

ROBUST EXPANSION AND HAMILTON DECOMPOSITIONS

ABSTRACT. We investigate the recently coined notion of robust expansion and its strong ties to Hamilton decompositions. In particular, we prove and sketch aspects of the result that any sufficiently large regular robustly expanding digraph has a Hamilton decomposition. The long-standing conjecture of Kelly (1968), that any regular tournament on n vertices can be decomposed into $(n - 1)/2$ edge-disjoint Hamilton cycles, then follows as a simple corollary.

In this essay I have included some original work, proving that asymptotically almost every graph is a robust expander (with certain parameters). Furthermore, I have found an error in one of the most important lemmas in the essay and provided two fixes (one for the statement and one for the proof).

1. PRELIMINARIES

1.1. Introduction. Informally, we think of the expansion property as meaning that for a set that is ‘not too large’ its neighbourhood or boundary is ‘large’, suggesting strong connectivity properties. These ‘regular’ or ‘standard’ expanders have been researched heavily, and their applications reach far beyond combinatorics: they have extensive uses in theoretical computer science and error-correcting codes (see [5] for examples).

There are several notions of expansion that are prevalent in the literature. We may like to call an undirected graph G an expander if for all $S \subset V(G)$ with $0 < |S| < |G|/2$, we have $|N(S)| \geq \lambda|S|$ for some $\lambda > 0$, or we may ask that $e(S, S^c) \geq \lambda|S|$. We could even slightly adapt this idea and instead look for sets which are ‘not too small nor too large’ which have neighbourhoods that are ‘at least a bit larger than them’, i.e. for all $S \subset V(G)$ with $\tau|G| \leq |S| \leq (1 - \tau)|G|$ we have $|N(S)| \geq |S| + \nu|G|$.

However, problems arise when one attempts to find a Hamilton cycle. Note that for the first definition to hold, then we need $\lambda \leq 2$, so if we add an extra vertex to our graph G and join it to two ‘poorly connected’ vertices, then it will remain an expander (possibly with a slightly worse parameter $\tilde{\lambda} \geq \lambda/2$), but it is extremely likely that we will not be able to find a Hamilton cycle here. We can imagine that the problem is that we are not requiring the vertices in the neighbourhood to have any properties beyond simply being in the neighbourhood. What ‘good’ is a vertex $y \in N(S)$ for S , if it only has one neighbour in there? We require much stronger ‘connectivity’ properties if we hope to find a general result about finding Hamilton cycles due to the expansion property. This is where we introduce the concept of a *robust expander*, where instead of considering the standard neighbourhood, we opt for the *robust neighbourhood* which is the set of vertices with at least νn neighbours in S - it is highly connected to S . Requiring this to be of size at least $|S| + \nu n$ is a much stronger statement. With this new concept, we can easily prove the existence of a Hamilton cycle, and we can even go as far as finding a Hamilton decomposition of the whole graph.

This then begs the question of whether this notion is too strong and if we are restricting ourselves too much here. I show, through my own original work, that, as we let n grow, almost every graph on n vertices indeed is a robust expander.

The concept of robust expansion can be attributed to Daniela Kühn and Deryk Osthus, who developed this in order to tackle open problems in Hamilton cycles. They had tremendous success with this, settling several long-standing conjectures, including that of Kelly (1968) and Sumner’s Universal Tournament Conjecture (1971).

In this essay, our main goal is to develop robust expansion to the point that it can be used to prove that any sufficiently large regular robust outexpander has a Hamilton decomposition, which trivially implies Kelly's conjecture. We begin by extracting the important result from [1] (the paper in which robust expansion was introduced) that robust expanders have a Hamilton cycle. This result is omnipresent in the aforementioned proof, being utilised directly or via another lemma which relies deeply on it. Afterwards, we dedicate a section to the conversion of 1-factors (a vertex-disjoint union of paths) into full Hamilton cycles by borrowing edges from superregular graphs using robust expansion properties. This, and the result above, have their own sections due to the fact that these results can easily be applied elsewhere, and are not merely tools for proving one theorem. The rest of the essay is concerned with sketching the proof of Kelly's conjecture. This will be spread over several sections due to its length ([2, 3] sum to about 140 pages) and we will aim to extract the key ideas - above all those that include robust expansion.

1.2. Main results. Before giving statements of the key theorems, we will first precisely define the concept of a robustly expanding digraph; a note on notation will be given afterwards.

Definition 1.1. Given $0 < \nu < 1$, a digraph G and a set $S \subset V(G)$ the ν -robust outneighbourhood of S is defined to be

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

Definition 1.2. Given $0 < \nu \leq \tau < 1$, we call a digraph G an *robust (ν, τ) -outexpander* if $|RN_{\nu, G}^+(S)| \geq |S| + \nu|G|$ for all $S \subset V(G)$ with $\tau|G| \leq |S| \leq (1 - \tau)|G|$.

The main result of this essay was proved by Kühn and Osthus in [2] and states that every sufficiently large regular robustly expanding digraph has a full Hamilton decomposition; the precise statement is given below.

Theorem 1.3. 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

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

Then G has a Hamilton decomposition.

In order to prove this, we shall rely on an approximate version of the above that was proved by Kühn, Osthus and Staden in [3], which formed part of Staden's PhD thesis.

Theorem 1.4. 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

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

Then G contains at least $(1 - \eta)r$ edge-disjoint Hamilton cycles.

The above two theorems will be looked at in some depth, however due to their length we will not be able to go too deep. In light of this, we will isolate the robust expansion and sketch the rest of the argument, picking out the key, interesting elements.

We will then prove the following corollary of Theorem 1.3, which is given in much more familiar terms. This will be the subject of Section 8 at the end of the essay, and has little substance to it.

Theorem 1.5. For every $\epsilon > 0$, there exists n_0 such that every r -regular oriented graph G on $n \geq n_0$ vertices with $r \geq 3n/8 + \epsilon n$ has a Hamilton decomposition.

From this, it is trivial to deduce Kelly's long-standing conjecture from 1968.

Conjecture 1.6. (Kelly) Every regular tournament on n vertices can be decomposed into $(n - 1)/2$ edge-disjoint Hamilton cycles.

This paper is organised as follows. The remainder of this section will be spent introducing the concept of robust expansion. After, I will then prove, through my own original work, that asymptotically almost every graph is a robust expander. In Section 3 we will show that any robustly expanding digraph has a Hamilton cycle; this is a fundamental tool for the remainder of the paper. An equally important tool will be developed in Section 4, in which we will show that under certain conditions we can easily turn 1-factors into Hamilton cycles. In this section we also point out the error that was committed in [2] and two fixes, which were confirmed by Kühn. The two Sections that follow will include sketches of Theorems 1.4 and 1.3, respectively, highlighting the use of robust expanders in each proof. Section 8 will show that every large tournament satisfies the hypotheses of Theorem 1.3, thus settling the long-standing conjecture from 1968.

1.3. Notation. Most of the notation in this essay will be standard, identical to what was used in Part II Graph Theory. Given that the course did not stretch to directed graphs (henceforth abbreviated to digraphs), we will explain this notation; it is a mixture of what is found from Kühn's papers and Staden's PhD thesis.

If $x \in V(G)$ where G is a digraph, we let the *inneighbourhood* and *outneighbourhood*, respectively, of x be denoted by $N_G^-(x)$ and $N_G^+(x)$, respectively, i.e. the $y \in V(G)$ for which $yx \in E(G)$ and $xy \in E(G)$, respectively. Note that elsewhere, if it is clear from the context, we will often omit the subscript G . We write $d_G^+(x) := |N_G^+(x)|$ for the *outdegree* of x and $d_G^-(x) := |N_G^-(x)|$ for the *indegree*. We denote the *minimum outdegree* by $\delta^+(G) := \min_{x \in V(G)} d_G^+(x)$ and the *minimum indegree* by $\delta^-(G) := \min_{x \in V(G)} d_G^-(x)$, and the notion of degree we are most interested in is the *minimum semidegree* $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$.

