ such that they are not covered by M. We say that an interval $I \in \mathcal{I}$ is *bad for H_j* if H_j contains a vertex lying in a cluster on I. Let $\mathcal{I}'$ denote the set of all those intervals $I \in \mathcal{I}$ which are bad for at least one of $H_{i-10}, H_{i-9}, \dots, H_{i+10}$. Thus $|\mathcal{I}'| \leq 21f/100$ by (a). But every complete exceptional path system which spans an interval $I \in \mathcal{I} \setminus \mathcal{I}'$ is vertex-disjoint from each of $H_{i-10}, H_{i-9}, \dots, H_{i+10}$. Since for each such I the set $\mathcal{CEPS}$ contains precisely qt/f complete exceptional path systems spanning I, it follows that the degree of H_i in B is at least $(1 - 21/100)tq$. On the other hand, for each $CEPS \in \mathcal{CEPS}$ there are precisely $3t-1$ other complete exceptional path systems in $\mathcal{CEPS}$ which are not vertex-disjoint from $CEPS$. This implies that at most $20 \cdot 3t$ neighbours $CEPS$ of H_i in B are not vertex-disjoint from each of the at most 20 complete exceptional path systems matched to $H_{i-10}, \dots, H_{i-1}, H_{i+1}, \dots, H_{i+10}$ in M. Call a neighbour $CEPS$ of H_i in B *nice* if $CEPS$ is vertex-disjoint from each of these at most 20 complete exceptional path systems. So H_i has at least

$$(1 - 21/100)tq - 60t \geq (1 - 22/100)tq = (1 - 22/100)|B_2| \tag{9.16}$$

nice neighbours in B. Note that each nice neighbour $CEPS$ of H_i has to be covered by M (otherwise we could enlarge M into a bigger matching satisfying ($\heartsuit$) by adding the edge between H_i and $CEPS$). Thus in particular,

$$|M| \geq (1 - 22/100)|B_2|. \tag{9.17}$$

Let $I^* \in \mathcal{I}$ be the interval which $CEPS^*$ spans. Then (b) implies that I^* is bad for at most $tq/100$ of the H_j. But this implies that there are at most $21tq/100$ indices j with $1 \leq j \leq tq$ and such that I^* is bad for at least one of $H_{j-10}, H_{j-9}, \dots, H_{j+10}$. Thus the degree of $CEPS^*$ in B is at least $(1 - 21/100)tq$. Together with (9.17) this implies that $CEPS^*$ has at least $(1 - 43/100)|B_1|$ neighbours in B which are covered by M. We call such a neighbour H_j *useful* if $CEPS^*$ is vertex-disjoint from each of the (at most) 21 complete exceptional path systems matched to $H_{j-10}, H_{j-9}, \dots, H_{j+10}$ in M. Recall $\mathcal{CEPS}$ contains precisely $3t-1$ other complete exceptional path systems which are not vertex-disjoint from $CEPS^*$. But each of these can force at most 21 neighbours H_j of $CEPS^*$ to become useless (by being matched to one of $H_{j-10}, H_{j-9}, \dots, H_{j+10}$). So $CEPS^*$ has at least

$$(1-43/100)|B_1| - 63t = (1-43/100)|B_1| - 63|B_1|/q \geq (1-44/100)|B_1|$$

useful neighbours which are covered by M. Together with (9.16) this implies that there is a matching edge $e \in M$ such that its endpoint in B_1 is a useful neighbour of $CEPS^*$ while its endpoint in B_2 is a nice neighbour of H_i. Let H_j and $CEPS$ be the endpoints of e. Let M' be the matching obtained from M by deleting e and adding the edge between H_i and $CEPS$ and the edge between H_j and $CEPS^*$. Then M' is a larger matching which still satisfies ($\heartsuit$), a contradiction.

This shows that B has a perfect matching satisfying ($\heartsuit$). For each $i = 1, \dots, tq$ we take $CEPS_i$ to be the complete exceptional path system which is matched to H_i. Then (c), the definition of our auxiliary graph B and ($\heartsuit$) together imply that the $CEPS_i$ are as desired. □

To obtain an algorithmic version of the above proof, we simply start with an empty matching in the auxiliary graph B and use the above argument to gradually extend the matching into a perfect one.

For the final lemma of this section, we are given an r-factor H of $G - V_0$. H is then split into 1-factors F_i and these 1-factors are split further into small matchings H^i_j. We use Lemma 9.6 to assign a suitable complete exceptional path system $CEPS^i_j$ to each H^i_j. We then apply Lemma 9.5 to extend each $CEPS^i_j \cup H^i_j$ into a Hamilton cycle using edges of CA. Since Lemma 9.5 allows us to prescribe a regular subgraph $\mathcal{B}'(U')$ of $\mathcal{B}(U')$ whose edges will all be used for the Hamilton cycles, this means we can use up all edges of $\mathcal{B}(U')$ in the process, so the leftover of the entire process is a blow-up of C, as required.

Lemma 9.7. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll 1/q \ll 1/f \ll r_1/m \ll d \ll 1/\ell', 1/g \ll 1$ and that $rk \leq m/f^2$. Let*

$$s := rfk, \quad r_2 := 96\ell' g^2 kr, \quad r_3 := s/q$$

and suppose that $k/f, k/g, q/f, m/4\ell', fm/q, 2fk/3g(g-1) \in \mathbb{N}$. Suppose that

$$(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$$

is an $(\ell', k, m, \varepsilon, d)$-setup with $|G| = n$ and $C = V_1 \dots V_k$. Suppose that H is an r-factor of $G - V_0$ and that $CA = \mathcal{B}(C)^* \cup \mathcal{B}(U') \cup CA^{\mathrm{exc}}$ is a chord absorber for C, U' with parameters $(\varepsilon, r_1, r_2, r_3, q, f)$ whose original version is edge-disjoint from H. Then there are edge-disjoint Hamilton cycles $C_1, \dots, C_s$ in G which satisfy the following conditions:

(i) *Altogether $C_1, \dots, C_s$ contain all the edges of $H \cup \mathcal{B}(U') \cup (CA^{\mathrm{exc}})^{\mathrm{orig}}$. Moreover, all remaining edges in $C_1, \dots, C_s$ are contained in $\mathcal{B}(C)^*$.*
(ii) *$CA^{\mathrm{orig}} \setminus \bigcup_i E(C_i) = \mathcal{B}(C)^* \setminus \bigcup_i E(C_i)$ is an $(r_1 + r_2 + r - (q-1)s/q)$-blow-up of C.*
(iii) *Each C^{basic}_i contains one of the s complete exceptional path systems contained in CA^{exc}.*

The analogue holds for an $(\ell', k, m, \varepsilon, d)$-bi-setup $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ if we assume in addition that H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$ (where $\mathcal{V}_{\rm even}$ is the set of all those V_i such that i is even and $\mathcal{V}_{\rm odd}$ is defined analogously).

Proof. Define new constants by

$$q' := fk \quad \text{and} \quad \phi := 8g^2k/q' = 8g^2/f.$$

Thus

$$s = rq' \quad \text{and} \quad r_2 = 12\ell'\phi q' r.$$

Recall that $V_1, \dots, V_k$ denote the clusters in $\mathcal{P}$. Apply Proposition 6.1 to decompose the edges of H into r edge-disjoint 1-factors $F_1, \dots, F_r$ of $G - V_0$. We now consider each F_{i^*} with $i^* = 1, \dots, r$. Apply Lemma 8.7 with F_{i^*} and q' playing the roles of H and q^* to decompose F_{i^*} into q' matchings $H_1^{i^*}, \dots, H_{q'}^{i^*}$ which satisfy the following properties:

- (a_1) For all $i = 1, \dots, q'$, $H_i^{i^*}$ consists of at most $2g^2km/q' = \phi m/4$ edges. Moreover $|H_i^{i^*} \cap V_j| \le 4g^2km/q' \le \phi m/2$ for all $j = 1, \dots, k$.
- (a_2) If $|i - j| \le 10$, then $H_i^{i^*}$ and $H_j^{i^*}$ are vertex-disjoint, with the indices considered modulo q'.
- (a_3) Each $H_i^{i^*}$ consists entirely of edges of the same double-type and for each $t \in \binom{g}{2}$ the number of $H_i^{i^*}$ of double-type t is $q'/\binom{g}{2}$.

Note that the 'moreover part' of (a_1) follows immediately from the first part of (a_1). For (a_3), recall that we considered a canonical interval partition $\mathcal{I}_g$ of C into g edge-disjoint intervals of equal length and for each $j = 1, \dots, g$ we denote the union of the clusters in the jth interval by X_j. Then $H_i^{i^*}$ has double-type ab (where $a, b \le g$) if all its vertices are contained in $X_a \cup X_b$.

For each $i^* = 1, \dots, r$ and each $i = 1, \dots, q'$ we now assign a suitable complete exceptional path system $CEPS_i^{i^*}$ from $CA^{\rm exc}$ to $H_i^{i^*}$. We do this in such a way that all these complete exceptional path systems are distinct from each other and the following properties hold:

- (b_1) For all $i = 1, \dots, q'$ and all $j = 1, \dots, k$ we have $|(H_i^{i^*} \cup CEPS_i^{i^*}) \cap V_j| \le \phi m$. Moreover, each $H_i^{i^*} \cup CEPS_i^{i^*}$ consists of at most ϕm paths.
- (b_2) $H_i^{i^*} \cup CEPS_i^{i^*}$ and $H_j^{i^*} \cup CEPS_j^{i^*}$ are pairwise vertex-disjoint whenever $|i - j| \le 10$ (for each $i^* = 1, \dots, r$).
- (b_3) $H_i^{i^*}$ and $CEPS_i^{i^*}$ are vertex-disjoint.
- (b_4) $H_i^{i^*} \cup (CEPS_i^{i^*})^{\rm orig}$ and $H_j^{i^*} \cup (CEPS_j^{i^*})^{\rm orig}$ are pairwise edge-disjoint whenever $i \ne j$ (for each $i^* = 1, \dots, r$).

Note that since $CA^{\rm exc}$ consists of exceptional factors with parameters $(q/f, f)$, the $CEPS_i^{i^*}$ will always satisfy $|CEPS_i^{i^*} \cap V_j| \le fm/q \le m/f \le \phi m/2$ and each $CEPS_i^{i^*}$ will always consist of $fm/q \le \phi m/2$ paths. So (b_1) will follow from (a_1) and (b_3). (b_4) follows immediately from the fact that H and $(CA^{\rm exc})^{\rm orig}$ are edge-disjoint.

In order to choose the $CEPS_i^{i^*}$ we proceed as follows. Let $t := q'/q = r_3/r$. Let $\mathcal{I}$ be the canonical interval partition of C into f intervals of equal length. So $CA^{\rm exc}$ consists of r_3 exceptional factors $EF_1, \dots, EF_{r_3}$, where each EF_j induces the disjoint union of q/f complete exceptional path systems on each interval $I \in \mathcal{I}$. For each

$i^* = 1, \dots, r$ let $\mathcal{CEPS}^{i^*}$ denote the set of all complete exceptional path systems contained in $EF_{(i^*-1)t+1}, \dots, EF_{i^*t}$. Each of these exceptional factors contains q complete exceptional path systems, so altogether we have $tq = q'$ of them in $\mathcal{CEPS}^{i^*}$. We will take $CEPS_1^{i^*}, \dots, CEPS_{q'}^{i^*}$ to be the complete exceptional path systems in $\mathcal{CEPS}^{i^*}$.

To choose a suitable labeling of the $CEPS_i^{i^*}$, we aim to apply Lemma 9.6 with $H_1^{i^*}, \dots, H_{q'}^{i^*}$ playing the roles of $H_1, \dots, H_{tq}$ and $EF_{(i^*-1)t+1}, \dots, EF_{i^*t}$ playing the roles of $EF_1, \dots, EF_t$. So we need to check that conditions (a)–(c) of Lemma 9.6 hold. Condition (c) follows from (a_2). To check (a), consider any $H_i^{i^*}$ and let ab denote its double-type. Note that $\mathcal{I}$ can be obtained from $\mathcal{I}_g$ by splitting each interval in $\mathcal{I}_g$ into f/g intervals of equal length. Thus at most $4 + 2f/g \le f/100$ intervals $I \in \mathcal{I}$ have the property that $X_a \cup X_b \supseteq V(H_i^{i^*})$ contains a vertex lying in a cluster on I (the extra 4 accounts for those intervals sharing exactly one cluster with X_a or X_b). To check (b), consider any interval $I \in \mathcal{I}$. Then there are at most $2g$ double-types ab such that the set $X_a \cup X_b$ does not avoid all the clusters on I. Since by (a_3) for each double-type precisely $q'/\binom{g}{2}$ of $H_1^{i^*}, \dots, H_{q'}^{i^*}$ have that double-type, this implies at most $2gq'/\binom{g}{2} \le q'/100$ of $H_1^{i^*}, \dots, H_{q'}^{i^*}$ contain a vertex lying in a cluster on I. Thus we can indeed apply Lemma 9.6 to find a labeling $CEPS_1^{i^*}, \dots, CEPS_{q'}^{i^*}$ of the complete exceptional path systems in $\mathcal{CEPS}^{i^*}$ as described there. Then the $CEPS_i^{i^*}$ also satisfy (b_2) and (b_3).

Our aim now is to apply Lemma 9.5. Let $r^* := 12\ell'\phi q'$. Note that $r_2 = r^* r$. Recall that (CA2) implies that $\mathcal{B}(U')$ is an r_2-blow-up of U' so that for each edge AB of U', there is a partition of both A and B into four subclusters $A_1, \dots, A_4$ and $B_1, \dots, B_4$ of equal size so that $\mathcal{B}(U')[A_j, B_j]$ consists of exactly r_2 edge-disjoint perfect matchings between each pair A_j, B_j (for all $j = 1, \dots, 4$). So we can decompose the edges of $\mathcal{B}(U')$ into edge-disjoint graphs $S_1, \dots, S_r$ so that each of these contains exactly r^* of these perfect matchings for each pair A_j, B_j of subclusters of each edge AB. (So S_{i^*} is an r^*-blow-up of U' for each $i^* = 1, \dots, r$.) Thus we can satisfy condition (a) of Lemma 9.5 if we let S_{i^*} and r^* play the roles of $\mathcal{B}'(U')$ and r_2.

In particular, we can now apply Lemma 9.5 with S_1 playing the role of $\mathcal{B}'(U')$, F_1 playing the role of H, and ϕ, q', $2/f^{1/2}$, r^* playing the roles of ϕ, q, ε', r_2 to obtain a collection $\mathcal{C}_1$ of q' edge-disjoint Hamilton cycles in $G^{\rm orig}$. We next apply Lemma 9.5 for each of $F_2, \dots, F_r$ in turn to find collections $\mathcal{C}_2, \dots, \mathcal{C}_r$, each consisting of q' edge-disjoint Hamilton cycles in $G^{\rm orig}$. For each F_{i^*} we use only S_{i^*}, the unused part of $\mathcal{B}(C)^*$ and the complete exceptional path systems $CEPS_i^{i^*}$ guaranteed by (b_1)–(b_4). Note that in each of the applications of Lemma 9.5 the in- and outdegrees of a vertex in $\mathcal{B}(C)^*$ decrease by at most q'. So in total the in- and outdegrees will decrease by at most $rq' = rfk \le m/f$. Thus Proposition 4.3(ii) applied with $d' := 1/f$ implies that in each step the remainder of $\mathcal{B}(C)^*$ will still be $(2/f^{1/2}, r_1/m)$-superregular. So this means we can indeed apply Lemma 9.5.