If $X, Y \subset V(G)$, we let $e_G(X, Y)$ be the number of edges which have their initial vertex in X and their final in Y . In the case that $X \cap Y = \emptyset$, we define the subdigraph $G[X, Y]$ to be the bipartite graph with all edges of G directed from X to Y . In light of this, we can often view it as undirected (since all edges go from X to Y); indeed, many quoted results in [2] are for undirected bipartite graphs.

An r -factor of a digraph G is a subdigraph such that all vertices have out- and indegree 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 $r \otimes R$ obtained by replacing each vertex x of R by r vertices and replacing each edge xy of R by the oriented complete bipartite graph $K_{r,r}$. A slightly looser notion of a blowing up is the following: if R is a graph on clusters V_1, \dots, V_k , then a *blow-up* $\mathcal{B}(R)$ of R is obtained by replacing each vertex V_i of R with the vertices in it and each edge $V_i V_j$ with a certain bipartite graph with all edges from V_i to V_j . If R is a directed cycle, e.g. $R = C = V_1 \dots V_k$, and G is a digraph with $V(G) \subset V_1 \cup \dots \cup V_k$, then we say that (*the edges of*) G *wind(s) around* C if for each edge $xy \in G$, there is an i such that $x \in V_i$ and $y \in V_{i+1}$.

In sketches and motivating paragraphs, we will often write G and R without explaining what they mean. G will always refer to the original digraph we are given, and R will always be its reduced digraph. This is simply to remove clutter and repetition, when it is clear from the context to what we are referring.

When choosing parameters we will adopt the notation of Kühn and Osthus: $a \ll b \ll c$, which essentially means that we can choose the parameters from left to right. Precisely, there exist non-decreasing functions $f_1, f_2 : (0, 1] \rightarrow (0, 1]$ such that the result holds if $b \leq f_1(c)$ and $a \leq f_2(b)$. In order to avoid clutter, we will only find *examples* of such functions in Section 3 and the odd proof later on in the paper. (It must be noted that the functions we find will be specific to our argument and may be different than the ones Kühn et al. had in mind; furthermore, our bounds may be somewhat crude in comparison.) The reason for not calculating these all the time is that

finding these relations often detracts from the ‘heart’ of the argument, so it is often not desirable to see the explicit calculation, but rather better to be aware of the fact that such parameters do exist. However, since this is an essay focussed on robust expansion, the details may be included if the constants are specific to this, but if they are about, say, regularity, then it is unlikely that I will include the parameter calculations.

1.4. Definitions. There are a few notions of regularity that we are interested in.

Definition 1.7. Let $G = (A, B)$ be an undirected bipartite graph and $\epsilon > 0$. We say that G is ϵ -regular if for every $X \subset A$ and $Y \subset B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ we have that $|d_G(X, Y) - d_G(A, B)| < \epsilon$.

The following two definitions will only be used in the sketches of Theorems 1.4 and Theorem 1.3. They are included here for completeness, but one should not spend too long thinking about exactly what they mean.

Definition 1.8. Let G be a digraph and \mathcal{P} be a partition of $V(G)$ into an exceptional set V_0 and clusters of equal size. Supposing \mathcal{P}' is another partition 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 partition each cluster in \mathcal{P} into l subclusters of equal size.

Definition 1.9. \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. Whenever x is a vertex of G , V is a cluster in \mathcal{P} and $|N_G^+(x) \cap V| \geq \epsilon|V|$ then $|N_G^+(x) \cap V'| = (1 \pm \epsilon)|N_G^+(x) \cap V|/l$ for each cluster $V' \in \mathcal{P}'$ with $V' \subset V$; the inneighbourhoods satisfy the analogous condition.

We can informally think of this as meaning that our refinement essentially distributes the edges evenly so that the subclusters have roughly the same number of edges to each vertex. A proof for the existence of such uniform refinements can be found in [2].

1.5. Probability estimates. We recall the following Chernoff-type bounds for a random variable $X \sim \text{Bin}(n, p)$.

Theorem A.1.12 from [6] states (after a slight modification) that for $\beta > 1$, we have that:

$$\mathbb{P}(X \geq \beta \mathbb{E}[X]) \leq e^{(\beta-1-\beta \log \beta) \mathbb{E}[X]}. \quad (1.1)$$

[6] also gives us (Corollary A.1.14) that for any $\epsilon > 0$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \epsilon \mathbb{E}[X]) \leq 2e^{-c_\epsilon \mathbb{E}[X]}, \quad (1.2)$$

where $c_\epsilon > 0$ depends only on ϵ .

From [7] (Corollary 2.3), we have the following estimate for any $\delta > 0$:

$$\mathbb{P}(X \geq (1 + \delta) \mathbb{E}[X]) \leq e^{-\delta^2 \mathbb{E}[X]/3}. \quad (1.3)$$

.

On the other hand, we have the following estimate from [8] for $\delta \in (0, 1)$:

$$\mathbb{P}(X \leq (1 - \delta) \mathbb{E}[X]) \leq e^{-\delta^2 \mathbb{E}[X]/2}, \quad (1.4)$$

which we can actually weaken to find get

$$\mathbb{P}(X \leq np - nh) \leq e^{-h^2 n/2}. \quad (1.5)$$

1.6. Szemerdi's regularity lemma. All of the work of this essay is reliant on the following lemma, henceforth abbreviated to the (di)regularity lemma or just Szemerédi's lemma. This provides us with a partition of G into vertex sets (called clusters) of equal size (with a small exceptional set) such that the bipartite graph restricted to two clusters either is sufficiently dense or has no edges. This is a *very* useful partition since we will go on to show that the *reduced digraph*, whose vertex set and has edges if the corresponding bipartite graph is dense, inherits this robust expansion property. At the end of Section 3 we will prove that any robust expander has a Hamilton cycle, so we will thus know that any reduced digraph also has a Hamilton cycle, which one can imagine is a very good starting point for trying to find Hamilton cycles of G .

Lemma 1.10. (Degree form of the diregularity lemma.) For every $\epsilon \in (0, 1)$ and every integer M' there are integers M and n_0 such that if G is a digraph on $n \geq n_0$ vertices and $d \in [0, 1]$, then there is a partition of the vertex set of G into V_0, V_1, \dots, V_k and a spanning subdigraph G' of G such that the following holds:

- $M' \leq k \leq M$;
- $|V_0| \leq \epsilon n$;
- $|V_1| = \dots = |V_k| =: m$;
- $d_{G'}^+ > d_G^+ - (d + \epsilon)n$ for all vertices $x \in V(G)$;
- $d_{G'}^- > d_G^- - (d + \epsilon)n$ for all vertices $x \in V(G)$;
- for all $i = 1, \dots, k$ the digraph $G'[V_i]$ is empty; and
- for all $1 \leq i, j \leq k$ with $i \neq j$ the pair $(V_i, V_j)_{G'}$ is ϵ -regular and has density either 0 or at least d .

A short sketch by Kühn and Osthus can be found in [9]. This construction gives rise to a couple of extra definitions, with the latter being of special interest to us.

Definition 1.11. We call V_1, \dots, V_k *clusters*, V_0 the *exceptional set* and its vertices *exceptional vertices*. G' is called the *pure digraph*.

Definition 1.12. The *reduced digraph* R of G with parameters ϵ, d, M' is the digraph whose vertices are V_1, \dots, V_k and in which $V_i V_j$ is an edge precisely when $(V_i, V_j)_{G'}$ has density at least d .

2. HOW COMMON ARE ROBUST EXPANDERS?

It is well known that, despite their difficulty to construct, expanding graphs are very common; almost every graph is an expander. However, being a robustly expanding graph is much stronger, so this begs the question of whether the same result holds. The result below has not been proved in any paper online or been published in any journal that I could find - it is my own work.

Theorem 2.1. Let $0 \leq \nu \ll \tau \leq 1/2$. Then asymptotically almost surely $G \in \mathcal{G}(n, 1/2)$ is a robust (ν, τ) -expander.