We take $C_1, \dots, C_s$ to be the Hamilton cycles in $\mathcal{C}_1 \cup \dots \cup \mathcal{C}_r$. Then clearly $C_1, \dots, C_s$ are pairwise edge-disjoint and they satisfy (i) and (iii). To check (ii), consider any vertex $x \in V(G) \setminus V_0$. Then x has outdegree s in $C_1 \cup \dots \cup C_s$ and all the $r_2 + r_3 + r$ outedges at x in $\mathcal{B}(U') \cup (CA^{\rm exc})^{\rm orig} \cup H$ are covered by $C_1 \cup \dots \cup C_s$.

Moreover, the outdegree in $\mathcal{B}(C)^*$ is r_1. Thus the outdegree of x in $\mathcal{B}(C)^* \setminus \bigcup_i E(C_i)$ is

$$r_1 + r_2 + r_3 + r - s = r_1 + r_2 + r - \frac{(q-1)s}{q}.$$

Since the analogue also holds for the indegree of x, this proves (ii). The proof of the bipartite analogue goes through unchanged. □

10. Absorbing a blown-up cycle via switches

Our main aim in this section is to define (and find) a 'cycle absorber' CyA which will be removed from the original digraph G at the start of the proof of Theorem 1.2. We would like to find a Hamilton decomposition of the union of several cycle absorbers CyA and the 'leftover' G' of the chord absorber obtained by an application of Lemma 9.7. Recall that this leftover G' is a blow-up of C. Consider a 1-factor H in a 1-factorization of the leftover G' – so the edges of H wind around C. CyA will also be a blow-up of C (if one ignores the edges in the complete exceptional path systems which will be contained in CyA). We will first find a special 1-factorization of $H \cup CyA$ which makes use of this property. In particular, either half or all the edges of each 1-factor will come from CyA. We will then successively switch pairs of edges between pairs of these 1-factors of $H \cup CyA$ with the goal of turning each of them into a Hamilton cycle after a certain number of these switches (see Figure 2). These switches will always involve edges from CyA and not from H.

However, it will turn out that if these switches only involve pairs of 1-factors, then the parity of the total number of cycles in a 1-factorization is preserved. In particular, this will imply that we cannot find a Hamilton decomposition of $H \cup CyA$ if we start of with a 1-factorization into an odd number of cycles. So in Section 10.4, we also define a 'parity switcher' which involves switches between triples of 1-factors to overcome this problem. We then extend the cycle absorber CyA into a 'parity extended cycle absorber' PCA and find a Hamilton decomposition of $H \cup PCA$. We proceed in the same way for each 1-factor H in the above 1-factorization of G'.

10.1. **Definition of the cycle absorber.** Let C_4 denote the orientation of a 4-cycle in which two vertices have outdegree 2 (and thus the two other vertices have indegree 2). Given digraphs H and H', we say they form a *switchable pair* if there are vertices x, x^+, y, y^+ so that xx^+, yy^+ are edges of H and xy^+, yx^+ are edges of H'. So the union C_4^* of these four edges forms a copy of C_4. We say that C_4^* is a *HH'-switch*. More generally, we also say that a copy of C_4 in a digraph G (again with the above orientation) is a *potential switch*. A *C_4^*-exchange* consists of moving the edges xx^+, yy^+ from H to H' and moving the edges xy^+, yx^+ from H' to H (see Figure 2). The following proposition (whose proof follows immediately from the definition of a C_4^*-exchange) states the crucial property of switches.

Proposition 10.1. *Given 1-regular digraphs H and H', suppose there is a HH'-switch C_4^* and let $H_{\rm new}$ and $H'_{\rm new}$ be obtained from H and H' via a C_4^*-exchange.*

(i) *If the two edges of $C_4^* \cap H$ lie on the same cycle of H, then $H_{\rm new}$ has one more cycle than H.*

(ii) *If the two edges of $C_4^* \cap H$ lie on different cycles D_1 and D_2 of H, then $H_{\rm new}$ has one less cycle than H. More precisely, the set of cycles of $H_{\rm new}$ is the same as that of H except that the vertices of D_1 and D_2 now lie on a common cycle.*

Moreover, the analogous assertions hold for H'.

Consider a (k, m, ε, d)-scheme $(G, \mathcal{P}, R, C)$. As usual, let $C = V_1, \dots, V_k$ and recall that V_0 denotes the exceptional set in $\mathcal{P}$. Throughout this section, when referring to 'clusters', we will mean the clusters in $\mathcal{P}$, i.e. $V_1, \dots, V_k$. We assume that k is a multiple of 14.

Given a subdigraph F of $G - V_0$, we say the *top half of F* is the subdigraph $F_{\rm top}$ of F induced by all the vertices in $V_1 \cup V_2 \cup \dots \cup V_{k/2+1}$. The *lower half of F* is the subdigraph $F_{\rm low}$ of F induced by all the vertices in $V_{k/2+1} \cup \dots \cup V_{k-1} \cup V_k \cup V_1$.

Roughly speaking, a cycle absorber consists of three edge-disjoint 1-factors F, S and S' whose edges wind around C (we ignore exceptional vertices and edges in this explanatory paragraph). There will be switches $C_{4,j}$ between F and S and $C'_{4,j}$ between F and S'. Suppose we are given a 1-factor H which also winds around C. In the proof of Lemma 10.2, we will construct two 1-factors $T := H_{\rm low} \cup F_{\rm top}$ and $T' := H_{\rm top} \cup F_{\rm low}$. The switches $C_{4,j}$ between F and S will then correspond to switches between T and S. We will use these to turn T into a Hamilton cycle in Lemma 10.2. Moreover, after these switches, the resulting 1-factor obtained from S will be either a Hamilton cycle or will consist of two cycles. We will proceed similarly for T' and S'. If necessary, S and S' will then be transformed into Hamilton cycles using the parity switcher in Section 10.4.

A *bicycle* B on V is a digraph with $V(B) = V$ which consists of exactly two vertex-disjoint (directed) cycles. A *spanning bicycle* B in a digraph G is a 1-factor of G which consists of exactly two vertex-disjoint cycles.

Let $I_1, \dots, I_7$ be a canonical interval partition of C into 7 intervals of equal length. Recall that a complete exceptional path system $CEPS$ completely spans I_i if $CEPS$ spans I_i and the vertex set of $CEPS$ is the union of all the clusters in I_i. Suppose that H is a digraph on $V(G) \setminus V_0$ which contains s complete exceptional path systems (for some s) and whose other edges lie in G. We say that H *agrees with* C if for every edge vv' of H which does not lie in one of the s complete exceptional path systems there is an i with $1 \leq i \leq k$ so that $v \in V_i$ and $v' \in V_{i+1}$. So if $s = 0$ then H agrees with C if and only if H winds around C.

Below we assume that the vertices in V_1 and in $V_{k/2+1}$ are ordered. A *cycle absorber* CyA (with respect to C) in G is a digraph on $V(G) \setminus V_0$ with the following properties:

(CyA0) CyA is the union of three 1-regular digraphs F, S and S', each with vertex set $V(G) \setminus V_0$. $F_{\rm top}$, $F_{\rm low}$, S and S' each contain a complete exceptional path system (labelled $CEPS_3$, $CEPS_5$, $CEPS_4$ and $CEPS_2$ respectively) and $CEPS_i$ completely spans the interval I_i. Moreover, all the edges of $F \cup S \cup S'$ which are not contained in $CEPS_2 \cup \dots \cup CEPS_5$ lie in $G - V_0$ and each of F, S and S' agrees with C. Finally, $F^{\rm orig}$, $S^{\rm orig}$ and $(S')^{\rm orig}$ are pairwise edge-disjoint subdigraphs of G.

(CyA1) For each $j = 1, \dots, m$, let P_j denote the path of length $k/7$ in $F_{\rm top}$ starting at the jth vertex of V_1 and ending in $V_{k/7+1}$. (So each P_j contains precisely one vertex from each cluster in I_1.) Then $P_j \cup P_{j+1}$ forms a switchable pair with S for all $j = 1, \dots, m$ (with indices considered modulo m). Denote the switch by $C_{4,j}$.

(CyA2) There is a potential switch $C_{4,m+1}$ in $G - V_0$ so that S contains two independent edges of $C_{4,m+1}$ but the other two edges of $C_{4,m+1}$ do not lie in $CyA^{\rm orig}$. Moreover, $V(C_{4,j}) \subseteq \bigcup_{V \in I_1} V$ for each $j = 1, \dots, m+1$ and all the $C_{4,j}$ are pairwise vertex-disjoint.

(CyA3) For each $j = 1, \dots, m+1$, denote the edges of $C_{4,j}$ which are contained in S by ℓ_j and r_j. Let L be the (ordered) sequence of edges $\ell_1, \dots, \ell_{m+1}$ and R be the (ordered) sequence of edges $r_1, \dots, r_{m+1}$. S is a bicycle on $V(G) \setminus V_0$ where one cycle contains all edges of L in the given order and the other cycle contains all edges of R in the given order.

Moreover, $F_{\rm low}$ and S' will satisfy the following conditions which are analogous to (CyA1)–(CyA3). In (CyA2′), we define $I_{4,\rm low}$ to be the subinterval $V_{k/2+1} \dots V_{4k/7+1}$ of I_4.

(CyA1′) For each $j = 1, \dots, m$, let P'_j denote the path of length $k/14$ in $F_{\rm low}$ starting at the jth vertex of $V_{k/2+1}$ and ending in $V_{4k/7+1}$. (So each P'_j contains precisely one vertex from each cluster in $I_{4,\rm low}$.) Then $P'_j \cup P'_{j+1}$ forms a switchable pair with S' for all $j = 1, \dots, m$ (with indices considered modulo m). Denote the switch by $C'_{4,j}$.

(CyA2′) There is a potential switch $C'_{4,m+1}$ in $G - V_0$ so that S' contains two independent edges of $C_{4,m+1}$ but the other two edges of $C'_{4,m+1}$ do not lie in $CyA^{\rm orig}$. Moreover, $V(C'_{4,j}) \subseteq \bigcup_{V \in I_{4,\rm low}} V$ for each $j = 1, \dots, m+1$ and all the $C'_{4,j}$ are pairwise vertex-disjoint.

(CyA3′) For each $j = 1, \dots, m+1$, denote the edges of $C'_{4,j}$ which are contained in S' by ℓ'_j and r'_j. Let L' be the (ordered) sequence of edges $\ell'_1, \dots, \ell'_{m+1}$ and R' be the (ordered) sequence of edges $r'_1, \dots, r'_{m+1}$. S' is a bicycle on $V(G) \setminus V_0$ where one cycle contains all edges of L' in the given order and the other cycle contains all edges of R' in the given order.

The switches $C_{4,j}$ and $C'_{4,j}$ with $j < m$ will be used in the proof of Lemma 10.2 to 'transform' a given 1-regular graph H whose edges wind around C into a Hamilton cycle. We will not actually use the two switches $C_{4,m}$ and $C'_{4,m}$ defined above. However, they make our description of the construction of F a little simpler. The potential switches $C_{4,m+1}$ and $C'_{4,m+1}$ will be used to 'attach' the cycle absorber to the parity switcher defined in Section 10.4. This will ensure that the 'leftover' of the cycle absorber after the above transformation step also has a Hamilton decomposition.

Note that $CyA^{\rm orig}$ is a spanning subdigraph of G in which the vertices in V_0 have in- and outdegree 4, while the others have in- and outdegree 3. However, CyA is not actually a subdigraph of G, so saying that it is a cycle absorber in G is a slight abuse of notation.

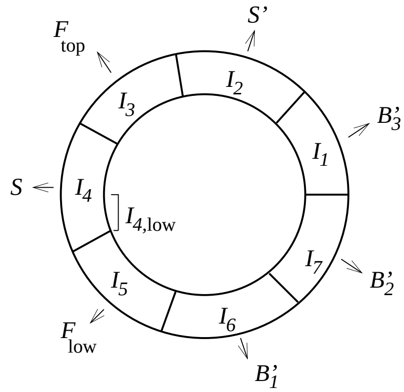

FIGURE 5. The exceptional factor which contains the complete exceptional path systems $CEPS_i$ of the (parity extended) cycle absorber. Each $CEPS_i$ spans the interval I_i and the diagram shows how these are assigned to the 1-factors of the cycle absorber in (CyA0) and to the bicycles B_i' of the parity switcher in the proof of Lemma 10.5.

The complete exceptional path systems $CEPS_i$ contained in CyA will be chosen within a single exceptional factor, which has parameters $(1, 7)$ (see Figure 5). The assignment of the $CEPS_i$ to the different 1-factors of the cycle absorber (and the parity switcher) is chosen in such a way that the switches of the 1-factor can be chosen to be vertex disjoint from the $CEPS_i$ contained in this 1-factor.

10.2. **Using the cycle absorber.** The following lemma shows that given an arbitrary 1-factor H of $G - V_0$ which winds around C and a cycle absorber CyA, we can 'almost' decompose $H \cup CyA$ into Hamilton cycles: we obtain a decomposition into at least two Hamilton cycles and at most two spanning bicycles in G. The final step of transforming the bicycles into Hamilton cycles is done by means of a 'parity switcher', defined in Section 10.4. (As discussed at the beginning of Section 10.4, the difficult case is when Lemma 10.2 yields a decomposition with exactly three Hamilton cycles and exactly one bicycle.)

Lemma 10.2. *Suppose that* $0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll 1$, *that* $k/14 \in \mathbb{N}$ *and that* $(G, \mathcal{P}, R, C)$ *is a* (k, m, ε, d)*-scheme with* $|G| = n$. *Suppose that* H *is a* 1*-factor of* $G - V_0$ *which winds around* C. *Let* CyA *be a cycle absorber with respect to* C *in* G *such that* $CyA^{\rm orig}$ *and* H *are edge-disjoint. Then* $H \cup CyA^{\rm orig}$ *has a decomposition into four* 1*-factors* F_1, F_2, F_3 *and* F_4 *of* G *satisfying the following conditions:*

(i) F_1 *contains* $CEPS_3^{\rm orig}$, F_2 *contains* $CEPS_5^{\rm orig}$, F_3 *contains* $CEPS_4^{\rm orig}$ *and* F_4 *contains* $CEPS_2^{\rm orig}$ *(where the* $CEPS_i$ *are as defined in (CyA0)).*
(ii) F_1 *and* F_2 *are Hamilton cycles in* G.
(iii) *Each of* F_3 *and* F_4 *is either a Hamilton cycle or a spanning bicycle in* G. *If* F_3 *is a bicycle, then one of the cycles contains* ℓ_{m+1} *and the other contains* r_{m+1}. *Similarly, if* F_4 *is a bicycle, then one of the cycles contains* ℓ'_{m+1} *and the other contains* r'_{m+1}.

Proof. Similarly as for F (as defined in (CyA0)), we partition H into $H_{\rm top}$ and $H_{\rm low}$. So both $F_{\rm top}$ and $H_{\rm top}$ consist of m vertex-disjoint paths from V_1 to $V_{k/2+1}$ and both $F_{\rm low}$ and $H_{\rm low}$ consist of m vertex-disjoint paths from $V_{k/2+1}$ to V_1. Let $T := H_{\rm low} \cup F_{\rm top}$ and let $T' := H_{\rm top} \cup F_{\rm low}$. Then (CyA0) implies that both T and T' are 1-regular digraphs on $V(G) \setminus V_0$ which agree with C and correspond to 1-factors $T^{\rm orig}$ and $(T')^{\rm orig}$ of G.

Our first aim is to perform switches between T and S to transform T into a Hamilton cycle on $V(G) \setminus V_0$. ($T^{\rm orig}$ will then turn out to be a Hamilton cycle of G.) Suppose that T is not a Hamilton cycle. For $j = 1, \dots, m$, let P_j be as defined in (CyA1). Recall from (CyA0) that $CEPS_3$ is the complete exceptional path system contained in $T_{\rm top} = F_{\rm top}$. But since $CEPS_3$ completely spans I_3, the paths in $CEPS_3$ link all vertices of $V_{2k/7+1}$ to those in $V_{3k/7+1}$. It follows that every cycle D in T visits every cluster on C except possibly $V_{2k/7+2}, \dots, V_{3k/7}$. In particular, the following assertion holds (where the P_j are as defined in (CyA1)):

For each $j = 1, \dots, m$, any cycle D in T either contains P_j or it avoids all vertices of P_j. Moreover, D contains at least one of the P_j and so it contains the jth vertex x_j of V_1 for some j with $1 \le j \le m$. $(\star)$

We say that i with $1 \le i < m$ is a *switch index for* T if x_i and x_{i+1} lie on different cycles of T (i.e. if the initial vertices of P_i and P_{i+1} lie on different cycles of T). Since T is not a Hamilton cycle, $(\star)$ implies that there must be an i with $1 \le i < m$ which is a switch index. Our approach will be to perform a switch between the cycles D and D' which contain x_i and x_{i+1} respectively. This will reduce the number of cycles of T and turn S into a Hamilton cycle. We continue in this way until T is a Hamilton cycle. The only difference in the later steps is that S might already be a Hamilton cycle, in which case it is transformed into a bicycle after the switch.

More precisely, for $i = 1, \dots, m+1$, we define $L_i := (\ell_i, \dots, \ell_{m+1})$ and $R_i := (r_i, \dots, r_{m+1})$. So $L_1 = L$ and $R_1 = R$ (where ℓ_i, r_i, L, R are as defined in (CyA3)). Suppose that $1 \le i < m$ and that T_i and S_i are 1-regular digraphs on $V(G) \setminus V_0$ which satisfy (a$_i$) and (b$_i$) below as well as either (CyA3$_i^-$) or (CyA3$_i^+$):

- (a$_i$) Let D be any cycle of T_i. For each $j = i+1, \dots, m$, D either contains P_j or it avoids all vertices of P_j. Moreover, let e_{i-1} denote the edge in $C_{4,i-1} \cap P_i$. Then D either contains all edges in $E(P_i) \setminus \{e_{i-1}\}$ (possibly D even contains P_i) or D avoids all vertices of P_i.
- (b$_i$) All of $x_1, \dots, x_i$ lie on a common cycle in T_i.
- (CyA3$_i^+$) S_i is a bicycle on $V(G) \setminus V_0$ where one cycle contains all edges of L_i in the given order and the other cycle contains all edges of R_i in the given order.
- (CyA3$_i^-$) S_i is a Hamilton cycle on $V(G) \setminus V_0$ which contains all edges of L_i and R_i in the given order and where all edges of L_i come before all edges of R_i.

Thus if $i = 1$, $T_1 := T$ and $S_1 := S$, then $(\star)$ and (CyA3) together imply that (a$_i$), (b$_i$) and (CyA3$_i^+$) hold. (Note that we view $\{e_0\}$ as being the empty set, so the last part of (a$_1$) says that D either contains all edges of P_1 or D avoids all vertices of P_1.) So suppose first that (CyA3$_i^+$) holds for some $1 \le i < m$. We define a switch index for T_i in the same way as for T. If i is not a switch index, let $S_{i+1} := S_i$ and $T_{i+1} := T_i$. Then clearly (a$_{i+1}$), (b$_{i+1}$) and (CyA3$_{i+1}^+$) hold. If i is a switch

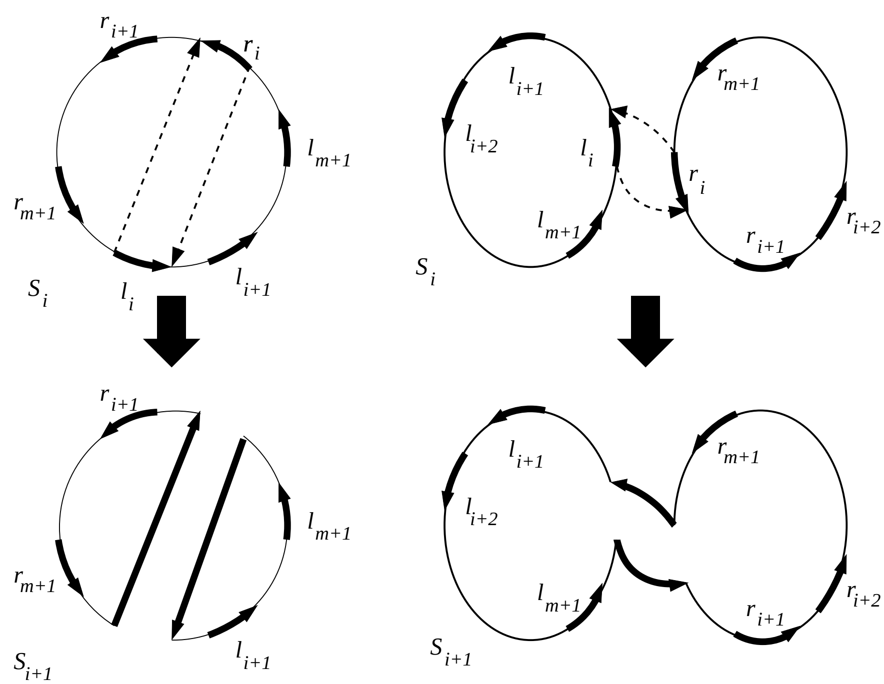

FIGURE 6. Transforming S_i into S_{i+1}. The left hand side illustrates the case when S_i satisfies (CyA3$_i^-$) and the right hand side illustrates the case when S_i satisfies (CyA3$_i^+$).

index, let D be the cycle of T_i which contains x_i and let D' be the cycle of T_i which contains x_{i+1}. Then (a$_i$) implies that D contains all edges in $E(P_i) \setminus \{e_{i-1}\}$ and D' contains P_{i+1}. But $e_{i-1} \notin C_{4,i}$ since $C_{4,i-1}$ and $C_{4,i}$ are vertex-disjoint by (CyA2). Thus we can carry out the $C_{4,i}$-exchange to obtain S_{i+1} and T_{i+1}. Then Proposition 10.1 implies that the vertices of D and D' now lie on a common cycle D'' of T_{i+1} In particular, T_{i+1} satisfies (b$_{i+1}$). Moreover, this new cycle D'' will contain all edges in $E(P_{i+1}) \setminus \{e_i\}$. Together with the fact that $C_{4,i}$ avoids all the P_j for $j = i+2, \dots, m$, it follows that T_{i+1} satisfies (a$_{i+1}$). Moreover, it is easy to see that S_{i+1} satisfies (CyA3$_{i+1}^-$) (see Figure 6). Suppose next that (CyA3$_i^-$) holds for some $1 \le i < m$. If i is not a switch index, let $S_{i+1} := S_i$ and $T_{i+1} := T_i$. Then clearly (a$_{i+1}$), (b$_{i+1}$) and (CyA3$_{i+1}^-$) hold. If i is a switch index, define D and D' as above. Again carry out the $C_{4,i}$-exchange to obtain S_{i+1} and T_{i+1}. Again, Proposition 10.1 implies that the vertices of D and D' now lie on a common cycle of T_{i+1}. Also, it is easy to see that (a$_{i+1}$), (b$_{i+1}$) and (CyA3$_{i+1}^+$) hold (see Figure 6).

So, by induction, T_m and S_m satisfy (a$_m$) and (b$_m$) as well as one of (CyA3$_m^+$) and (CyA3$_m^-$). Moreover, in both of the above cases, Proposition 10.1(ii) implies that all vertices which lie on a common cycle in T_i still lie on a common cycle in T_{i+1}. Together with (b$_m$) and the fact that by ($\star$) every cycle in $T = T_1$ contains x_j for some $1 \le j \le m$ this means that T_m is a Hamilton cycle. Note that both $CEPS_3$ and $CEPS_4$ are edge-disjoint (actually even vertex-disjoint) from all $C_{4,j}$ (by (CyA2) and the fact that $CEPS_3$ spans I_3 and $CEPS_4$ spans I_4). Thus both

$CEPS_3$ and $CEPS_4$ are unaffected by the switches we carried out and so T_m and S_m still contain $CEPS_3$ and $CEPS_4$ respectively. So we can take $F_1 := T_m^{\rm orig}$ and $F_3 := S_m^{\rm orig}$. (In particular, Observation 7.4 implies that F_1 is a Hamilton cycle of G and the argument for F_3 is similar.).

In a similar way, we obtain F_2 from T' and F_4 from S'. This time, we let $T'_1 := T'$ and $S'_1 := S'$. We then carry out the above procedure with S'_i, T'_i, P'_j playing the roles of S_i, T_i, P_j to obtain a Hamilton cycle T'_m on $V(G) \setminus V_0$. Since by (CyA2′) both $CEPS_2$ and $CEPS_5$ are edge-disjoint from all $C'_{4,j}$, both $CEPS_2$ and $CEPS_5$ are unaffected by the switches we carried out and so T'_m and S'_m still contain $CEPS_5$ and $CEPS_2$ respectively. We let $F_2 := (T'_m)^{\rm orig}$ and $F'_4 := (S'_m)^{\rm orig}$. □

10.3. **Finding the cycle absorber.** The following lemma guarantees the existence of a cycle absorber in a (k, m, ε, d)-scheme.

Lemma 10.3. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll 1$, that $k/14, m/2 \in \mathbb{N}$ and that $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme with $|G| = n$ and $C = V_1 \dots V_k$. Let $I_1, \dots, I_7$ be the canonical interval partition of C into 7 intervals of equal length. Let EF be an exceptional factor with parameters $(1, 7)$ with respect to C, $\mathcal{P}$. For each $i = 1, \dots, 7$, let $CEPS_i$ be the complete exceptional path system which is contained in EF and completely spans I_i. Then there is a cycle absorber CyA with respect to C in G which satisfies the following properties:*

- (i) *$CEPS_2, \dots, CEPS_5$ are the complete exceptional path systems described in (CyA0).*
- (ii) *$C_{4,m+1} \subseteq G[V_{12}, V_{13}]$ and $C'_{4,m+1} \subseteq G[V_{k/2+12}, V_{k/2+13}]$ (where $C_{4,m+1}$ and $C'_{4,m+1}$ are as defined in (CyA2) and (CyA2′) respectively).*

Proof. We first construct $F_{\rm top}$. Below, we assume the existence of an ordering of the vertices in any cluster V_i. Let $\mathcal{B}(C)$ be the union of $G[V_i, V_{i+1}]$ over all $i = 1, \dots, k$. So $\mathcal{B}(C)$ is a blow-up of C in which every edge of C corresponds to an $[\varepsilon, \geq d]$-superregular pair.

Claim 1. *$\mathcal{B}(C)$ contains a system $Q_1, \dots, Q_m$ of vertex-disjoint paths and m vertex-disjoint copies $C_{4,1}, \dots, C_{4,m}$ of C_4 such that the following properties are satisfied:*

- *Q_j joins the jth vertex of V_1 to the jth vertex of $V_{2k/7+1}$.*
- *For each $j = 1, \dots, m$, $C_{4,j}$ shares exactly one edge with Q_j and one edge with Q_{j+1} (where $Q_{m+1} := Q_1$).*
- *$V(C_{4,j}) \subseteq \bigcup_{V \in I_1} V$ for each $j = 1, \dots, m$.*

To prove the claim, we first apply the Blow-up lemma (Lemma 4.13) to $G[V_4, V_5]$ to find $m/2$ vertex-disjoint copies $C_{4,1}, C_{4,3}, \dots, C_{4,m-1}$ of C_4. Now we choose a perfect matching in each of the $C_{4,j}$. Next apply the Blow-up lemma to $G[V_8, V_9]$ to find $m/2$ vertex-disjoint copies $C_{4,2}, C_{4,4}, \dots, C_{4,m}$ of C_4. As before we choose a perfect matching in each of the $C_{4,j}$. This gives a perfect matching M_1 in $G[V_4, V_5]$ and a perfect matching M_2 in $G[V_8, V_9]$. Order the edges of each M_i such that the following properties are satisfied:

- Edges belonging to the same C_4 are consecutive.
- The edges of $C_{4,j+2}$ come directly after those of $C_{4,j}$ (modulo m).

- The edges of $C_{4,1}$ are the first two edges of M_1 and the edges of $C_{4,2}$ are the second and third edge of M_2.

These matchings induce a new ordering on the vertices in the clusters V_4, V_5, V_8, V_9 with which we replace the orderings chosen initially.

Now we apply Corollary 4.15 three times to obtain a system of m vertex-disjoint paths in $\mathcal{B}(C)$ which for each $j = 1, \dots, m$ link the jth vertex of V_1 to the jth vertex of V_4, a system of m vertex-disjoint paths which link the jth vertex of V_5 to the jth vertex of V_8 and a system of m vertex-disjoint paths which link the jth vertex of V_9 to the jth vertex of $V_{2k/7+1}$. The union of M_1, M_2 and all these path systems forms a system $Q_1, \dots, Q_m$ of paths as required in Claim 1.

Recall that the paths in $CEPS_3$ join $V_{2k/7+1}$ to $V_{3k/7+1}$. Now let $Q'_1, \dots, Q'_m$ be a system of vertex-disjoint paths in $\mathcal{B}(C)$ linking the vertices in $V_{3k/7+1}$ to those $V_{k/2+1}$ (in an arbitrary way). We let $F_{\text{top}} := Q_1 \cup \dots Q_m \cup CEPS_3 \cup Q'_1 \cup \dots Q'_m$.

F_{low} is constructed similarly, but this time we choose $C'_{4,1}, C'_{4,3}, \dots, C'_{4,m-1}$ in $G[V_{k/2+4}, V_{k/2+5}]$ and $C'_{4,2}, C'_{4,4}, \dots, C'_{4,m}$ in $G[V_{k/2+8}, V_{k/2+9}]$ (and F_{low} will contain $CEPS_5$).

So let us now construct S. Let $\mathcal{B}'(C)$ be obtained from $\mathcal{B}(C)$ by deleting all the edges in $F = F_{\text{top}} \cup F_{\text{low}}$ and define G' similarly (since this decreases the in- and the outdegrees by at most 1, the superregularity of the pairs $G'[V_i, V_{i+1}]$ is not affected significantly). Choose a copy $C_{4,m+1}$ of C_4 in $G'[V_{12}, V_{13}]$. Let ℓ_{m+1} and r_{m+1} be a matching in $C_{4,m+1}$. For each $j = 1, \dots, m$ let ℓ_j and r_j be the edges of $C_{4,j}$ which are not contained in F.

When we refer to an endvertex of an edge e below, this is allowed to be either the initial or the final vertex of e (it will be clear from the context which one is meant). For all odd j with $1 \le j < m$, apply Corollary 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}'(C)$ linking the endvertex of ℓ_j in V_5 to the endvertex of ℓ_{j+1} in V_8 and the endvertex of r_j in V_5 to the endvertex of r_{j+1} in V_8. Apply Corollary 4.15 again to find a system of m vertex-disjoint paths in $\mathcal{B}'(C)$ such that one of these paths links the endvertex of ℓ_m in V_9 to the endvertex of ℓ_{m+1} in V_{12}, another path links the endvertex of r_m in V_9 to the endvertex of r_{m+1} in V_{12} and the remaining $m-2$ paths join the remaining vertices in V_9 to those in V_{12}. We also choose a matching M_3 in $G'[V_{12}, V_{13}]$ which consists of $m-2$ edges and avoids the endvertices of ℓ_{m+1} and r_{m+1} (we find M_3 by applying Proposition 4.14 to the subgraph of $G'[V_{12}, V_{13}]$ obtained by deleting the endvertices of ℓ_{m+1} and r_{m+1}). Next we apply Corollary 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}'(C)$ joining the vertices in V_{13} to those in $V_{3k/7+1}$.