Proof. In this proof we shall assume omit floors and ceilings, since the argument is unaffected.

Take some $G \in \mathcal{G}(n, 1/2)$ and some $S \subset V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$. We see that

$$\begin{aligned}
 \mathbb{P}(|RN_{\nu, G}(S)| < |S| + \nu n) &= \mathbb{P}\left(\sum_{y \in G} \mathbb{1}_{\{|N(y) \cap S| < \nu n\}} \geq n - (|S| + \nu n)\right) \\
 &\leq \mathbb{P}\left(\sum_{y \in G} \mathbb{1}_{\{|N(y) \cap S| < \nu n\}} \geq (\tau - \nu)n\right) \\
 &\leq \mathbb{P}(\exists Y \subset V(G) \text{ s.t. } |Y| = (\tau - \nu)n \text{ and } \forall y \in Y \ |N(y) \cap S| < \nu n) \\
 &\leq \mathbb{P}(\exists Y \subset V(G) \text{ s.t. } |Y| = (\tau - \nu)n \text{ and } e(Y, S) \leq 2\nu n|Y|),
 \end{aligned}$$

where in the last line we doubled the RHS of the edge-counting inequality to account for the double counting of $xy \in E(G)$ if $x, y \in Y \cap S$.

Take $Y \subset V(G)$ with $|Y| = (\tau - \nu)n$. Write $Y' := Y \cap S =: S'$, $Y'' := Y \setminus S$ and $S'' := S \setminus Y$. Then we may write

$$e(Y, S) = e(Y', S') + e(Y', S'') + e(Y'', S).$$

For ease, write $|S| = \eta n$ and $|Y'| = \xi n$. The distributions of the three terms are

$$\begin{aligned} e(Y', S') &\sim \text{Bin}(\xi n(\xi n - 1)/2, 1/2) \\ e(Y', S'') &\sim \text{Bin}(\xi(\eta - \xi)n^2, 1/2) \\ e(Y'', S) &\sim \text{Bin}((\tau - \nu - \xi)\eta n^2, 1/2), \end{aligned}$$

which is clear from examining the subgraphs with the given vertex sets. All of the terms are independent, and, using basic properties of binomial distributions, we thus know that

$$e(Y, S) \sim \text{Bin}\left(\underbrace{(((\tau - \nu)\eta - \xi^2/2)n - \xi/2)n}_{=: N}, 1/2\right).$$

Note that $\tau \leq \eta \leq (1 - \tau)$ and $\xi \leq \min\{\tau - \nu, \eta\}$. We can use these inequalities to find a simple lower bound of N :

$$(\tau - \nu)\eta - \xi^2/2 \geq (\tau - \nu)(\eta - (\tau - \nu)/2) \geq (\tau - \nu)(\tau - (\tau - \nu)/2) \geq (\tau - \nu)^2/2,$$

and so

$$N \geq ((\tau - \nu)^2 n/2 - \xi/2)n \geq (\tau - \nu)^2 n^2/4 =: M$$

for sufficiently large n . Let X_1, X_2 be binomial distributions with N and M trials respectively and probability $1/2$ for each trial. Then for any k we have that $\mathbb{P}(X_1 \leq k) \leq \mathbb{P}(X_2 \leq k)$. In particular, we have that

$$\begin{aligned} \mathbb{P}(e(Y, S) \leq 2\nu(\tau - \nu)n^2) &\leq \mathbb{P}(X \leq 2\nu(\tau - \nu)n^2) \text{ where } X \sim \text{Bin}(M, 1/2) \\ &= \mathbb{P}(X \leq M/2 - M[1/2 - 8\nu/(\tau - \nu)]) \\ &\leq \exp\left\{-\underbrace{\frac{1}{8}(\tau - \nu)^2(1/2 - 8\nu/(\tau - \nu))n^2}_{=: \lambda}\right\}. \end{aligned}$$

By taking, say, $\nu < \tau/17$ we have that $\lambda > 0$. Note that this bound is independent of the set Y .

We thus conclude that

$$\begin{aligned} \mathbb{P}(G \text{ is not a robust } (\nu, \tau)\text{-expander}) &= \mathbb{P}(\exists S \subset V(G), \tau n \leq |S| \leq (1 - \tau)n \text{ s.t. } |RN_{\nu, G}(S)| < |S| + \nu n) \\ &\leq 2^n \mathbb{P}(|RN_{\nu, G}(S)| < |S| + \nu n) \\ &\leq 2^n \mathbb{P}(\exists Y \subset V(G) \text{ s.t. } |Y| = (\tau - \nu)n \text{ and } e(Y, S) \leq 2\nu n|Y|) \\ &\leq 4^n \mathbb{P}(e(Y, S) \leq 2\nu(\tau - \nu)n^2) \\ &\leq 4^n \exp\{-\lambda n^2\} \rightarrow 0. \end{aligned}$$

Thus, $\mathbb{P}(G \in \mathcal{G}(n, 1/2) \text{ is a robust } (\nu, \tau)\text{-expander}) \rightarrow 1$, as desired. \square

It is worth noting that the parameter relations that I offered in my proof are not optimised, which is why I stated the theorem with $\nu \ll \tau$ rather than $\nu < \tau/17$. One must also note that in the above proof I provided no bound on τ , so we just have the standard bound $\tau < 1$; or, to remove the trivial robust expansion property, $\tau \leq 1/2$. By taking much tighter bounds in the proof, one could most definitely obtain much better parameter relations $\nu \ll \tau$.

3. FINDING A HAMILTON CYCLE IN A ROBUST EXPANDER

The paper [1] in which the notion of robust expansion was formally introduced. Kühn, Osthus and Treglown proved the fundamental result that a robust outexpander with sufficiently large minimum semidegree has a Hamilton cycle. This result is used extensively throughout the remaining chapters, both directly and indirectly.

One aspect of its utility is the fact that if we remove a Hamilton cycle from a robust outexpander, we are still left with a robust outexpander, albeit with worse parameters. So, instead of just obtaining one Hamilton cycle, we can obtain many this way. Naturally, as we continue to remove them we do lose the structure of the graph, so this does not get us near a full decomposition, but this fact will be used when finding Hamilton cycles that cover all edges in $G[V_0]$ in Section 7.

The main use of the following theorem is that we can always find a Hamilton cycle in the reduced digraph, since our first lemma will prove that the reduced digraph inherits the robust expansion property, providing us with a concrete base from which we can try to find many in the original graph G . Furthermore, the property of being a robust outexpander is somewhat ‘fluid’, in the sense that we can often guarantee it for cleverly constructed digraphs derived from G , such that a Hamilton cycle in said digraph corresponds to something that we are seeking in G , for example, in Section 4 said cycle will correspond to the special matching that helps us turn 1-factors into cycles.

The precise statement is given below.

Theorem 3.1. Let n_0 be a positive integer and ν, τ, η be positive constants such that $1/n_0 \ll \nu \leq \tau \ll \eta < 1$. Let G be a digraph on $n \geq n_0$ vertices with $\delta^0(G) \geq \eta n$ which is a robust (ν, τ) -outexpander. Then G contains a Hamilton cycle.

To prove this, we will first prove two lemmas. The first of which will be used in Sections 6 and 7 and it states that if G is a robust outexpander, then the reduced digraph R inherits this property (with slightly worse parameters); the utility of this lemma stretches much further than proving Theorem 3.1. This highlights the strength of robustness, since we could not run through a similar proof if G were just an ordinary outexpander - we will touch on this in the following subsection. The second lemma asserts that if we have a robust outexpander R and a spanning digraph H with sufficiently large minimum semidegree, then we can find a spanning oriented subdigraph R^* which is a robust expander and contains a lot of the edges of H .