Let $\mathcal{Q}$ denote the system of paths obtained from the union of all the paths chosen previously, of $\ell_1, \dots, \ell_{m+1}$, of $r_1, \dots, r_{m+1}$ and of the edges in M_3. So each path in $\mathcal{Q}$ joins some vertex in V_4 to some vertex in $V_{3k/7+1}$. Consider the system $\mathcal{Q}'$ of paths obtained by concatenating the paths in $\mathcal{Q}$ with those in $CEPS_4$ (recall that $CEPS_4$ completely spans the interval $I_4 = V_{3k/7+1} \dots V_{4k/7+1}$). For each even j with $1 \le j \le m$, let r'_j denote the vertex in $V_{4k/7+1}$ which is connected to the edge r_j by a path in $\mathcal{Q}'$ and define ℓ'_j similarly. Apply Corollary 4.15 to find a system $\mathcal{Q}''$ of m vertex-disjoint paths in $\mathcal{B}'(C)$ from $V_{4k/7+1}$ to V_4 such that for each even j with

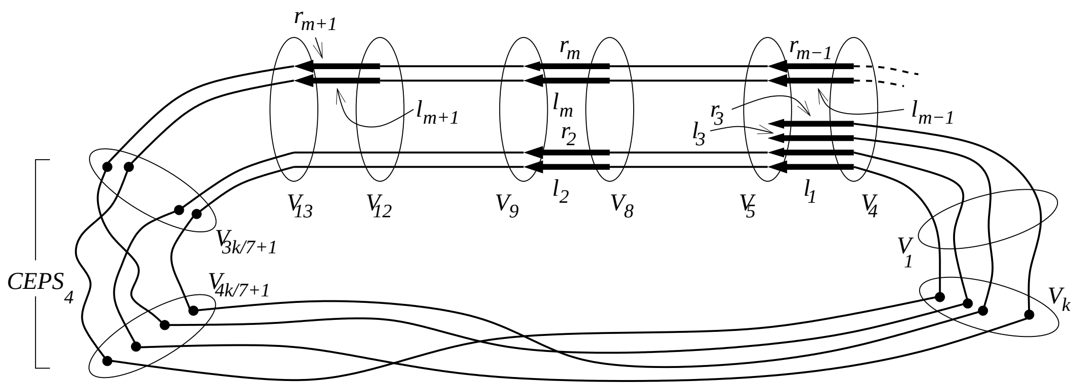

FIGURE 7. The construction of S.

$1 \leq j < m$ the vertex r'_j is linked to the endvertex of r_{j+1} in V_4 and ℓ'_j is linked to the endvertex of ℓ_{j+1} in V_4. Also, ℓ'_m is linked to ℓ_1 and r'_m is linked to r_1 (see Figure 7). We let S be the union of all the paths in $\mathcal{Q}'$ and $\mathcal{Q}''$. Then F and S satisfy (CyA0)–(CyA3).

S' is constructed similarly. This is possible as the switches $C'_{4,1}, \dots, C'_{4,m+1}$ lie in $I_{4,\mathrm{low}} \subseteq I_4$ whereas the complete exceptional path system $CEPS_2$ contained in S' spans I_2 (see Figure 5). $\square$

10.4. **The parity switcher.** Suppose that we are given a decomposition $\mathcal{D}$ of a regular digraph into r edge-disjoint 1-factors and suppose that the total number of cycles is K, say. If we carry out C_4-exchanges between the cycles in these 1-factors, then Proposition 10.1 implies that the resulting total number of cycles either stays the same, increases by two or decreases by two after each exchange. So if e.g. r is odd and K is even, then we will never be able to transform $\mathcal{D}$ into a set of edge-disjoint Hamilton cycles if we rely only on C_4-exchanges between cycles in $\mathcal{D}$. The following concept of 'triple switches' will allow us to change the parity of the total number of cycles in a decomposition. In particular, the resulting parity switcher will allow us to transform the spanning bicycles which are potentially returned by Lemma 10.2 into Hamilton cycles.

Let $K_{2,3}$ denote the orientation of a complete bipartite graph with vertex classes of size 2 and 3 in which every edge is oriented towards the vertex class of size 3. Note that $K_{2,3}$ is the edge-disjoint union of three matchings M_1, M_2 and M_3 of size 2. Given edge-disjoint bicycles B_1, B_2 and B_3 on the same vertex set, we say that they are *triply-switchable* if for each $i = 1, 2, 3$ there are independent edges ℓ_i and r_i lying on different cycles of B_i such that the union of ℓ_i and r_i over all $i = 1, 2, 3$ forms a copy $K^*_{2,3}$ of $K_{2,3}$. We say that $K^*_{2,3}$ is a *$B_1B_2B_3$-switch*. A *$K^*_{2,3}$-exchange* consists of deleting the edges $\ell_i =: xx_i$ and $r_i =: yy_i$ from B_i and adding the edges xy_i and yx_i for each $i = 1, 2, 3$ (see Figure 8). Thus for each $i = 1, 2, 3$ one of the edges ℓ_i, r_i will be added to B_{i+1} and the other edge will be added to B_{i+2} (where the $B_4 := B_1$ and $B_5 := B_2$). Note that the digraph obtained from B_i via a $K^*_{2,3}$-exchange is a Hamilton cycle (for each $i = 1, 2, 3$). Thus the union of three

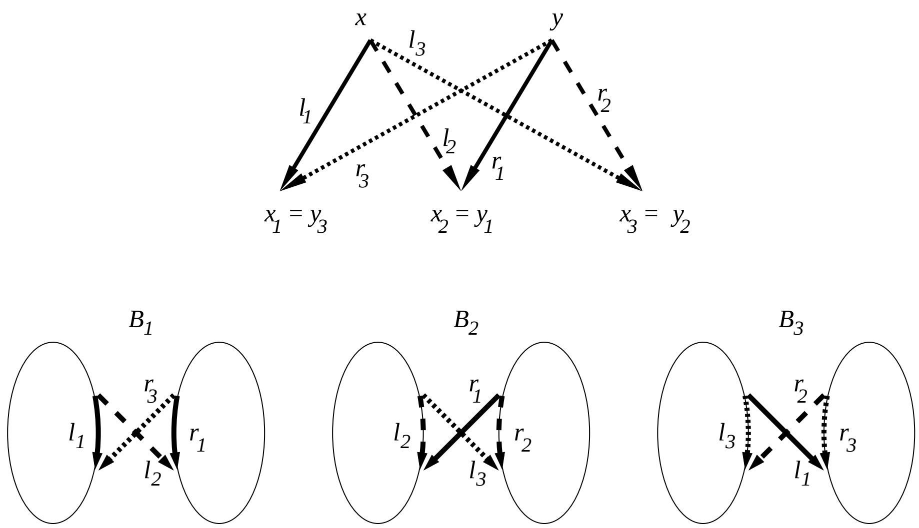

FIGURE 8. Illustrating a $K^*_{2,3}$-exchange.

edge-disjoint triply-switchable bicycles has a decomposition into three edge-disjoint Hamilton cycles as well as a decomposition into six cycles (two for each B_i). As mentioned above, this parity difference will enable us to turn the bicycle(s) which are potentially returned by Lemma 10.2 into Hamilton cycles (without creating any additional bicycles elsewhere).

A *parity extended cycle absorber PCA in G* with respect to C is digraph on $V(G) \setminus V_0$ with the following properties:

(PCA1) PCA is the union of two digraphs CyA and TSB, each with vertex set $V(G) \setminus V_0$. $CyA = F \cup S \cup S'$ is a cycle absorber in G. TSB is the union of three 1-regular digraphs B'_1, B'_2 and B'_3 on $V(G) \setminus V_0$ such that each B'_i contains a complete exceptional path system $CEPS(B'_i)$ (but no other exceptional edges) and such that B'_i is a bicycle on $V(G) \setminus V_0$. Let $B_i := (B'_i)^{\rm orig}$. Then $CyA^{\rm orig}$, B_1, B_2 and B_3 are pairwise edge-disjoint subdigraphs of G.

(PCA2) B_1, B_2 and B_3 are triply-switchable. Let $K^*_{2,3}$ denote the corresponding $B_1B_2B_3$-switch and let ℓ^*_i and r^*_i denote the edges of $K^*_{2,3}$ contained in B_i. (So ℓ^*_i and r^*_i lie on different cycles of B_i.)

(PCA3) Recall that ℓ_{m+1} and r_{m+1} are the two edges of the switch $C_{4,m+1}$ which are contained in S (where $C_{4,m+1}$ is as defined in (CyA2)). Then B_1 contains the other two edges ℓ_S and r_S of $C_{4,m+1}$. Similarly, recall that ℓ'_{m+1} and r'_{m+1} are the two edges of the switch $C'_{4,m+1}$ which are contained in S' (where $C'_{4,m+1}$ is as defined in (CyA2′)). Then B_1 also contains the other two edges $\ell_{S'}$ and $r_{S'}$ of $C'_{4,m+1}$.

(PCA4) B_2 and B_3 are pairwise switchable. Let C_{4,B_2B_3}-denote the corresponding switch.

(PCA5) All the switches $K_{2,3}^*$, $C_{4,m+1}$, $C'_{4,m+1}$ and C_{4,B_2B_3} are pairwise edge-disjoint. Both $C_{4,m+1}$ and $C'_{4,m+1}$ are vertex-disjoint from $CEPS(B'_1)$ while both $K_{2,3}^*$ and C_{4,B_2B_3} are vertex-disjoint from $CEPS(B'_i)$ for all $i = 1, 2, 3$.

(PCA6) One cycle of the bicycle B_1 contains the edges $\ell_S, \ell_{S'}, \ell_1^*$ in that order while the other cycle of B_1 contains $r_S, r_{S'}, r_1^*$ in that order.

An *s-fold parity extended cycle absorber $PCA(s)$ with respect to C in G* consists of s parity extended cycle absorbers whose original versions are pairwise edge-disjoint. Note that $PCA(s)$ is a $6s$-regular spanning subdigraph of G^{basic} which contains precisely $7s$ complete exceptional path systems (and no other exceptional edges). Moreover, $PCA(s)^{\text{orig}}$ is a spanning subdigraph of G and

$$d^{\pm}(x) = 7s \quad \forall x \in V_0 \qquad \text{and} \qquad d^{\pm}(y) = 6s \quad \forall y \in V(G) \setminus V_0. \tag{10.1}$$

Lemma 10.4. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll 1$, that $k/14 \in \mathbb{N}$ and that $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme with $|G| = n$. Suppose that H is an s-factor of $G - V_0$ such that H is a blow-up of C. Let $PCA(s)$ be an s-fold parity extended cycle absorber with respect to C in G such that $PCA(s)^{\text{orig}}$ and H are edge-disjoint. Then $H \cup PCA(s)^{\text{orig}}$ has a decomposition into $7s$ edge-disjoint Hamilton cycles of G. Moreover, each of these Hamilton cycles contains the original version of one of the $7s$ complete exceptional path systems contained in $PCA(s)^{\text{orig}}$.*

Proof. Let $PCA_1, \dots, PCA_s$ denote the parity extended cycle absorbers contained in $PCA(s)$. We apply Proposition 6.1 to find a 1-factorization of H into $H_1, \dots, H_s$. Since H is a blow-up of C, each H_i winds around C. We claim that $H_i \cup PCA_i^{\text{orig}}$ has a decomposition into Hamilton cycles $C_{7(i-1)+1}, \dots, C_{7i}$ of G.

Note that the claim follows if we can prove it for the case $i = 1$. So let CyA and $TSB = B'_1 \cup B'_2 \cup B'_3$ be as defined in (PCA1)–(PCA6). Apply Lemma 10.2 to obtain a decomposition of $H_1 \cup CyA^{\text{orig}}$ into 1-factors $F_1, \dots, F_4$ satisfying the following conditions:

(a) F_1 and F_2 are Hamilton cycles in G.

(b) Each of F_3 and F_4 is either a Hamilton cycle or a spanning bicycle in G. If F_3 is a bicycle, then one of the cycles contains ℓ_{m+1} and the other contains r_{m+1}. Similarly, if F_4 is a bicycle, then one of the cycles contains ℓ'_{m+1} and the other contains r'_{m+1}.

(c) Each F_i contains (the original version of) a complete exceptional path system $CEPS_i^*$ (which is one of those contained in $PCA(1)^{\text{orig}}$). Moreover, $CEPS_3^*$ spans I_4 and $CEPS_4^*$ spans I_2 (where I_i is the ith interval of the canonical interval partition into 7 intervals defined earlier).

If both F_3 and F_4 are Hamilton cycles, we perform the $K_{2,3}^*$-exchange to decompose $TSB^{\text{orig}} = B_1 \cup B_2 \cup B_3$ into three edge-disjoint Hamilton cycles of G. Then these Hamilton cycles together with $F_1, \dots, F_4$ form a Hamilton decomposition of $H_1 \cup PCA_1^{\text{orig}}$.

So suppose next that both F_3 and F_4 are bicycles. First perform the $C_{4,m+1}$-exchange. This turns both F_3 and B_1 into Hamilton cycles. Let B_1^1 denote the Hamilton cycle obtained from B_1. Then (PCA6) implies that the edges $\ell_{S'}, \ell_1^*, r_{S'}, r_1^*$ appear on B_1^1 in that order. Next we perform the $C'_{4,m+1}$-exchange. This turns F_4

into a Hamilton cycle and B_1^1 into a bicycle B_1^2. Note that one of the cycles of B_1^2 contains ℓ_1^* while the other cycle contains r_1^* (this is a special case of the argument illustrated in Figure 6). So we can now perform the $K_{2,3}^*$-exchange. This turns each of B_1^2, B_2 and B_3 into a Hamilton cycle. Altogether this gives a Hamilton decomposition of $H_1 \cup PCA_1^{\rm orig}$.

So suppose next that F_3 is a bicycle but F_4 is a Hamilton cycle. In this case we perform the $C_{4,m+1}$-exchange. This turns both F_3 and B_1 into Hamilton cycles. We then perform the C_{4,B_2B_3}-exchange. This turns both B_2 and B_3 into Hamilton cycles. As before, altogether this gives a Hamilton decomposition of $H_1 \cup PCA_1^{\rm orig}$.

The case when F_3 is a Hamilton cycle but F_4 is a bicycle is similar to the previous case, but we perform the $C'_{4,m+1}$-exchange instead of the $C_{4,m+1}$-exchange.