Proving the theorem will be reasonably straightforward with the above lemmas in hand. What we eventually want to do is find a subdigraph satisfying the hypotheses of Lemma 3.4 from [10], which will be stated later and provides sufficient conditions for finding a Hamilton cycle in a normal outexpander. We will first find a robustly expanding oriented spanning subdigraph R^* such that it shares many edges with a carefully chosen spanning subdigraph H ; we choose this by keeping the edges in R which correspond to particularly high densities. Having done that, we then obtain G^* from G by only including edges corresponding to R^* . We then control the minimum degrees of the vertices in $G^* - V_0$ by moving poorly-connected vertices to the exceptional set to ensure that $\delta^0(G^* - V_0)$ is sufficiently large. Finally, we add edges from V_0 to G^* to ensure $\delta^0(G^*)$ is large enough, and then we apply Lemma 3.4 to obtain the Hamilton cycle.

We will now state and prove the first lemma.

Lemma 3.2. Let M', n_0 be positive integers and $\epsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll \epsilon \ll d \ll \nu, \tau, \eta < 1$ and such that $M' \ll n_0$. Let G be a digraph on $n \geq n_0$ vertices with $\delta^0(G) \geq \eta n$ 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 $\delta^0(R) \geq \eta |R|/2$ and R is a robust (ν, τ) -outexpander.

Proof. Let G' be the pure digraph, $k := |R|$, V_1, \dots, V_k be the clusters of G (i.e. the vertices of R) and V_0 the exceptional set. Let $m := |V_1|, \dots, |V_k|$.

Firstly, recall that if $V_{i_1} V_{i_2} \in E(R)$, then $d_{G'}(V_{i_1}, V_{i_2}) \geq d$ and each $x \in V_{i_1}$ can trivially have at most m outneighbours in V_{i_2} , and if $V_{i_1} V_{i_2} \notin E(R)$ then $e_{G'}(V_{i_1}, V_{i_2}) = 0$. Furthermore, $G'[V_{i_1}]$ has no edges. Any outedges to V_0 are not included in the information from R . So, after doing the same thing for inedges, we get the following upper bound:

$$\delta^0(G') \leq m\delta^0(R) + |V_0|.$$

Then

$$\delta^0(R) \geq (\delta^0(G') - |V_0|)/m \geq (\delta^0(G) - (d + 2\epsilon)n)/m \geq (\eta - (d + 2\epsilon))n/m \geq \eta k/2,$$

where we used Lemma 1.10 for the second inequality and required $\eta \geq 2(d + 2\epsilon)$ in the last.

Consider any $S \subset V(R)$ with $2\tau k \leq |S| \leq (1 - 2\tau)k$, and let $S' \subset V(G)$ be the union of all the clusters in S . Then $\tau n \leq \tau 2km \leq |S'| \leq (1 - 2\tau)km \leq (1 - 2\tau)n \leq (1 - \tau)n$, using that $n \leq 2km$ if $\epsilon < 1/2$. Using the robust expansion property of G , and requiring $\nu \geq 2(d + \epsilon)$, for each $x \in RN_{\nu, G}^+(S')$ we see $|N_{G'}^-(x) \cap S'| \geq |N_G^-(x) \cap S'| - (d + \epsilon)n \geq \nu n/2$. We hence see that $x \in RN_{\nu, G}^+(S')$ implies that $x \in RN_{\nu/2, G'}^+(S')$, and so

$$|RN_{\nu/2, G'}^+(S')| \geq |RN_{\nu, G}^+(S')| \geq |S'| + \nu n = |S|m + \nu mk.$$

Now, any $x \in RN_{\nu/2, G'}^+(S') \setminus V_0$ must receive edges from at least $|N_{G'}^-(x) \cap S'|/m \geq (\nu n/2)/m \geq \nu k/2$ clusters $V_i \in S$. As discussed above, the cluster V_j containing x must be an outneighbour of each such V_i in R , since the density is non-zero, hence $V_j \in RN_{\nu/2, R}^+(S)$. We thus know $|RN_{\nu/2, G'}^+(S')| \leq m|RN_{\nu/2, R}^+(S)| + |V_0|$, and so, providing $\nu \geq 2\epsilon$, we have that

$$|RN_{\nu/2, R}^+(S)| \geq (|RN_{\nu/2, G'}^+(S')| - |V_0|)/m \geq |S| + \nu k - |V_0|/m \geq |S| + (\nu - \epsilon)k \geq |S| + k\nu/2,$$

confirming that R is a robust $(\nu/2, 2\tau)$ -outexpander. \square

Lemma 3.3. Given positive constants $\nu \leq \tau \leq \eta$, there exists a positive integer n_0 such that the following holds. Let R be a digraph on $n \geq n_0$ vertices which is a robust (ν, τ) -outexpander. Let H be a spanning subdigraph of R with $\delta^0(H) \geq \eta n$. Then R has a spanning orientated subgraph R^* which is a robust $(\nu/12, \tau)$ -outexpander and such that $\delta^0(R^* \cap H) \geq \eta n/4$.

The method of proof is probabilistic. We will form R^* by removing one of xy or yx at random if $xy, yx \in E(R)$. We will then show that the probability that $\delta^0(R^* \cap H) < \nu n/4$ or that there is some $S \subset V(R)$ such that its external $\nu/3$ -robust outneighbourhood in R contains more than $\nu n/6$ vertices that are not in its external $\nu/12$ -robust outneighbourhood in R^* is vanishingly small. This therefore means that (for large n) we have a spanning oriented subgraph that has neither of these two properties. The last thing to do is, for any $S \subset V(R)$ split up $RN_{\nu, R}^+(S) = EN \cup N$ into the vertices outside S and those inside (respectively) and consider a random partition of $S = S_1 \cup S_2$ such that $|N_R^-(x) \cap S_i| \geq \nu n/3$ for all $x \in N$; it is clear why we choose such a partition, given the property our spanning subgraph now has. We then note that $S_1 \cap N \subset RN_{\nu/3, R}(S_2) \setminus S_2$ by our choice of partition, but then we instantly know that it must therefore be the case that all but at most $\nu n/6$ vertices of $S_1 \cap N$ are in $RN_{\nu/12, R^*}(S_2) \setminus S_2 \subset RN_{\nu/12, R^*}^+(S)$. We have the same for $S_2 \cap N$, and we trivially have the same conclusion for EN . From this, we thus see that $|RN_{\nu/12, R^*}^+(S)| \geq |S| + \nu n/2$, i.e. R^* is a robust $(\nu/12, \tau)$ -outexpander with the required minimum semidegree condition.

Proof of Lemma 3.3. As mentioned above, we begin by considering a random spanning oriented subgraph R^* by removing one of xy or yx if they both are in $E(G)$; we do this independently from all other such pairs. To analyse this properly, we will introduce some notation. Take any $x \in V(R)$. Let $N_R^\pm(x) := N^+(x) \cap N^-(x)$ for each $x \in V(R)$ and write $H^* := H \cap R$.

Suppose that $|N_H^\pm(x)| \leq 3\eta n/4$. So, upon removing edges at random, at most $3\eta n/4$ vertices can be lost in $R^* \cap H$ from each of its in- and outneighbourhoods of x , thus $d_{H^*}^+(x), d_{H^*}^-(x) \geq \eta n/4$. We thus assume that $|N_H^\pm(x)| \geq 3\nu/4$.

Write $X := |N_H^\pm(x) \cap N_{H^*}^+(x)|$ for the number of $y \in N_H^\pm(x)$ such that the inedge yx gets deleted. Writing $N := |N_H^\pm(x)|$, we see $X \sim \text{Bin}(N, 1/2)$ and so $\mathbb{E}[X] \geq 3\nu n/8$. Also note that $X \geq d_{H^*}(x)$ since x may have had vertices in $N^+(x) \setminus N^-(x)$. Then using the Chernoff-type bound (1.2), we have that

$$\mathbb{P}(d_{H^*}^+(x) < \nu n/4) \leq \mathbb{P}(X < \nu n/4) \leq \mathbb{P}(X < 2\mathbb{E}[X]/3) < 2e^{-c\mathbb{E}[X]} \leq 2e^{-3\nu n/8},$$

where c is the constant guaranteed by (1.1) that is independent of ν, τ and η . We can do the exact same thing for the indegree.