The 'moreover part' follows since the switches involved in the above argument are edge-disjoint from the complete exceptional path systems of the bicycles and Hamilton cycles involved in the corresponding exchanges. More precisely, for the Hamilton cycles originating from the B_i, the 'moreover part follows from (PCA1) and (PCA5). For F_1 and F_2, it follows from (a) and (c) above. For F_3, it follows from (CyA2) and (c), and for F_4 it follows from (CyA2$'$) and (c). □

Lemma 10.5. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll d \ll 1$, that $s/m \ll d$ and that $k/14, m/2 \in \mathbb{N}$. Let $(G, \mathcal{P}, R, C)$ be a (k, m, ε, d)-scheme with $|G| = n$. Suppose that $EF_1, \dots, EF_s$ are exceptional factors with parameters $(1, 7)$ with respect to C, $\mathcal{P}$ whose original versions are pairwise edge-disjoint. Then there exists an s-fold parity extended cycle absorber $PCA(s)$ with respect to C in G such that the $7s$ complete exceptional path systems contained in $PCA(s)$ are precisely those in $EF_1 \cup \dots \cup EF_s$.*

Proof. Choose an additional constant ε' with $\varepsilon, s/m \ll \varepsilon' \ll d$. Let G' be the digraph obtained from G by deleting all the edges in $EF_1^{\rm orig}, \dots, EF_s^{\rm orig}$. Note that $|N_G^+(x) \setminus N_{G'}^+(x)| \le 7s \le (\varepsilon'/3)^2 m$ for every vertex x of G and the analogous condition holds for the inneighbourhoods of x. Thus Lemma 7.1(ii) implies that $(G', \mathcal{P}, R, C)$ is still a (k, m, ε', d)-scheme.

Let us first show how to find a single parity extended cycle absorber PCA_1 with respect to C in G'. Let $I_1, \dots, I_7$ be the canonical interval partition of C into 7 intervals $I_1, \dots, I_7$ of equal length and let $CEPS_i$ denote the complete exceptional path system in EF_1 which completely spans I_i (for all $i = 1, \dots, 7$). Apply Lemma 10.3 to find a cycle absorber CyA with respect to C in G' which contains $CEPS_2, \dots, CEPS_5$ (in the way described in (CyA0)).

Recall that ℓ_S and r_S denote the edges of the (potential) switch $C_{4,m+1}$ described in (CyA2) and (PCA3) which are not contained in CyA. Similarly, recall $\ell_{S'}$ and $r_{S'}$ denote the edges of the (potential) switch $C'_{4,m+1}$ described in (CyA2$'$) and (PCA3) which are not contained in CyA. Recall from Lemma 10.3 that ℓ_S and r_S lie in $G'[V_{12}, V_{13}]$ and that $\ell_{S'}$ and $r_{S'}$ lie in $G'[V_{k/2+12}, V_{k/2+13}]$. Remove the edges of $CyA^{\rm orig}$ from G' to obtain G''. Let $\mathcal{B}''(C)$ be the union of $G''[V_i, V_{i+1}]$ over all $i = 1, \dots, k$. Proposition 4.3(iii) implies that $\mathcal{B}''(C)$ is a blow-up of C in which every edge of C corresponds to an $[2\sqrt{\varepsilon'}, \ge d]$-superregular pair.

When we refer to an endvertex of an edge e below, this is allowed to be either the initial or the final vertex of e (it will be clear from the context which one is meant). Our next aim is to find TSB. We start by finding B'_1. First we apply Proposition 4.14 to choose a perfect matching in $G''[V_{12}, V_{13}]$ which extends ℓ_S and r_S. Now apply Corollary 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}''(C)$ which link all vertices of V_{13} to those in $V_{k/2+12}$ such that the endvertex of ℓ_S in V_{13} is linked to the endvertex of $\ell_{S'}$ in $V_{k/2+12}$ and the endvertex of r_S in V_{13} is linked to the endvertex of $r_{S'}$ in $V_{k/2+12}$. Choose a perfect matching in $G''[V_{k/2+12}, V_{k/2+13}]$ which extends $\ell_{S'}$ and $r_{S'}$.

Next choose a copy $K^*_{2,3}$ of $K_{2,3}$ in $G''[V_{k/2+16}, V_{k/2+17}]$. For each $i = 1, 2, 3$ let ℓ^*_i and r^*_i be two independent edges in $K^*_{2,3}$ such that all these six edges are distinct from each other (and so $K^*_{2,3}$ is the union of all these six edges). Now apply Lemma 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}''(C)$ which link the vertices in $V_{k/2+13}$ to those in $V_{k/2+16}$ in such a way that the endvertex of $\ell_{S'}$ in $V_{k/2+13}$ is linked to the endvertex of ℓ^*_1 in $V_{k/2+16}$ and the endvertex of $r_{S'}$ in $V_{k/2+13}$ is linked to the endvertex of r^*_1 in $V_{k/2+16}$. Extend ℓ^*_1 and r^*_1 into a perfect matching of $G''[V_{k/2+16}, V_{k/2+17}]$. Now apply Lemma 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}''(C)$ which (arbitrarily) link all vertices of $V_{k/2+17}$ to those in $V_{5k/7+1}$. Altogether this gives us a system $\mathcal{Q}$ of m vertex-disjoint paths joining all vertices in V_{12} to all vertices in $V_{5k/7+1}$.

Now extend this path system by concatenating the paths in $\mathcal{Q}$ with the ones forming $CEPS_6$. (Recall that $CEPS_6$ spans the interval $I_6 = V_{5k/7+1} \dots V_{6k/7+1}$ completely.) Denote the resulting path system by $\mathcal{Q}'$. Order the vertices of V_{12} so that the first vertex is the endvertex of ℓ_S and the last vertex is the endvertex of r_S. For each $j = 1, \dots, m$, let $q_j \in V_{6k/7+1}$ denote the vertex which is linked to the jth vertex of V_{12} by a path in $\mathcal{Q}'$. Thus the path in $\mathcal{Q}'$ ending in q_1 contains ℓ_S, $\ell_{S'}$, ℓ^*_1 (in that order) while the path in $\mathcal{Q}'$ ending in q_m contains r_S, $r_{S'}$, r^*_1 (in that order).

Now apply Lemma 4.15 to find a system $\mathcal{Q}''$ of m vertex-disjoint paths in $\mathcal{B}''(C)$ which link each q_j to the $(j+2)$nd vertex in V_{12} (where the indices are considered modulo m). Let B'_1 be the union of all the paths in $\mathcal{Q}'$ and $\mathcal{Q}''$. Then B'_1 is a bicycle on $V(G) \setminus V_0$ and $B_1 := (B'_1)^{\rm orig} = \mathcal{Q} \cup CEPS_6^{\rm orig} \cup \mathcal{Q}''$ satisfies (PCA3) and (PCA6).

We next show how to find B'_2. Remove all edges of B_1 from G'' to obtain G'''. Let $\mathcal{B}'''(C)$ be the union of $G'''[V_i, V_{i+1}]$ over all $i = 1, \dots, k$. Extend ℓ^*_2 and r^*_2 into a perfect matching of $G'''[V_{k/2+16}, V_{k/2+17}]$. Choose a copy C_{4,B_2B_3} of C_4 in $G'''[V_{k/2+20}, V_{k/2+21}]$. Choose two independent edges $\ell^\diamond_2$ and $r^\diamond_2$ of C_{4,B_2B_3} and let $\ell^\diamond_3$ and $r^\diamond_3$ denote the other two edges. Now apply Lemma 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}'''(C)$ which link all vertices of $V_{k/2+17}$ to those in $V_{k/2+20}$ such that the endvertex of ℓ^*_2 in $V_{k/2+17}$ is linked to the endvertex of $\ell^\diamond_2$ in $V_{k/2+20}$ and the endvertex of r^*_2 in $V_{k/2+17}$ is linked to the endvertex of $r^\diamond_2$ in $V_{k/2+20}$. Extend $\ell^\diamond_2$ and $r^\diamond_2$ into a perfect matching of $G'''[V_{k/2+20}, V_{k/2+21}]$. Now apply Lemma 4.15 to find a system of m vertex-disjoint paths in $\mathcal{B}'''(C)$ which (arbitrarily) link all vertices of $V_{k/2+21}$ to those in $V_{6k/7+1}$. Altogether this gives us a system $\mathcal{S}$ of m vertex-disjoint paths joining all vertices in $V_{k/2+16}$ to all vertices in $V_{6k/7+1}$.

Now extend this path system by concatenating the paths in $\mathcal{S}$ with the ones forming $CEPS_7$. (Recall that $CEPS_7$ spans the interval $I_7 = V_{6k/7+1} \dots V_1$ completely.)

Denote the resulting path system by $\mathcal{S}'$. Order the vertices of $V_{k/2+16}$ so that the first vertex is the endvertex of ℓ_2^* and the last vertex is the endvertex of r_2^*. For each $j = 1, \dots, m$, let $s_j \in V_1$ denote the vertex which is linked to the jth vertex of $V_{k/2+16}$ by a path in $\mathcal{S}'$. Now apply Lemma 4.15 to find a system $\mathcal{S}''$ of m vertex-disjoint paths in $\mathcal{B}'''(C)$ which link each s_j to the $(j+2)$nd vertex in $V_{k/2+16}$ (where the indices are considered modulo m). Let B_2' be the union of all the paths in $\mathcal{S}'$ and $\mathcal{S}''$. Then B_2' is a bicycle on $V(G) \setminus V_0$. B_3' can be chosen in a similar way as B_2' except that it contains $\ell_3^*, r_3^*, \ell_3^\diamond, r_3^\diamond$ and $CEPS_1$ instead of $\ell_2^*, r_2^*, \ell_2^\diamond, r_2^\diamond$ and $CEPS_7$. Let $TSB := B_1 \cup B_2 \cup B_3$ and $PCA_1 := CyA \cup TSB$. Then PCA_1 satisfies (PCA1)–(PCA6)

We now remove the edges of $PCA_1^{\rm orig}$ from G' and repeat the above process to find the remaining $s-1$ edge-disjoint parity extended cycle absorbers $PCA_2, \dots, PCA_s$. To see that this is possible, denote the digraph obtained from G' by the removal of the edges of $PCA_1^{\rm orig}, \dots, PCA_i^{\rm orig}$ by G_i (where $i < s$). Note that $|N^+_{G'}(x) \setminus N^+_{G_i}(x)| \le 7s \le \varepsilon' m$ for every vertex x of G' and the analogous condition holds for the inneighbourhoods of x. Thus by Lemma 7.1(ii) $(G_i, \mathcal{P}, R, C)$ is still a $(k, m, 3\sqrt{\varepsilon'}, d)$-scheme. □

11. Proof of Theorem 1.2

The following lemma shows that we can cover all edges induced by a small exceptional set using a small number of edge-disjoint Hamilton cycles.

Lemma 11.1. *Suppose that $0 < 1/n \ll \varepsilon \ll \nu \le \tau \ll \alpha \le 1$. Let G be a robust (ν, τ)-outexpander with $\delta^0(G) \ge \alpha n$ and let V_0 be a set of vertices in G with $|V_0| \le \varepsilon n$. Then there is a set of εn edge-disjoint Hamilton cycles in G which contain all edges of $G[V_0]$.*

Proof. Note that when viewed as an undirected graph, $G[V_0]$ has maximum degree less than εn. So by Vizing's theorem, we can partition the edges of $G[V_0]$ into $t := \varepsilon n$ matchings $M_1, \dots, M_t$ (some of these may be empty). For each matching M_i in turn, we find a Hamilton cycle C_i which contains all edges of M_i and which is edge-disjoint from $C_1, \dots, C_{i-1}$. Suppose that we have found $C_1, \dots, C_{i-1}$. Let G_i be the graph obtained from G by removing the edges of $C_1, \dots, C_{i-1}$. Note that G_i is still a robust $(\nu/2, \tau)$-outexpander with $\delta^0(G_i) \ge \alpha n - (i-1) \ge \alpha n/2$. Let G_i' be the graph obtained from G_i by contracting the edges of M_i. More precisely, we successively replace each directed edge ab of M_i by a vertex whose outneighbours are the current outneighbours of b and whose inneighbours are the current inneighbours of a. Then G_i' is still a robust $(\nu/3, 2\tau)$-outexpander with $\delta^0(G_i') \ge \alpha n/3$. Thus G_i' contains a Hamilton cycle C_i' by Theorem 6.2. But C_i' corresponds to a Hamilton cycle C_i in G_i (and thus G) containing M_i. So we can continue until we have found $C_1, \dots, C_t$ as required. □

Before proving Theorem 1.2, we will combine the chord absorbing step and the cycle absorbing step into a single 'robust decomposition' lemma. Roughly speaking, this means that we can find a sparse 'robustly decomposable digraph' $G^{\rm rob}$, so that for an arbitrary very sparse regular digraph H on the same vertex set, the digraph

$H \cup G^{\rm rob}$ always has a Hamilton decomposition. We will state a variant of this lemma in Section 12. This variant will be used in [14, 32]. The main advantage of combining the chord absorbing step and the cycle absorbing step into a single lemma is that in future applications one will not need to know the definitions of the chord absorber and (the more involved) definition of the cycle absorber in order to apply it.

Lemma 11.2. *Suppose that $0 < 1/n \ll 1/k \ll \varepsilon \ll 1/q \ll 1/f \ll r_1/m \ll d \ll 1/\ell', 1/g \ll 1$ and that $rk^2 \leq m$. Let*

$$r_2 := 96\ell' g^2 kr, \quad r_3 := rfk/q, \quad r^\diamond := r_1 + r_2 + r - (q-1)r_3, \quad s' := rfk + 7r^\diamond$$

and suppose that $k/14, k/f, k/g, q/f, m/4\ell', fm/q, 2fk/3g(g-1) \in \mathbb{N}$. Suppose that $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an $(\ell', k, m, \varepsilon, d)$-setup with $|G| = n$ and $C = V_1 \dots V_k$. Suppose that $\mathcal{P}^*$ is a (q/f)-refinement of $\mathcal{P}$ and that $EF_1, \dots, EF_{r_3}$ are exceptional factors with parameters $(q/f, f)$ with respect to C, $\mathcal{P}^*$ whose original versions are pairwise edge-disjoint. Let $\mathcal{EF}$ be the union of the EF_i over all $i = 1, \dots, r_3$. Then there exists a spanning subdigraph $CA^\diamond(r)$ of $G - V_0$ for which the following holds:

- (i) *$CA^\diamond(r)$ is an $(r_1 + r_2)$-regular spanning subdigraph of $G - V_0$ which is edge-disjoint from $\mathcal{EF}^{\rm orig}$.*
- (ii) *Suppose that $EF'_1, \dots, EF'_{r^\diamond}$ are exceptional factors with parameters $(1, 7)$ with respect to C, $\mathcal{P}$ such that the original versions of all these exceptional factors are pairwise edge-disjoint from each other and edge-disjoint from $CA^\diamond(r) \cup \mathcal{EF}^{\rm orig}$. Let $\mathcal{EF}'$ be the union of the EF'_i over all $i = 1, \dots, r^\diamond$. Then there exists a spanning subdigraph $PCA^\diamond(r)$ of $G - V_0$ for which the following holds:*
  - (a) *$PCA^\diamond(r)$ is a $5r^\diamond$-regular spanning subdigraph of $G - V_0$ which is edge-disjoint from $CA^\diamond(r) \cup (\mathcal{EF} \cup \mathcal{EF}')^{\rm orig}$.*
  - (b) *Let $\mathcal{CEPS}$ be the set consisting of all the s' complete exceptional path systems contained in $\mathcal{EF} \cup \mathcal{EF}'$. Whenever H is an r-factor of $G - V_0$ which is edge-disjoint from $G^{\rm rob} := CA^\diamond(r) \cup PCA^\diamond(r) \cup (\mathcal{EF} \cup \mathcal{EF}')^{\rm orig}$, then $H \cup G^{\rm rob}$ has a decomposition into s' edge-disjoint Hamilton cycles $C_1, \dots, C_{s'}$ of G. Moreover, for each $i = 1, \dots, s'$, the basic version $C_i^{\rm basic}$ of C_i contains one of the complete exceptional path systems from $\mathcal{CEPS}$.*

The analogue holds for an $(\ell', k, m, \varepsilon, d)$-bi-setup $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ if we assume in addition that H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$ (where $\mathcal{V}_{\rm even}$ is the set of all those V_i such that i is even and $\mathcal{V}_{\rm odd}$ is defined analogously).

Note that the definition of an exceptional factor and (i) together imply that in the original version $CA^\diamond(r) \cup \mathcal{EF}^{\rm orig}$ of $CA^\diamond(r) \cup \mathcal{EF}$ we have

$$d^\pm(x) = r_3 q \ \ \forall x \in V_0 \quad \text{and} \quad d^\pm(y) = r_1 + r_2 + r_3 \ \ \forall y \in V(G) \setminus V_0. \tag{11.1}$$

Similarly, in the original version $PCA^\diamond(r) \cup (\mathcal{EF}')^{\rm orig}$ of $PCA^\diamond(r) \cup \mathcal{EF}'$ we have

$$d^\pm(x) = 7r^\diamond \ \ \forall x \in V_0 \quad \text{and} \quad d^\pm(y) = 6r^\diamond \ \ \forall y \in V(G) \setminus V_0. \tag{11.2}$$

In order to construct $CA^\diamond(r)$ we will choose a chord absorber CA with $CA^{\rm exc} = \mathcal{EF}$. Similarly, in order to construct $PCA^\diamond(r)$ we will choose a parity extended cycle absorber $PCA(r^\diamond)$ such that the complete exceptional path systems contained in $PCA(r^\diamond)$ are those in $\mathcal{EF}'$.

Proof of Lemma 11.2. Choose new constants ε_1, d_1 such that $\varepsilon \ll \varepsilon_1 \ll 1/q \ll 1/f \ll r_1/m \ll d_1 \ll d$. Note that

$$r/m, r_2/m, qr_3/m \ll \varepsilon \quad \text{and} \quad r^\diamond \leq d_1 m. \tag{11.3}$$

We first apply Lemma 9.4 with $\varepsilon_1, r_1, r_2, r_3$ playing the roles of $\varepsilon', r_0, r_0', r_0''$ to find a chord absorber CA for C, U' with parameters $(\varepsilon_1, r_1, r_2, r_3, q, f)$ such that $CA^{\rm exc} = \mathcal{EF}$. Let $CA^\diamond(r) := CA \setminus CA^{\rm exc}$. Then (9.1) and the fact that $CA^{\rm exc}$ is r_3-regular imply that $CA^\diamond(r)$ satisfies (i).

Note that by definition of a setup, $(G, \mathcal{P}, R, C)$ is a (k, m, ε, d)-scheme. Let G_1 be obtained from G by deleting all edges in $CA^{\rm orig}$. Thus (9.1) implies that G_1 is obtained from G by deleting r_3q outedges and r_3q inedges at every vertex in V_0 and deleting $r_1+r_2+r_3$ outedges and $r_1+r_2+r_3$ inedges at every vertex in $V(G)\setminus V_0$. But $r_3q \leq \varepsilon m$ and $r_1+r_2+r_3 \leq d_1 m$ by (11.3). So Lemma 7.1 implies that $(G_1, \mathcal{P}, R, C)$ is still a $(k, m, 3\sqrt{d_1}, d)$-scheme. Since $r^\diamond/m \leq d_1 \ll d$ by (11.3), we can apply Lemma 10.5 to $(G_1, \mathcal{P}, R, C)$ to obtain an $r^\diamond$-fold parity extended cycle absorber $PCA(r^\diamond)$ such that the complete exceptional path systems contained in $PCA(r^\diamond)$ are precisely those in $\mathcal{EF}'$. Let $PCA^\diamond(r) := PCA(r^\diamond) \setminus \mathcal{EF}'$. Then (10.1) implies that $PCA^\diamond(r)$ satisfies (ii)(a).

To check (ii)(b), suppose that H is an r-factor of $G - V_0$ which is edge-disjoint from $G^{\rm rob} = CA^{\rm orig} \cup PCA(r^\diamond)^{\rm orig}$. Note that an $(\ell', k, m, \varepsilon, d)$-setup is also an $(\ell', k, m, \varepsilon_1, d)$-setup. So we can apply Lemma 9.7 to the $(\ell', k, m, \varepsilon, d)$-setup

$$(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$$

and the chord absorber $CA = \mathcal{B}(C)^* \cup \mathcal{B}(U') \cup CA^{\rm exc}$ with parameters $(\varepsilon_1, r_1, r_2, r_3, q, f)$ chosen before. This gives us a set $\mathcal{C}_1$ of rfk edge-disjoint Hamilton cycles in G such that the following conditions hold:

- Altogether the Hamilton cycles in $\mathcal{C}_1$ contain all the edges of $H \cup \mathcal{B}(U') \cup (CA^{\rm exc})^{\rm orig}$. Moreover, all remaining edges of these Hamilton cycles are contained in $\mathcal{B}(C)^*$.
- The digraph H_1 obtained from $CA^{\rm orig}$ by deleting all the edges lying on Hamilton cycles in $\mathcal{C}_1$ is a regular blow-up of C of degree $(r_1 + r_2 + r - (q-1)rfk/q) = r^\diamond$.
- The basic version of each cycle in $\mathcal{C}_1$ contains one of the s complete exceptional path systems contained in $CA^{\rm exc} = \mathcal{EF}$.

Finally, we apply Lemma 10.4 to the (k, m, ε, d)-scheme $(G, \mathcal{P}, R, C)$ with H_1 playing the role of H and with the $r^\diamond$-fold parity extended cycle absorber $PCA(r^\diamond)$ chosen before to find a Hamilton decomposition $\mathcal{C}_2$ of $H_1 \cup PCA(r^\diamond)^{\rm orig}$ such that each Hamilton cycle in $\mathcal{C}_2$ contains one of the $7r^\diamond$ complete exceptional path systems contained in $PCA(r^\diamond)$, and thus one of the complete exceptional path systems contained in $\mathcal{EF}'$. Then $\mathcal{C}_1 \cup \mathcal{C}_2$ is a Hamilton decomposition of $H \cup G^{\rm rob}$ as required in (ii)(b).

For the bipartite analogue, we apply the bipartite version of Lemmas 9.4 and 9.7 in the above argument. Since a $(\ell', k, m, \varepsilon, d)$-bi-setup is a (k, m, ε, d)-scheme, the remainder of the proof is the identical. $\square$

Finally, we can put together all of the previous results in order to prove Theorem 1.2. First we apply Szemerédi's regularity lemma to G. Based on the resulting partition, we find a consistent system and a universal walk (so that we have a corresponding setup). Within these, we find and remove a preprocessing graph PG, a chord absorber CA and a parity extended cycle absorber PCA. (Actually, instead of choosing CA and PCA separately, we choose suitable exceptional factors and apply Lemma 11.2 to find a robustly decomposable graph $G^{\rm rob}$ which is essentially the union of CA and PCA.) Then we apply Lemma 11.1 to cover all edges of $G[V_0]$. Next we find an approximate Hamilton decomposition using Theorem 1.3 in the remainder $G^{\#}$ of G, which leaves a sparse leftover H_0. We then find a Hamilton decomposition of $H_0 \cup PG^{\rm orig} \cup G^{\rm rob}$.

Proof of Theorem 1.2. Let $\tau^* := \tau(\alpha/2)$ be as defined in Theorem 1.3. Choose $\tau \le \tau^*$ such that $0 < \tau \ll \alpha$. Note that whenever $\nu' \le \nu$, every robust (ν,τ)-outexpander is also a robust (ν',τ)-outexpander. So we may assume that $0 \ll \nu \ll \tau$. Choose $n_0 \in \mathbb{N}$ so that $0 < 1/n_0 \ll \nu$. (τ and n_0 will be the constants returned by Theorem 1.2.) Now let G be an r-regular robust (ν,τ)-outexpander on $n \ge n_0$ vertices with $r \ge \alpha n$. We have to show that G has a Hamilton decomposition. Choose additional positive constants so that

$$\begin{aligned} &0 < 1/n_0 \ll \eta \ll d_0' \ll d_0'' \ll 1/k_2^* \ll 1/k_1^* \ll \varepsilon_0 \ll \varepsilon_1 \ll \varepsilon_2 \ll \varepsilon \ll \varepsilon' \ll \\ &\ll 1/q \ll 1/f \ll d_1 \ll d \ll \nu \ll \tau \ll \alpha < 1 \ \text{ and } \ d \ll 1/g \ll \zeta \le 1/2, \end{aligned} \tag{11.4}$$

and where $n_0, k_1^*, k_2^*, q, f, g \in \mathbb{N}$. Let

$$\ell' := \frac{64 \cdot 36}{\nu^6}, \quad s := \frac{64 \cdot 10^7}{\nu^6} \ \text{ and } \ \ell^* := f^2. \tag{11.5}$$

Note that we can choose $q, f \in \mathbb{N}$ and ν in such a way that $q/f, 50\ell^*/(s-1), \ell^*/7 \in \mathbb{N}$. As before, we can replace ν with a suitable $\nu' < \nu$ if necessary (and then prove the theorem for all (ν',τ)-outexpanders), so we can ensure that these divisibility conditions hold.

Apply Szemerédi's regularity lemma (Lemma 4.1) with parameters ε_0, d, k_1^* to obtain a partition $\mathcal{P}_0 = \{V_0', \dots, V_{k_0}'\}$ of the vertices of G into k_0 clusters and an exceptional set V_0', where $1/k_2^* \ll 1/k_0 \le 1/k_1^*$ and $|V_0'| \le \varepsilon_0 n$. Note that by adding at most $42g(g-1)f$ clusters to the exceptional set if necessary, we may assume that $k_0/14, k_0/f, k_0/g, 2fk_0/3g(g-1) \in \mathbb{N}$. Moreover, by moving at most ℓ^* vertices from each cluster V_i' to the exceptional set we may assume that the cluster size is divisible by ℓ^*. Note that we still have that $|V_0'| \le 2\varepsilon_0 n$. Let R_0 denote the corresponding reduced digraph. So every edge of R_0 corresponds to an $(2\varepsilon_0, \ge d)$-regular pair. Let

$$k := \ell^* k_0.$$

So

$$1/k_2^* \le 1/k \le 1/k_1^*$$

and $k/14, k/f, k/g, 2fk/3g(g-1), 50k/(s-1) \in \mathbb{N}$. By Lemma 5.1, R_0 is a robust $(\nu/2, 2\tau)$-outexpander with minimum semidegree at least $\alpha k_0/2$. So by Theorem 6.2, R_0 contains a Hamilton cycle C_0.

Apply Lemma 4.7 with C_0 playing the role of C to obtain an ε_0-uniform ℓ^*-refinement $\mathcal{P} = \{V_0', V_1 \dots, V_k\}$ of $\mathcal{P}_0$. Let R be the digraph obtained from R_0 by replacing every V_i' by the ℓ^* subclusters in $\mathcal{P}$ which are contained in V_i' and by replacing every edge $V_i'V_j'$ of R_0 by a complete bipartite graph K_{ℓ^*,ℓ^*} between the two corresponding sets of subclusters in $\mathcal{P}$, where all the edges of K_{ℓ^*,ℓ^*} are oriented towards those subclusters which are contained in V_j'. So R is an ℓ^*-fold blow-up of R_0 and $|R| = k$. By Lemma 4.7(ii) each edge of R corresponds to an $(\varepsilon_1, \geq d)$-regular pair in G. Moreover, Lemma 5.3 implies that R is still a robust $(\nu^3/8, 4\tau)$-outexpander with minimum semidegree at least $\alpha k/2$. Let C be a Hamilton cycle in R obtained from C_0 by winding ℓ^* times around C_0. By relabeling the clusters in $\mathcal{P}$ if necessary, we may assume that $C = V_1 \dots V_k$. Later on will use that this construction satisfies (CSys8) with $1/2$ playing the role of θ (since $\mathcal{P}$ is an ε_0-uniform ℓ^*-refinement of $\mathcal{P}_0$).

Apply Lemma 9.1 to R and C in order to obtain a universal walk U for C with parameter ℓ'. Let H be the spanning subgraph of R which consists of all the edges contained in $C \cup U$. Thus $\Delta(H) \leq 2(1+\ell') \leq 1/\nu^7$. Since each edge of R corresponds to an $(\varepsilon_1, \geq d)$-regular pair in G, Lemma 4.4 implies that we can move $\sqrt{\varepsilon_1}|V_i|$ vertices from each V_i to the exceptional set V_0' to achieve that every edge of H (and thus of $C \cup U$) corresponds to an $[\varepsilon_2/2, \geq d]$-superregular pair in G. We denote the modified exceptional set by V_0 and still denote the modified clusters by $V_1, \dots, V_k$. From now on, we will view C as a cycle on these clusters and U as a universal walk on these clusters. We also still write $\mathcal{P}$ for the partition of $V(G)$ into the exceptional set V_0 and clusters $V_1, \dots, V_k$. Let

$$m := |V_1| = \dots = |V_k|.$$

By adding at most $200\ell' q/f$ vertices from each cluster in $\mathcal{P}$ to V_0 if necessary, we may assume that $m/50, m/4\ell', fm/q \in \mathbb{N}$ (in particular, m is even). Note that

$$|V_0| \leq |V_0'| + \sqrt{\varepsilon_1} n + 200\ell' q/f \leq \varepsilon_2 n \tag{11.6}$$

and that every edge of R still corresponds to an $(\varepsilon_2, \geq d)$-regular pair in G and every edge of $C \cup U$ still corresponds to an $[\varepsilon_2, \geq d]$-regular pair in G. Let

$$r_0 := \eta r, \quad r_0' := d_0' m, \quad r_0'' := d_0'' m \quad \text{and} \quad r_1 := d_1 m. \tag{11.7}$$

(11.4) and the fact that $\eta \alpha n \leq r_0 = \eta r \leq \eta n$ together imply that

$$1/n_0 \ll r_0/m \ll r_0'/m \ll r_0''/m \ll 1/k. \tag{11.8}$$

Let G_0 denote the digraph obtained from G by deleting all the edges between vertices in V_0. Then

$$(G_0, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$$

is a consistent $(\ell^*, k, m, \varepsilon_2, d, \nu^3/8, 4\tau, \alpha/2, 1/2)$-system.

Apply Lemma 4.7 to obtain an ε_2-uniform ℓ'-refinement $\mathcal{P}'$ of $\mathcal{P}$. Let U' be the universal subcluster walk with respect to C, U and $\mathcal{P}'$. (So U' satisfies (ST2).) Then

$$(G_0, \mathcal{P}, \mathcal{P}', R, C, U, U')$$

is an $(\ell', k, m, \varepsilon, d)$-setup. Here we use that Lemma 4.7(i) implies that (ST3) is satisfied.

Apply Lemma 4.7 to obtain an η-uniform 50-refinement $\mathcal{P}''$ of $\mathcal{P}$. Apply Lemma 8.6 to the consistent system $(G_0, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ to find a preprocessing graph PG

in $G_0^{\rm basic}$ with parameters $(s,\varepsilon',d,r_0',r_0'',r_0,\zeta)$ with respect to C, R, $\mathcal{P}''$. (Here (11.4) and (11.8) imply that the conditions of the lemma are satisfied.)