Take any $S \subset V(R^*) = V(R)$. We define the external ν -robust outneighbourhood in the obvious way: $ERN_{\nu, R}^+(S) := RN_{\nu, R}^+(S) \setminus S$. For want of a better word, we will say S is *good* if $|ERN_{\nu/3, R}^+(S) \setminus ERN_{\nu/12, R^*}^+(S)| \leq \nu n/6$, i.e. all but at most $\nu n/6$ vertices in $ERN_{\nu/3, R}^+(S)$ are in $ERN_{\nu/12, R^*}^+(S)$. The goal is now to show that the probability of S not being good is vanishingly small as $n \rightarrow \infty$. Now, S can be any subset of $\mathcal{P}([n])$, so we will need to probability that S is not good to go to zero quicker than 2^n grows, so trying to prove that it is less than e^{-n} seems like a reasonable goal.

We begin by writing $ERN_R^\pm(S) := \{x \in ERN_{\nu/3, R}^\pm(S) : |N_R^\pm(x) \cap S| \geq \nu n/4\}$, i.e. the elements in the external $\nu/3$ -robust outneighbourhood that also have edges in both directions with a significant subset of S . Take some $x \in ERN_{\nu/3, R}^+(S) \setminus ERN_R^\pm(S)$, then $|N_R^\pm(x) \cap S| < \nu n/4$ by definition, and so in R^* we lose at most $\nu n/4$ of its inneighbours inside of S , i.e. $|N_{R^*}^-(x) \cap S| \geq (1/3 - 1/4)\nu n = \nu n/12$, thus $x \in ERN_{\nu/12, R^*}^+(S)$. We say that $x \in ERN_R^\pm(S)$ *fails* if $x \notin ERN_{\nu/12, R^*}^\pm(S)$. Write $Y = |N_{R^*}^-(x) \cap N_R^\pm(x) \cap S|$, then $\mathbb{E}[Y] \geq \nu n/8$ trivially. So,

$$\mathbb{P}(x \text{ fails}) \leq \mathbb{P}(Y < \nu n/12) \leq 2e^{-c\nu n/8} =: p.$$

Let Z be the number of $x \in ERN_R^\pm(S)$ that fail. Trivially, $\mathbb{E}[Z] \leq pn$. By (1.1) have that,

$$\mathbb{P}(Z \geq \beta \mathbb{E}[Z]) \leq e^{(\beta - (\beta+1) \log(\beta+1))\mathbb{E}[Z]} \leq e^{(\beta - \beta \log \beta)\mathbb{E}[Z]} \leq e^{-\beta \log \beta \mathbb{E}[Z]/2},$$

where in the last inequality we require $\beta \geq e^2$. Setting $\beta := \nu n/(6\mathbb{E}[Z]) \geq \nu/(6p) = \nu e^{c\nu n/8}/6$ (which is greater than e^2 for large n) gives

$$\mathbb{P}(S \text{ is not good}) = \mathbb{P}(Z > \nu n/6) = \mathbb{P}(Z > \beta \mathbb{E}[Z]) \leq e^{-\nu n(\log \beta)/12} \leq e^{-n},$$

where the last inequality is true for large n (substitute the lower bound for β and it falls out easily).

So, putting all of this together, $\mathbb{P}(\delta^0(H^*) < \eta n/4 \text{ or } \exists S \text{ that is not good}) \leq 4ne^{-3c\eta n/8} + 2^n e^{-n} \rightarrow 0$, using a simple union bound, and so for sufficiently large n we can find an R^* with the required minimum semidegree and with the property that every $S \subset V(R)$ is good. The only thing that remains to be checked is that this is a robust $(\nu, 12)$ -outexpander.

Consider any $S \subset V(R)$ with $\tau n \leq |S| \leq (1 - \tau)n$. For ease we will write the external and internal neighbourhoods as $EN := ERN_{\nu, R}^+(S)$ and $N := RN_{\nu, R}^+(S) \cap S$, so $RN_{\nu, R}^+(S) = EN \cup N$. Since S is good and $EN \subset ERN_{\nu/3, R}^+(S)$, it follows that all but at most $\nu n/6$ vertices in EN are in $ERN_{\nu/12, R^*}^+(S) \subset RN_{\nu/12, R^*}^+(S)$. We now have the following important claim that will help us split up N .

Claim 1. There is a partition of S into S_1, S_2 such that every vertex $x \in N$ satisfies $|N_R^-(x) \cap S_i| \geq \nu n/3$ for $i = 1, 2$.

Consider a random partition S_1, S_2 of S such that each $y \in S$ is in either set with probability $1/2$, independently of all other vertices. Let $X \sim \text{Bin}(\nu n, 1/2)$. Take some $x \in N$, then

$$\begin{aligned} \mathbb{P}(|N_R^-(x) \cap S_i| < \nu n/3 \text{ for some } i) &\leq 2\mathbb{P}(|N_R^-(x) \cap S_1| < \nu n/3) \\ &\leq 2\mathbb{P}(X \leq \nu n/3) \leq 2e^{-n/72}, \end{aligned}$$

using (1.5). We thus conclude that

$$\mathbb{P}(\text{No such partition exists}) \leq 2|N|e^{-n/72} \leq 2ne^{-n/72} \rightarrow 0,$$

and so for sufficiently large n we can find such a partition, proving the claim.

Use any such partition. Then we see that $S_1 \cap N \subset ERN_{\nu/3, R}(S_2)$, but S_2 is good, so all but at most $\nu n/6$ vertices of $S_1 \cap N$ are also in $ERN_{\nu/12, R^*}^+(S_2) \subset RN_{\nu/12, R^*}^+(S)$. Similarly, all but at most $\nu n/6$ vertices in $S_2 \cap N$ are in $RN_{\nu/12, R^*}^+(S)$. Compiling all of this, we thus have that

$$|RN_{\nu/12, R^*}^+(S)| \geq |EN \cup (S_1 \cap N) \cup (S_2 \cap N)| - 3 \cdot \nu n/6 = |RN_{\nu, R}^+(S)| - \nu n/2 \geq |S| + \nu n/2,$$

as required. \square

3.1. Proof of Theorem 3.1. With these two lemmas in our arsenal, we will be able to prove Theorem 3.1 with the following lemma from [10] (it is Theorem 2 in the paper). For this we will introduce a simple notion of outexpansion that will not be used in the rest of the essay: we call a digraph G a (ν, τ) -outexpander if $|N^+(S)| \geq |S| + \nu|G|$ for all $S \subset V(G)$ with $\tau|G| < |S| < (1 - \tau)|G|$.

Lemma 3.4. Let M', n_0 be positive integers and let $\epsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll \epsilon \ll d \ll \nu, \tau, \eta < 1$ and such that $M' \ll n_0$. Let G be an oriented graph on $n \geq n_0$ vertices such that $\delta^0(G) \geq 2\eta n$. Let R be the reduced digraph of G with parameters ϵ, d and M' . Suppose there exists a spanning orientated subgraph R^* of R with $\delta^0(R^*) \geq \eta|R^*|$ which is a (ν, τ) -expander. Then G contains a Hamilton cycle.

Before proving Theorem 3.1, we will expand on the brief sketch offered at the start of the section and motivate some of the ideas.

The broad goal is to find oriented spanning subdigraphs $G^* \subset G$ and a $R^* \subset R$ such that R^* has sufficiently large minimum degree, is an outexpander and corresponds to the reduced digraph of G^* with some parameters. Having this, we can simply apply Lemma 3.4 with G^* playing the roles of G' and G and R^* those of R and R^* .