Let G_1 be obtained from G_0 by deleting all edges in $PG^{\rm orig}$. Thus (8.3) implies that G_1 is obtained from G_0 by deleting $r_0(s-1)$ outedges and $r_0(s-1)$ inedges at every vertex in V_0 and by deleting r_0'' outedges and r_0'' inedges at every vertex in $V(G)\setminus V_0$. But $r_0(s-1), r_0''\le \varepsilon m$ by (11.4) and (11.8). So Lemma 7.1(i) implies that

$$(G_1,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$$

is still a consistent $(\ell^*,k,m,3\sqrt{\varepsilon},d,\nu^3/16,4\tau,\alpha/4,1/4)$-system. Furthermore, Lemma 9.2 implies that $(G_1,\mathcal{P},\mathcal{P}',R,C,U,U')$ is still an $(\ell',k,m,\varepsilon^{1/3},d)$-setup. Let

$$r^*:=r_0''-(s-1)r_0,\quad r_2:=96\ell' g^2kr^*\quad\text{and}\quad r_3:=\frac{r^*fk}{q}.$$

Note that

$$r^*k^2\le r_0''k^2\le r_0''(k_2^*)^2=d_0''m(k_2^*)^2\le d_1m=r_1 \tag{11.9}$$
$$\le m. \tag{11.10}$$

In particular, (11.9) implies that

$$r^*,r_0'',r_2,r_3\le r_1\overset{(11.7)}{=}d_1m. \tag{11.11}$$

So

$$\frac{r_2}{m}\overset{(11.11)}{\le} d_1\overset{(11.4)}{\ll} d\quad\text{and}\quad \frac{r_3}{m}\le\frac{r^*k}{m}\overset{(11.9)}{\le}\frac{r_0''k_2^*}{m}\overset{(11.7)}{=}d_0''k_2^*\overset{(11.4)}{\ll}\varepsilon. \tag{11.12}$$

Also, note that (11.5) implies that $f/\ell^*=1/f\ll 1$. Moreover, $r_3q/fm=r^*k/m\ll d$ by (11.12). Apply Lemma 4.7 to obtain an ε-uniform q/f-refinement $\mathcal{P}'''$ of $\mathcal{P}$. Apply Lemma 7.6 to $(G_1,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ in order to obtain exceptional factors $EF_1,\dots,EF_{r_3}$ with parameters $(q/f,f)$ with respect to C, $\mathcal{P}'''$ such that the original versions of all these exceptional factors are pairwise edge-disjoint. Let $\mathcal{EF}$ denote the union of the EF_i over all $i=1,\dots,r_3$. Let

$$r^\diamond:=r_1+r_2+r^*-(q-1)r_3.$$

Note that

$$r^\diamond/m\le(r_1+r_2+r^*)/m\overset{(11.11)}{\le}3d_1\overset{(11.4)}{\ll}d. \tag{11.13}$$

(11.10) guarantees that we can now apply Lemma 11.2 to the $(\ell',k,m,\varepsilon^{1/3},d)$-setup $(G_1,\mathcal{P},\mathcal{P}',R,C,U,U')$ with r^* playing the role of r and with the exceptional factors $EF_1,\dots,EF_{r_3}$ chosen before to find a spanning subdigraph $CA^\diamond(r^*)$.

Let G_2 be obtained from G_1 by deleting all edges in $CA^\diamond(r^*)\cup\mathcal{EF}^{\rm orig}$. Thus (11.1) implies that G_2 is obtained from G_1 by deleting r_3q outedges and r_3q inedges at every vertex in V_0 and by deleting $r_1+r_2+r_3$ outedges and $r_1+r_2+r_3$ inedges at every vertex in $V(G)\setminus V_0$. But $r_3q=r^*fk\le d_1m$ by (11.4) and (11.9) while $r_1+r_2+r_3\le 3d_1m$ by (11.11). So Lemma 7.1 implies that $(G_2,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ is still a consistent $(\ell^*,k,m,3\sqrt{3d_1},d,\nu^3/32,4\tau,\alpha/8,1/8)$-system. Together with (11.13) this shows that we can apply Lemma 7.6 to $(G_2,\mathcal{P}_0,R_0,C_0,\mathcal{P},R,C)$ in order to obtain exceptional factors $EF_1',\dots,EF_{r^\diamond}'$ with parameters $(1,7)$ with respect to C, $\mathcal{P}$ such that the original versions of all these exceptional factors are pairwise edge-disjoint. Let $\mathcal{EF}'$ denote the union of the EF'_i over all $i = 1, \dots, r^\diamond$.

Let $PCA^\diamond(r^*)$ be as guaranteed by Lemma 11.2(ii). Let

$$G^{\rm rob} := CA^\diamond(r^*) \cup PCA^\diamond(r^*) \cup (\mathcal{EF} \cup \mathcal{EF}')^{\rm orig} \quad \text{and} \quad G^{\rm absorb} := PG^{\rm orig} \cup G^{\rm rob}.$$

Let $r_0^{\rm abs}$ be the outdegree of the exceptional vertices (i.e. those in V_0) in $G^{\rm absorb}$. Then (8.3), (11.1) and (11.2) imply that $r_0^{\rm abs}$ is also the indegree of the exceptional vertices. Moreover, they imply that

$$r_0^{\rm abs} = r_0(s-1) + r_3 q + 7r^\diamond.$$

Let $r^{\rm abs}$ be the outdegree of the non-exceptional vertices in $G^{\rm absorb}$. Again, (8.3), (11.1) and (11.2) imply that $r^{\rm abs}$ is also the indegree of the non-exceptional vertices Moreover,

$$r^{\rm abs} = r''_0 + (r_1 + r_2 + r_3) + 6r^\diamond.$$

But

$$\begin{aligned} r_0^{\rm abs} - r^{\rm abs} &= r_0(s-1) + (q-1)r_3 + r^\diamond - r''_0 - r_1 - r_2 \\ &= r_0(s-1) + (q-1)r_3 + (r_1 + r_2 + r^* - (q-1)r_3) - r''_0 - r_1 - r_2 \\ &= r_0(s-1) + r^* - r''_0 = 0. \end{aligned}$$

So $G^{\rm absorb}$ is $r^{\rm abs}$-regular. Moreover,

$$\text{(11.14)} \quad r^{\rm abs} \le r''_0 + (r_1 + r_2 + r_3) + 6(r_1 + r_2 + r^*) \overset{(11.11)}{\le} 22d_1 m \overset{(11.4)}{\ll} dm \le dn.$$

Let $G^\triangle$ denote the digraph obtained from G by removing the edges of $G^{\rm absorb}$. Let $r^\triangle := r - r^{\rm abs}$ be the degree of $G^\triangle$. Note that (11.14) implies that $G^\triangle$ is still a robust $(\nu/2, \tau)$-outexpander with $\delta^0(G^\triangle) = r^\triangle \ge \alpha n/2$. Let $r^\# := r^\triangle - \varepsilon n$. Then (11.6) implies that we can apply Lemma 11.1 to obtain a set $\mathcal{C}^\triangle$ of εn edge-disjoint Hamilton cycles in $G^\triangle$ which cover all the edges in $G^\triangle[V_0]$. Let $G^\#$ be the $r^\#$-regular digraph obtained from $G^\triangle$ by deleting all the edges in these Hamilton cycles.

Note that $r - r^\# = r^{\rm abs} + \varepsilon n \le dn + \varepsilon n \le 2dn$ by (11.14). Thus $G^\#$ is still a robust $(\nu/2, \tau)$-outexpander (and thus also a robust $(\nu/2, \tau^*)$-outexpander) with $\delta^0(G^\#) = r^\# \ge \alpha n/2$. Together with our choice of τ^*, ν, η and n_0 this shows that we can apply Theorem 1.3 to obtain a set $\mathcal{C}^\#$ of $r^\# - \eta r = r^\# - r_0$ edge-disjoint Hamilton cycles in $G^\#$. Let H_0 be the digraph obtained from $G^\#$ by deleting all the edges in these Hamilton cycles. Note that our application of Lemma 11.1 ensures that V_0 forms an independent set in H_0 and so H_0 is a r_0-regular subdigraph of G_0. Together with (11.4) and (11.8) this ensures that we can apply Corollary 8.5 to $(G_0, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ with H_0, $\mathcal{P}''$ playing the roles of H, $\mathcal{P}'$ and with the preprocessing graph PG with parameters $(s, \varepsilon', d, r'_0, r''_0, r_0, \zeta)$ chosen before to obtain a set $\mathcal{C}_0$ of $r_0 s$ edge-disjoint Hamilton cycles such that the following conditions hold:

- Altogether the Hamilton cycles in $\mathcal{C}_0$ contain all edges of H_0 and each of these Hamilton cycles lies in $H_0 \cup PG^{\rm orig}$.
- Let PG' be the digraph obtained from $PG^{\rm orig}$ by deleting all the edges lying in the Hamilton cycles in $\mathcal{C}_0$. Then every vertex $x \in V_0$ is isolated in PG'

and every vertex $x \in V(G) \setminus V_0$ has in- and outdegree $r''_0 - (s-1)r_0 = r^*$ in PG'.

The second condition above implies that $H_1 := PG' - V_0$ is an r^*-regular subdigraph of $G_0 - V_0$. Lemma 11.2(ii)(b) guarantees that $H_1 \cup G^{\rm rob}$ has a Hamilton decomposition $\mathcal{C}_1$. Then $\mathcal{C}^{\triangle} \cup \mathcal{C}^{\#} \cup \mathcal{C}_0 \cup \mathcal{C}_1$ is a Hamilton decomposition of G, as required. $\square$

12. Statement of the robust decomposition lemma for further use

We now present a 'standalone' version (Lemma 12.1) of the robust decomposition lemma (Lemma 11.2) which is suitable for further use. Instead of exceptional edges and exceptional path systems, it involves 'fictive edges' and 'special path systems'. One can use these in the same way as in the current proof to deal with edges at exceptional vertices. A crucial additional advantage is that one can also use them to deal with a small number of edges connecting G to another digraph. In particular, in [32] we can use it for a digraph G^* which consists of robust expanders G and G' which are connected by a small number of edges (so G^* is not a robust expander). Similarly, in the bipartite version, we can apply it to an 'almost bipartite' digraph and use the fictive edges to deal e.g. with the small number of edges which do not respect the (approximate) bipartition. This is the case in [14].

Suppose that $(G, \mathcal{P}, R, C)$ is an (k, m, ε, d)-scheme with $C = V_1 \dots V_k$. The next definition is a generalization of a complete exceptional path system. Suppose that $k/L, m/K \in \mathbb{N}$ and let $\mathcal{I}$ be the canonical interval partition of C into L intervals of equal length. A *special path system* SPS (with respect to C) with parameters (K, L) spanning an interval $I = U_j U_{j+1} \dots U_{j'}$ with $I \in \mathcal{I}$ consists of m/K vertex-disjoint paths $P_1, \dots, P_{m/K}$ such that the following conditions hold.

(SPS1) Every P_s has its initial vertex in U_j and its final vertex in $U_{j'}$.

(SPS2) SPS contains a matching $\mathrm{Fict}(SPS)$ such that all the edges in $\mathrm{Fict}(SPS)$ avoid the endclusters U_j and $U_{j'}$ of I and such that $E(P_s) \setminus \mathrm{Fict}(SPS) \subseteq E(G)$.

(SPS3) SPS contains precisely m/K vertices from every cluster in I and no other vertices.

The edges in $\mathrm{Fict}(SPS)$ are called *fictive edges of* SPS. Note that a complete exceptional path system $CEPS$ containing a complete exceptional sequence CES is a special path system where CES plays the role of the set of fictive edges.

Suppose that $\mathcal{P}^*$ is a K-refinement of $\mathcal{P}$. For each cluster $U \in \mathcal{P}$, let $U(1), \dots, U(K)$ denote the subclusters of U in $\mathcal{P}^*$. Consider a special path system SPS as above. We say that SPS has *style* b if its vertex set is $U_j(b) \cup \dots \cup U_{j'}(b)$. A *special factor* SF *with parameters* (K, L) (with respect to C, $\mathcal{P}^*$) is a 1-regular digraph on $V(G) \setminus V_0$ satisfying the following properties:

(SF1) On each of the L intervals $I \in \mathcal{I}$, SF induces the vertex-disjoint union of K special path systems.

(SF2) Moreover, for each $I \in \mathcal{I}$ and each $b = 1, \dots, K$, exactly one of the special path systems in SF spanning I has style b.

We write $\mathrm{Fict}(SF)$ for the union of the sets $\mathrm{Fict}(SPS)$ over all the KL special path systems SPS contained in SF and call the edges in $\mathrm{Fict}(SF)$ *fictive edges of* SF.

Note that an exceptional factor EF is a special factor where the exceptional edges in EF play the role of the fictive edges.

We will always view fictive edges as being distinct from each other and from the edges in other digraphs. So if we say that $SF_1, \dots, SF_r$ are pairwise edge-disjoint from each other and from some digraph Q on $V(G) \setminus V_0$, then this means that Q and all the $SF_i \setminus \mathrm{Fict}(SF_i)$ are pairwise edge-disjoint, but for example there could be an edge from x to y in Q as well as in $\mathrm{Fict}(SF_i)$ for five indices i (say). But these are the only instances of multiedges that we allow, i.e. if there is more than one edge from x to y, then all but at most one of these edges are fictive edges.

Given two multidigraphs M and M' on the same vertex set, we write $M + M'$ for the multidigraph whose vertex set is $V(M) = V(M')$ and in which the multiplicity of xy is the sum of the multiplicities of xy in M and in M' (for all $x, y \in V(M)$). So in the above example $Q + SF_1 + \dots + SF_r$ contains six edges from x to y.

We can now state the variant of the robust decomposition lemma. The proof is the same as that of Lemma 11.2 – the special factors play the role of the exceptional factors and the Hamilton cycles in Lemma 12.1 correspond to the basic versions of the Hamilton cycles returned by Lemma 11.2. The existence of fictive edges means that we formally consider multidigraphs (rather than digraphs) at several steps. However, this does not affect the argument. Indeed, fictive edges only occur within (pre-defined) special path systems and these are fixed building blocks that are never modified during the construction of the Hamilton cycles. The only other difference is that in Lemma 12.1, H need not be a subdigraph of G, but this does not affect the proof either.

Lemma 12.1. *Suppose that* $0 < 1/n \ll 1/k \ll \varepsilon \ll 1/q \ll 1/f \ll r_1/m \ll d \ll 1/\ell', 1/g \ll 1$ *and that* $rk^2 \le m$. *Let*

$$r_2 := 96\ell' g^2 kr, \quad r_3 := rfk/q, \quad r^\diamond := r_1 + r_2 + r - (q-1)r_3, \quad s' := rfk + 7r^\diamond$$

and suppose that $k/14, k/f, k/g, q/f, m/4\ell', fm/q, 2fk/3g(g-1) \in \mathbb{N}$. *Suppose that* $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ *is an* $(\ell', k, m, \varepsilon, d)$*-setup with* $|G| = n$, $V_0 = \emptyset$ *and* $C = V_1 \dots V_k$. *Suppose that* $\mathcal{P}^*$ *is a* (q/f)*-refinement of* $\mathcal{P}$ *and that* $SF_1, \dots, SF_{r_3}$ *are edge-disjoint special factors with parameters* $(q/f, f)$ *with respect to* C, $\mathcal{P}^*$. *Let* $\mathcal{SF} := SF_1 + \dots + SF_{r_3}$. *Then there exists a spanning subdigraph* $CA^\diamond(r)$ *of* G *for which the following holds:*

(i) $CA^\diamond(r)$ *is an* (r_1+r_2)*-regular spanning subdigraph of* G *which is edge-disjoint from* $\mathcal{SF}$.

(ii) *Suppose that* $SF'_1, \dots, SF'_{r^\diamond}$ *are special factors with parameters* $(1, 7)$ *with respect to* C, $\mathcal{P}$ *which are edge-disjoint from each other and from* $CA^\diamond(r) + \mathcal{SF}$. *Let* $\mathcal{SF}' := SF'_1 + \dots + SF'_{r^\diamond}$. *Then there exists a spanning subdigraph* $PCA^\diamond(r)$ *of* G *for which the following holds:*
   (a) $PCA^\diamond(r)$ *is a* $5r^\diamond$*-regular spanning subdigraph of* G *which is edge-disjoint from* $CA^\diamond(r) + \mathcal{SF} + \mathcal{SF}'$.
   (b) *Let* $\mathcal{SPS}$ *be the set consisting of all the* s' *special path systems contained in* $\mathcal{SF} + \mathcal{SF}'$. *Whenever* H *is an* r*-regular digraph on* $V(G)$ *which is edge-disjoint from* $G^{\mathrm{rob}} := CA^\diamond(r) + PCA^\diamond(r) + \mathcal{SF} + \mathcal{SF}'$, *then* $H + G^{\mathrm{rob}}$ *has a decomposition into* s' *edge-disjoint Hamilton cycles* $C_1, \dots, C_{s'}$.

Moreover, C_i contains one of the special path systems from $\mathcal{SPS}$, for each $i = 1, \dots, s'$.

The analogue holds for an $(\ell', k, m, \varepsilon, d)$-bi-setup $(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ if we assume in addition that H is bipartite with vertex classes $\bigcup \mathcal{V}_{\rm even}$ and $\bigcup \mathcal{V}_{\rm odd}$ (where $\mathcal{V}_{\rm even}$ is the set of all those V_i such that i is even and $\mathcal{V}_{\rm odd}$ is defined analogously).

13. Proofs of Theorems 1.1 and 1.4

The next result implies that a regular oriented graph with minimum semidegree at little larger than $3n/8$ is a robust outexpander. Together with Theorem 1.2 this implies Theorem 1.1.

Lemma 13.1. *Let $0 < 1/n \ll \nu \ll \tau \le \varepsilon/2 \le 1$ and suppose that G is an oriented graph on n vertices with $\delta^+(G) + \delta^-(G) + \delta(G) \ge 3n/2 + \varepsilon n$. Then G is a robust (ν, τ)-outexpander.*

Proof. Suppose not and let $X \subseteq V(G)$ be a set of vertices such that $\tau n \le |X| \le (1-\tau)n$ and $|RN^+_\nu(X)| < |X| + \nu n$. Let $A := X \cap RN^+_\nu(X)$, $B := RN^+_\nu(X) \setminus X$, $D := X \setminus RN^+_\nu(X)$ and $C := V(G) \setminus (A \cup B \cup D)$. Note that

$$|B| < |D| + \nu n. \tag{13.1}$$

Claim 1. $|A| + |B| + |D| \ge 2\delta^+(G) - 2\tau n$

To prove the claim, let us first assume that $|A| \ge \tau n/2$. Note that $e(A, C \cup D) \le \nu n(|C| + |D|)$ since every vertex in $C \cup D$ has at most νn inneighbours in $X \supseteq A$. Thus

$$\begin{aligned} e(A, A \cup C \cup D) = e(A, A) + e(A, C \cup D) &\le \frac{|A|^2}{2} + \nu n(|C| + |D|) \\ &\le \frac{|A|^2}{2} + \frac{2\nu}{\tau}|A|(|C| + |D|) \le \frac{|A|^2}{2} + |A|\frac{\tau n}{2}. \end{aligned}$$

So there exists a vertex $x \in A$ such that $|N^+(x) \cap (A \cup C \cup D)| \le |A|/2 + \tau n/2$. Hence $\delta^+(G) \le d^+(x) \le |A|/2 + |B| + \tau n/2$. Together with (13.1) this implies Claim 1 in this case.

So let us next assume that $|A| \le \tau n/2$. Then $|D| \ge \tau n/2$ and

$$\begin{aligned} e(D, A \cup C \cup D) = e(D, A) + e(D, C \cup D) &\le |D|\frac{\tau n}{2} + \nu n(|C| + |D|) \\ &\le |D|\frac{\tau n}{2} + \frac{2\nu}{\tau}|D|(|C| + |D|) \le |D|\frac{3\tau n}{4}. \end{aligned}$$

So there exists a vertex $x \in D$ such that $|N^+(x) \cap (A \cup C \cup D)| \le 3\tau n/4$. Hence $\delta^+(G) \le d^+(x) \le |B| + 3\tau n/4$. As before, together with (13.1) this implies Claim 1.

Claim 2. $|B| + |C| + |D| \ge 2\delta^-(G) - 3\nu n$

To prove Claim 2, we first consider the case when $C \ne \emptyset$. An averaging argument shows that there is a vertex $x \in C$ with $|N^-(x) \cap C| \le |C|/2$. But since $|N^-(x) \cap X| \le \nu n$ this means that $\delta^-(G) \le d^-(x) \le |B| + |C|/2 + \nu n$. Together with (13.1) this implies Claim 2.

So let us now assume that $C = \emptyset$. Together with the fact that $|A \cup B| = |RN^+_\nu(X)| < |X| + \nu n \le (1-\tau)n + \nu n < n$ this implies that $D \ne \emptyset$. But each vertex $x \in D$ satisfies $|N^-(x) \cap X| \leq \nu n$ and so $\delta^-(G) \leq d^-(x) \leq \nu n + |B|$. Together with (13.1) this implies Claim 2.

Claim 3. $|A| + |B| + |C| \geq \delta(G) - 2\nu n$

This clearly holds if $D = \emptyset$ (since $\delta(G) < n$). So suppose that $D \neq \emptyset$. Then $e(D, D) \leq e(X, D) \leq \nu n|D|$ and so there is a vertex $x \in D$ with $|N^+(x) \cap D| \leq \nu n$. But since $D \cap RN^+_\nu(X) = \emptyset$ we also have that $|N^-(x) \cap D| \leq \nu n$. Thus $d(x) \leq |A| + |B| + |C| + 2\nu n$, which in turn implies Claim 3.

Now Claims 1–3 together imply that

$$3n \overset{(13.1)}{\geq} 3|A| + 4|B| + 3|C| + 2|D| - \nu n \geq 2(\delta^+(G) + \delta^-(G) + \delta(G)) - 3\tau n > 3n,$$

a contradiction. □

Proof of Theorem 1.1. Let $\tau^* := \tau(3/8)$, where $\tau(3/8)$ is as defined in Theorem 1.2. Choose new constants $n_0 \in \mathbb{N}$ and ν, τ such that $0 < 1/n_0 \ll \nu \ll \tau \leq \varepsilon/2, \tau^*$. Lemma 13.1 implies that G is a robust (ν, τ)-outexpander and thus also a robust (ν, τ^*)-outexpander. So we can apply Theorem 1.2 with $\alpha := 3/8$ to find a Hamilton decomposition of G. □

The next result implies that a regular digraph with minimum semidegree at little larger than $n/2$ is a robust outexpander. Similarly as before, together with Theorem 1.2 this implies Theorem 1.4.

Lemma 13.2. *Suppose that $0 < \nu \leq \tau \leq \varepsilon < 1$ are such that $\varepsilon \geq 2\nu/\tau$. Let G be a digraph on n vertices with minimum semidegree $\delta^0(G) \geq (1/2 + \varepsilon)n$. Then G is a robust (ν, τ)-outexpander.*

Proof. Consider any set $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1-\tau)n$. Let $RN := RN^+_{\nu,G}(S)$. We have to show that $|RN| \geq |S| + \nu n$. Suppose first that $|S| \geq n/2$. Then every vertex of G has at least $\varepsilon n \geq \nu n$ inneighbours in S. So $RN = V(G)$. So we may assume that $|S| \leq n/2$. But

$$\begin{aligned}(1/2 + \varepsilon)n|S| \leq e(S, V(G)) = e(S, RN) + e(S, V(G) \setminus RN) &\leq |S||RN| + \nu n^2 \\ &\leq |S||RN| + \frac{\nu}{\tau} n|S|\end{aligned}$$

and so $|RN| \geq (1/2 + \varepsilon - \nu/\tau)n \geq (1 + \varepsilon)n/2 \geq |S| + \nu n$, as required. □

Proof of Theorem 1.4. Let $\tau^* := \tau(1/2)$, where $\tau(1/2)$ is as defined in Theorem 1.2. Choose new constants $n_0 \in \mathbb{N}$ and ν, τ such that $0 < 1/n_0 \ll \nu \ll \tau \leq \varepsilon, \tau^*$. Lemma 13.2 implies that G is a robust (ν, τ)-outexpander and thus also a robust (ν, τ^*)-outexpander. So we can apply Theorem 1.2 with $\alpha := 1/2$ to find a Hamilton decomposition of G. □

14. Acknowledgements

We would like to thank John Lapinskas for an idea which led to a simplification of the cycle absorbing argument. We are extremely grateful to Allan Lo for his detailed comments on a draft of this paper.

References

[1] N. Alon, R.A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the regularity lemma, *Journal of Algorithms* **16** (1994), 80–109.
[2] N. Alon, G. Gutin and M. Krivelevich, Algorithms with large domination ratio, *Journal of Algorithms* **50** (2004), 118–131.
[3] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *Journal of Computer and System Sciences* **69** (2004), 354–382.
[4] B. Alspach, The wonderful Walecki construction, *Bull. Inst. Combin. Appl.* **52** (2008), 7–20.
[5] B. Alspach, J.C. Bermond and D. Sotteau, Decompositions into cycles. I. Hamilton decompositions, *Cycles and rays (Montreal, PQ, 1987)*, Kluwer Acad. Publ., Dordrecht, 1990, 9–18.
[6] B. Alspach, D. Bryant and D. Dyer, Paley graphs have Hamilton decompositions, *Discrete Mathematics* **312** (2012), 113–118.
[7] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer 2000.
[8] J.C. Bermond and C. Thomassen, Cycles in digraphs – a survey, *J. Graph Theory* **5** (1981), 1–43.
[9] B. Bollobás and A. Frieze, On matchings and Hamiltonian cycles in random graphs. *Random graphs '83* (Poznan, 1983), North-Holland Math. Stud., 118, North-Holland, Amsterdam (1985), 23–46.
[10] A. Bondy, Basic graph theory: paths and circuits, in *Handbook of Combinatorics*, Vol. 1, Elsevier, Amsterdam (1995), 3–110.
[11] D. Christofides, P. Keevash, D. Kühn and D. Osthus, A semi-exact degree condition for Hamilton cycles in digraphs, *SIAM Journal Discrete Mathematics* **24** (2010), 709–756.
[12] D. Christofides, P. Keevash, D. Kühn and D. Osthus, Finding Hamilton cycles in robustly expanding digraphs, *Journal of Graph Algorithms and Applications* **16** (2012), 337–360.
[13] D. Christofides, D. Kühn and D. Osthus, Edge-disjoint Hamilton cycles in graphs, *J. Combin. Theory B* **102** (2012), 1035–1060.
[14] B. Csaba, D. Kühn, A. Lo, D. Osthus and A. Treglown, Proof of the 1-factorization and Hamilton decomposition conjectures II: the bipartite case, preprint.
[15] B. Csaba, D. Kühn, A. Lo, D. Osthus and A. Treglown, Proof of the 1-factorization and Hamilton decomposition conjectures III: approximate decompositions, preprint.
[16] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, *J. Combin. Theory B* **94** (2005), 159–172.
[17] A. Frieze and M. Krivelevich, On two Hamilton cycle problems in random graphs, *Israel Journal of Mathematics* **166** (2008), 221–234.
[18] 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.
[19] G. Gutin and A. Yeo, TSP tour domination and Hamilton cycle decompositions of regular graphs, *Oper. Res. Letters* **28** (2001), 107–111.
[20] R. Häggkvist, Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **2** (1993), 25–32.
[21] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in oriented graphs, in *Combinatorics, Geometry and Probability*, Cambridge University Press 1997, 339–353.
[22] B. Jackson, Long paths and cycles in oriented graphs, *J. Graph Theory* **5** (1981), 145–157.
[23] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.
[24] 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.

[25] L. Kelly, D. Kühn and D. Osthus, A Dirac-type result on Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **17** (2008), 689–709.
[26] F. Knox, D. Kühn and D. Osthus, Edge-disjoint Hamilton cycles in random graphs, preprint.
[27] J. Komlós, G.N. Sárközy and E. Szemerédi, Blow-up Lemma, *Combinatorica* **17** (1997), 109–123.
[28] J. Komlós, G.N. Sarközy and E. Szemerédi, An algorithmic version of the blow-up lemma, *Random Structures Algorithms* **12** (1998), 297–312.
[29] J. Kómlos, G.N. Sárközy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Ann. Combin* **2** (1998), 43–60.
[30] M. Krivelevich and W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, *SIAM J. Discrete Mathematics* **26** (2012), 964–982.
[31] D. Kühn, J. Lapinskas and D. Osthus, Optimal packings of Hamilton cycles in graphs of high minimum degree, *Combinatorics, Probability and Computing*, to appear.
[32] D. Kühn, A. Lo, D. Osthus and A. Treglown, Proof of the 1-factorization and Hamilton decomposition conjectures I: the two cliques case, preprint.
[33] D. Kühn, A. Lo, and D. Osthus, Proof of the 1-factorization and Hamilton decomposition conjectures IV: exceptional systems for the two cliques case, preprint.
[34] 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.
[35] D. Kühn and D. Osthus, Hamilton decompositions of regular expanders: applications, preprint.
[36] D. Kühn and D. Osthus, A survey on Hamilton cycles in directed graphs, *European J. Combinatorics* **33** (2012), 750–766.
[37] D. Kühn, D. Osthus and A. Treglown, Hamiltonian degree sequences in digraphs, *J. Combin. Theory B* **100** (2010), 367–380.
[38] D. Kühn, D. Osthus and A. Treglown, Hamilton decompositions of regular tournaments, *Proc. London Math. Soc.* **101** (2010), 303–335.
[39] E. Lucas, *Récréations Mathématiques*, Vol. 2, Gautheir-Villars, 1892.
[40] J.W. Moon, *Topics on tournaments*, Holt, Rinehart and Winston, New York, 1968.
[41] C. St. J. A. Nash-Williams, Hamiltonian arcs and circuits, in *Recent Trends in Graph Theory*, Springer 1971, 197–210.
[42] D. Osthus and K. Staden, Approximate Hamilton decompositions of regular robustly expanding digraphs, preprint.
[43] L. Perkovic and B. Reed, Edge coloring regular graphs of high degree, *Discrete Mathematics* **165/166** (1997), 567–579.
[44] A. Strivastav and P. Stangier, Algorithmic Chernoff-Hoeffding inequalities in integer programming, *Random Structures Algorithms* **8** (1996), 27–58.
[45] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in Combinatorics* (B. Bollobás ed.), *London Math. Soc. Lecture Notes* **38** (1979), 211–228.
[46] C. Thomassen, Edge-disjoint Hamiltonian paths and cycles in tournaments, *Proc. London Math. Soc.* **45** (1982), 151–168.
[47] T.W. Tillson, A Hamiltonian decomposition of K^*_{2m}, $2m \geq 8$, *J. Combin. Theory B* **29** (1980), 68–74.
[48] C.Q. Zhang, Every regular tournament has two arc-disjoint Hamilton cycles, *J. Qufu Normal College, Special Issue Oper. Research* (1980), 70–81.

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

E-mail addresses: {d.kuhn,d.osthus}@bham.ac.uk