Naturally, the first step is to apply Szemerédi's regularity lemma to obtain G' and R . We then find R^* . Lemma 3.3 will give us a (robust) outexpander for any H with large minimum semidegree. Since we want $G^* - V_0$ to have a large minimum semidegree with the clusters corresponding to regular pairs with high densities (in order to argue that the pair correspond to a partition using Szemerédi's lemma with certain parameters), it is logical to pick H to be correspond to edges $V_i V_j \in E(R)$ with a particularly high density. We then construct G^* by modifying G' , removing edges xy joining V_i to V_j for $V_i V_j \in E(R) \setminus E(R^*)$, which is now oriented. There is then the worry that some vertices in $G^* - V_0$ will have small in- and outdegrees. It turns out there are not that many, and we simply move them to V_0 . We then check the hypotheses of Szemerédi's regularity lemma to apply Lemma 3.4.

Proof of Theorem 3.1. We firstly pick $M' \in \mathbb{N}$ and ϵ, d such that $1/n_0 \ll 1/M' \ll \epsilon \ll d \ll \nu$. We next apply Szemerédi's regularity lemma with parameters ϵ, d and M' to G to obtain clusters V_1, \dots, V_k of size m , an exceptional set V_0 of size smaller than ϵn and a pure digraph G' . We instantly get (from the regularity lemma) that $\delta^0(G') \geq (\eta - (d + \epsilon))n$, and Lemma 3.2 says that $\delta^0(R) \geq \eta k/2$ and that R is a robust $(\nu/2, 2\tau)$ -outexpander.

We want to apply Lemma 3.3 to find a robustly expanding oriented spanning subgraph R^* of sufficiently large minimum semidegree. The natural thing to do, therefore, is find a spanning subdigraph of sufficiently large minimum semidegree. However, we can not pick any such spanning subdigraph, as we want the resulting H^* to correspond to clusters which have high-connectivity.

We let H be the spanning subdigraph of R in which $V_i V_j$ is an edge if $V_i V_j \in E(R)$ and $d_{G'}(V_i, V_j) \geq \eta/4$. Note that

$$\sum_{V_j \in N_R^+(V_i)} d_{G'}(V_i, V_j) m^2 = e_{G'}(V_i, V(G) \setminus V_0) \geq \delta^0(G') m - |V_0| m \geq (\eta - 2d) n m.$$

Claim 1. There are at least $\eta k/4$ outneighbours V_j of V_i in R such that $d'_G(V_i, V_j) \geq \eta/4$.

To prove this we suppose the contrary. Then we have

$$\sum_{V_j \in N_R^+(V_i)} d'_G(V_i, V_j) \leq \eta k/4 \cdot 1 + (d_R^+(V_i) - \eta k/4) \cdot \eta/4 \leq (2 - \eta/4) k \eta/4.$$

Now, consider the following inequality:

$$k \eta/4 (2 - \eta/4) < (\eta - 2d) k \Leftrightarrow \eta(8 - \eta) < 16(\eta - 2d) \Leftrightarrow d < (\eta^2 + 8\eta)/32.$$

And therefore, if we pick d in this range, then we contradict the fact that $\sum_{V_j \in N_R^+(V_i)} d'_G(V_i, V_j) \geq (\eta - 2d) n/m \geq (\eta - 2d) k$, thus completing the proof of the claim.

Now, each such V_j also belongs to $N_H^+(V_i)$, so $\delta^+(H) \geq \eta k/4$. We similarly get $\delta^-(G) \geq \eta k/4$, and as a result $\delta^0(H) \geq \eta k/4$. Applying Lemma 3.3 to obtain a spanning oriented subdigraph R^* which is a robust $(\nu/24, 2\tau)$ -outexpander with $\delta^0(H^*) \geq \eta k/16$, where $H^* := H \cap R^*$.

The goal now is to modify G' into a spanning oriented subgraph of G having minimum semidegree at least $\eta^2 n/100$. Let G^* be the spanning subdigraph of G' corresponding to R^* , i.e. deleting edges xy that join clusters V_i and V_j for $V_i V_j \in E(R) \setminus E(R^*)$. We begin by noting that $G^* - V_0$ is an oriented graph. However, we have little information regarding the degrees of vertices in $G^* - V_0$, which poses a problem when aiming to apply Lemma 3.4. The plan will thus be to identify vertices with not many neighbours and add them to V_0 so that the in- and outdegrees outside of V_0 are large.

Claim 2. Take any cluster V_i and any $V_j \in N_{H^*}^+(V_i)$, then at most ϵm vertices in V_i have less than $(d_{G'}(V_i, V_j) - \epsilon) m \geq \eta m/5$ outneighbours in V_j (in the digraph G'). We call these vertices of V_i *useless* for V_j .

As above, we prove this by contradiction. Supposing the above is not true, we have that

$$d_{G'}(V_i, V_j) m^2 = e_{G'}(V_i, V_j) \leq \epsilon m \cdot (d_{G'}(V_i, V_j) - \epsilon) m + (1 - \epsilon) m \cdot m.$$

If we have that $1 - \epsilon^2/(1 - \epsilon) < d_{G'}(V_i, V_j)$, then we have a contradiction. But we can definitely do this, since $d_{G'}(V_i, V_j) \geq d$ and $\epsilon \ll d$. The claim is now proved.

Claim 3. On average any vertex of V_i is useless for at most $\epsilon |N_{H^*}^+(V_i)|$ clusters $V_j \in N_{H^*}^+(V_i)$. Thus at most $\sqrt{\epsilon} m$ vertices in V_i are useless for more than $\sqrt{\epsilon} |N_{H^*}^+(V_i)|$ clusters.

Denote the clusters $V_j \in N_{H^*}^+(V_i)$ for which $x \in V_i$ is useless by $U(x)$. The average number of clusters that a vertex of V_i is useless for is given by $\sum_{x \in V_i} |U(x)|/m$ which we may write as

$$\frac{1}{m} \sum_{x \in V_i} \sum_{V_j \in N_{H^*}^+(V_i)} \mathbb{1}_{V_j \in U(x)} = \frac{1}{m} \sum_{V_j \in N_{H^*}^+(V_i)} \sum_{x \in V_i} \mathbb{1}_{V_j \in U(x)} \stackrel{(\text{Claim 2})}{\leq} \epsilon |N_{H^*}^+(V_i)|.$$

As for the final part of the claim, let $A_i := \{x \in V_i : |U(x)| > \sqrt{\epsilon} |N_{H^*}^+(V_i)|\}$.

$$\sum_{x \in V_i} |U(x)| \geq \sum_{x \in A_i} |U(x)| \geq \sqrt{\epsilon} |N_{H^*}^+(V_i)| |A_i|$$

We thus see that $|A_i| \leq \sqrt{\epsilon}m$, as desired.

Let $U_i^+ \subset V_i$ be a set of size $\sqrt{\epsilon}m$ which consists of all vertices in V_i that are useless for more than $\sqrt{\epsilon}|N_{H^*}^+(V_i)|$ clusters (and some extra vertices if necessary, so that the sizes match up). Similarly, choose $U_i^- \subset V_i \setminus U_i^+$ of size $\sqrt{\epsilon}m$ such that for each $x \in V_i \setminus U_i^+$ there are at most $\sqrt{\epsilon}|N_{H^*}^-(V_i)|$ clusters $V_j \in N_{H^*}^-(V_i)$ such that x has less than $\nu m/5$ inneighbours in V_j . For each $i = 1, \dots, k$, move the vertices in $U_i^+ \cup U_i^-$ from V_i to V_0 . For ease, we will denote the clusters by V_1, \dots, V_k and the exceptional set by V_0 . In doing this, we now have that $|V_0| \leq \epsilon n + 2\sqrt{\epsilon}km \leq 3\sqrt{\epsilon}n$. Moreover,

$$\begin{aligned} \delta^0(G^* - V_0) &\geq \frac{\eta m}{5}(1 - \sqrt{\epsilon})\delta^0(H^*) - |V_0| \\ &\geq \frac{\eta m}{5} \frac{\eta k}{17} - 3\sqrt{\epsilon}n \\ &\geq \left(\frac{(1 - \sqrt{\epsilon})\eta^2}{85} - 3\sqrt{\epsilon} \right) n \\ &\geq \frac{\eta^2 n}{100}, \end{aligned}$$

where in the penultimate equality we used that $n = |G| \leq mk + \epsilon n \leq mk + \sqrt{\epsilon}n$ and in the final equality we used that $\epsilon \ll \eta$. One possible such function could be $\epsilon \leq 9\eta^4/(1700^2(3 + \eta^2/85)^2) \leq \eta^4/1700^2$; however, I omit the tedious calculation.

The final step is to modify G^* by incorporating the exceptional set V_0 in such a way that we can apply Lemma 3.4. For each $x \in V_0$, we take a set of $\eta n/2$ outneighbours of x in G and $\eta n/2$ inneighbours in G such that these sets are disjoint, and we add edges between x and the selected vertices in G^* ; we still denote the oriented graph G^* . Clearly, $\delta^0(G^*) \geq \eta^2 n/100$.

Claim 4. The partition V_0, \dots, V_k of $V(G^*)$ described above is as in described in Szemerédi's Regularity Lemma with parameters $3\sqrt{\epsilon}$, $d - \epsilon$ and M' (with G^* playing the role of G' and G).

The first six properties are trivial. We only need to show that for $1 \leq i, j \leq k$ with $i \neq j$ the pair $(V_i, V_j)_{G^*}$ is $3\sqrt{\epsilon}$ -regular and has density 0 or at least $d - \epsilon$.

Take $1 \leq i, j \leq k$ with $i \neq j$. Now, note that in the rest of the proof we continued labelling the clusters as V_i despite removing vertices and adding them to V_0 . So, for the sake of this claim, let \bar{V}_i be the i th cluster in G' ; note that $\bar{V}_i = V_i \cup U_i^+ \cup U_i^-$. The first thing to note is that $G^*[V_i, V_j] = G'[V_i, V_j]$, since the edges from V_i to V_j were left untouched, and so we trivially have that $d_{G'}(V_i, V_j) = d_{G^*}(V_i, V_j)$.

Take $A \subset V_i$ and $B \subset V_j$ with $|A|, |B| \geq 3\sqrt{\epsilon}(1 - 2\sqrt{\epsilon})m \geq \epsilon m = \epsilon|\bar{V}_i|$ if $\epsilon \leq (3/7)^2$. By using the triangle-inequality and the fact that $G'[V_i, V_j]$ is ϵ -regular, we have that

$$\begin{aligned} |d_{G^*}(V_i, V_j) - d_{G^*}(A, B)| &= |d_{G'}(V_i, V_j) - d_{G'}(A, B)| \\ &\leq |d_{G'}(V_i, V_j) - d_{G'}(\bar{V}_i, \bar{V}_j)| + |d_{G'}(\bar{V}_i, \bar{V}_j) - d_{G^*}(A, B)| \\ &\leq 2\epsilon \leq 3\sqrt{\epsilon}, \end{aligned}$$

and so $(V_i, V_j)_{G^*}$ is $3\sqrt{\epsilon}$ -regular.

By construction, $d_{G^*}(V_i, V_j) = 0$ if $V_i V_j \notin E(R^*)$, and so it is sufficient to show that edges in R^* correspond to clusters in G^* of density at least $d - \epsilon$. But, this is immediate from the ϵ -regularity of $G'[V_i, V_j]$ (which we can use since $1 - 2\sqrt{\epsilon} \geq \epsilon$ for $\epsilon \leq 3 - 2\sqrt{2}$):

$$d_{G^*}(V_i, V_j) = d_{G'}(V_i, V_j) \geq d_{G'}(V_i, V_j) - \epsilon \geq d - \epsilon.$$

We can say thus say that R^* is the reduced digraph of G^* with these parameters. Apply Lemma 3.4 to R^* and G^* to obtain a Hamilton cycle in G^* and thus in G . \square

3.2. A note on the parameter relations. In this section, I have included possible functions governing the parameter relations ‘ \ll ’. Since they all were in the same direction, it is obvious that they can be made consistent. However, I included each one to be explicit, so that we could see exactly what was going on. Recall that these inequalities are specific to my argument, and Kühn, Osthus and Treglown may have had tighter bounds in mind. For the rest of the essay, these will, in general, be omitted.

3.3. What happens if we drop the robustness requirement? Kühn, Osthus and Treglown mention in [1] that for Lemmas 3.2 and 3.3 to work the graphs must be robust expanders, and not ordinary expanders. Unfortunately, since each proof relies on a decomposition with Szemerédi’s regularity lemma, of which there are a huge amount - which we roughly quantify in a moment - constructing a counterexample is essentially impossible. What we will try to do instead is run through the proof of Lemma 3.2 and suggest where it goes wrong, i.e. where being an ordinary (ν, τ) -outexpander fails. Although the form of Szemerédi’s lemma here is quite different than that in Gowers’ 1997 paper [4], he obtains a lower bound for the number of different partitions given by a tower of 2s of height proportional to $\log(1/\epsilon)$ - this should hopefully confirm the hypothesis that constructing an explicit counterexample is out of the question.

Everything up to taking the set $S \subset V(G)$ and defining S' is exactly the same. In the proof of Lemma 3.2, we then argue that if $x \in RN_{\nu, G}^+(S')$ then $x \in RN_{\nu/2, G'}^+(S')$. However, we have no control over this with an ordinary expander. The knowledge that $|N_G^+(S')| \geq |S|m + \nu n \geq |S|m + \nu mk$ does not tell us anything useful about $|N_{G'}^+(S')|$; all we know is that it is greater than $|S|m + \nu mk - (d + \epsilon)nm|S|$. This fact is not useful in the slightest - this number could even be negative - and does not let us progress with the rest of the proof.

Note also that if $x \in N_{G'}^+(S')$ and V_j is the cluster containing x , then $V_j \in N_R^+(S)$. So we obtain a similar inequality, $|N_{G'}^+(S')| \leq m|N_R^+(S)| + |V_0|$. So, if we were to continue in a similar way, we would write:

$$|N_R^+(S)| \geq (|N_{G'}^+(S')| - |V_0|)/m \geq |N_{G'}^+(S')|/m - \epsilon k.$$

However, as mentioned above, we have no control over the size of $N_{G'}^+(S')$, so obtaining that the above is at least $|S| + k\nu/2$ seems impossible without extra information.

In summary, I believe that the necessity of robustness is as follows. We want to be able to control the degrees in G' so that we can infer something about them in R . In the proof of Lemma 3.2, we immediately note that for $x \in RN_{\nu, G}^+(S')$, $|N_{G'}^-(x) \cap S'| \geq \nu n/2$ using the robustness property, concluding that $x \in RN_{\nu/2, G'}^+(S')$. This is since x loses at most $(d + \epsilon)n$ neighbours in G' , but we already know it has at least νn neighbours inside of S' , so we are all good. However, if $x \in N_G^+(S')$ we simply have no idea about the number of vertices of S' to which x was connected, so it is very possible that in G' any such edge may not exist, in which case $x \notin N_{G'}^+(S')$.

4. CONVERTING 1-FACTORS INTO HAMILTON CYCLES

Constructing many Hamilton cycles is very tough in general, so an intermediate step is to first find 1-factors F (which is not necessarily easy, but it is most definitely a lot easier) and then extend these to Hamilton cycles. Moreover, every single section of the proof of Theorem 1.3 explicitly concerned with finding Hamilton cycles (rather than setting up machinery) begins by finding 1-factors and using the lemmas of this section to extend them. But, the question is: How do we do this?

The general strategy will be to swap edges of F with edges in G in order to merge the cycles, with the goal of eventually making it a Hamilton cycle, i.e. merging all of them. To do this, we use the fact that G has very strong regularity properties; concretely, we use the *superregularity* of G .

4.1. Definitions. We introduce the notion of two types of superregularity in this section rather than at the beginning, since they are not required elsewhere. We will introduce this in terms of undirected bipartite graphs, since we tend to view $G[U, V]$ as undirected.

Definition 4.1. A bipartite $G = (U, V)$ is (ϵ, d) -superregular if it is ϵ -regular and $d_G(u) \geq (d - \epsilon)|V|$ for all $u \in U$ and $d_G(v) \geq (d - \epsilon)|U|$ for every $v \in V$.

Definition 4.2. Let G be a bipartite graph with vertex classes U and V , both of size m . Given $0 < \epsilon, d, c < 1$, we say that G is (ϵ, d, c) -regular if the following conditions are satisfied:

- (Reg1) Whenever $A \subset U$ and $B \subset V$ are sets of size at least ϵm , then $d(A, B) = (1 \pm \epsilon)d$.
- (Reg2) For all $u, u' \in U$ we have $|N(u) \cap N(u')| \leq c^2 m$. Similarly for $v, v' \in V$.
- (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 holds:

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

Even though these two notions of superregularity are different, in the sense that the latter does not imply the former, it is evident that the latter provides a much stricter framework. The point is that for sufficiently large subsets of U and V we know exactly the number of edges between them, but we also want the edges to be spread out in some sense, which is what (Reg2) achieves. Finally, (Reg3) and (Reg4) simply provide upper and lower bounds on the degrees, as we do not want any particular vertices receiving too many edges or too few.

4.2. Preliminary superregularity lemma. One reason for introducing the notion of superregularity is that one can show that, under specific parameter relations, such a bipartite graph can easily be made into a robustly expanding digraph. This, naturally, will be fundamental for our later purposes.

Lemma 4.3. Let $0 < 1/m \ll \nu \ll \tau \ll d \leq \epsilon \ll \mu, \zeta \leq 1/2$ and let G be a $(\epsilon, d, \zeta d, d/\mu)$ -superregular bipartite graph with vertex classes U and V of size m . Let $A \subset U$ be such that $\tau m \leq |A| \leq (1 - \tau)m$. Let $B \subset V$ be the set of all those vertices in V which have at least νm neighbours in A . Then $|B| \geq |A| + \nu m$.

To see how this can be made into a robustly expanding digraph, first orient all edges from U to V and then take any bijection $\phi : [m] \rightarrow [m]$. Identify each pair $(u_i, v_{\phi(i)})$ into one vertex whose in- and outneighbourhood are the points identified with $N_G(v_{\phi(i)})$ and $N_G(u_i)$ respectively; call this graph H . Lemma 4.3 then trivially implies that H is a robust (ν, τ) -outexpander with minimum semidegree at least ζdm .

Proof of Lemma 4.3. We first prove the following claim. It essentially says that if a subset of either vertex class is not too small then its *robust neighbourhood* is also not too small.

Claim 1. Let $U' \subset U$ be such that $|U'| \geq \tau m/2$ and let $RN(U')$ be the vertices in V which have at least νm neighbours in U' . Then $|RN(U')| \geq \min\{10\epsilon m, |U'|/\sqrt{\epsilon}\}$. The claim is identical replacing U by V and U' by $V' \subset V$.

The method here will be to assume $|RN(U')| \leq 10\epsilon m$ and then show that it must be the case that $|RN(U')| \geq |U'|/\sqrt{\epsilon}$ by considering paths of length two with end vertices in V and the middle vertex in U' .

Firstly, we define a function $f : \mathcal{P}(V) \times \mathcal{P}(V) \rightarrow \mathbb{N}_0$, where $\mathcal{P}(V)$ denotes the power set of V , such that $f(V', V'')$ counts the number of paths of length two which have one end

vertex in each of V' and V'' with the midpoint in U' . (Reg4) gives that $\delta(G) \geq \zeta dm$, and so

$$f(V, V) \geq |U'| \binom{\zeta dm}{2} \geq |U'| \frac{\zeta^2 d^2 m^2}{3} \quad (4.1)$$

if $1/m \leq d\zeta/3$ (recall $1/m \ll \nu \ll \dots \ll \zeta$). By the definition of $RN(U')$, it follows that $f(V \setminus RN(U'), V) \leq \nu |V \setminus RN(U')| m^2$.

Now, consider a path of length two starting and ending anywhere in V with its midpoint in U' . It can either have both endpoints in $RN(U')$, or at most endpoint in $RN(U')$. So, we can split $f(V, V)$ up as follows:

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

In the final inequality, we used that $|RN(U')| \leq 10\epsilon m$. Now, if we combine this with (4.1) and rearrange, we get that

$$\begin{aligned} |RN(U')| &\geq \frac{\mu^2}{10\epsilon d^2 m^2} \left(|U'| \frac{\zeta^2 d^2 m^2}{3} - \nu m^3 \right) \\ &\geq \frac{\mu^2}{10\epsilon d^2 m^2} |U'| \frac{\zeta^2 d^2 m^2}{4} \\ &\geq \frac{|U'|}{\sqrt{\epsilon}}, \end{aligned}$$

where in the final two inequalities are due to the relations $\nu \ll \tau, d, \zeta$ and $\epsilon \ll \mu, \zeta$, which we can justify by offering examples of such functions: $\nu \leq \frac{\tau \zeta^2 d^2}{12}$ and $\epsilon \leq \mu^4 \zeta^4 / 1600$. We can easily make these equations, amongst others, consistent, as we choose from right to left.

In order to prove the lemma, we split the work into several cases according to the size of A , and either use superregularity or the claim just proved. For the sake of notation, we will write $B = RN(A)$.

Suppose that $\tau m \leq |A| \leq \epsilon m$. Then the claim implies that $|B| \geq \min\{10\epsilon m, |A|/\sqrt{\epsilon}\}$. Note that $10\epsilon m \geq |A| + 9\epsilon m \geq |A| + \nu m$, and $|A|/\sqrt{\epsilon} \geq |A| + \nu m$ if $\nu \leq (1 - \sqrt{\epsilon})\tau/\sqrt{\epsilon}$, so we are done.

The next case is $\epsilon m \leq |A| \leq (1 - 2\epsilon)m$. Note that, by definition of B ,

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

which is implied by, say, $\nu < (1 - \epsilon)\epsilon d$. This implies that $d(A, V \setminus B) < (1 - \epsilon)d$. By (Reg1), it follows that $|V \setminus B| < \epsilon m$, and so $|B| \geq (1 - \epsilon)m \geq |A| + \nu m$.

The final case is $(1 - 2\epsilon)m \leq |A| \leq (1 - \tau)m$. We suppose that for a contradiction $|B| < |A| + \nu m$. Now, consider the previous case, then by restriction we know that $|B| \geq (1 - \epsilon)m \geq (1 - 2\epsilon)m + \nu m$. For simplicity, define $V' := V \setminus B$. Note that $m - |A| \geq \tau m$, and so, if $\nu \leq \tau/2$, we have that

$$\tau m/2 \leq m - |A| - \nu m < m - |B| = |V'| \leq 2\epsilon m - \nu m < 2\epsilon m. \quad (4.2)$$

Define $A' := A \cap RN(V')$. By definition, every vertex in A' has at least νm neighbours in V' . Recall that every vertex in V' has less than νm neighbours in $A \supset A'$. We deduce that

$|A'|\nu m \leq e(A', V') \leq |V'|\nu m$, and so $|A'| \leq |V'|$. Thus,

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

using (4.2) and the fact $\epsilon \leq \nu$ in the last three inequalities. However, since $|V'| > \tau m/2$, this contradicts the claim. \square

4.3. Obtaining Hamilton cycles. There are two lemmas here that are very important for later sections.

In [2, 3] the following lemmas were actually stated incorrectly. The proof given in [2] was also incorrect due to how they were stated, so I have provided a fix so that the proofs are now correct. Moreover, I also propose a simpler hypothesis for both proofs, so that the proof would be correct as it is written in [2]. I emailed Daniela Kühn, and she agreed with both the spot and the fixes.

Lemma 4.4. Suppose that $0 < 1/m \ll d' \ll \epsilon \ll d \ll \zeta, 1/t \leq 1/2$. 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 $J \subset E(G)$. For each edge $V_i V_{i+1} \in J$, let $V_i^1 \subset V_i$ and $V_{i+1}^2 \subset V_{i+1}$ be such that $|V_i^1| = |V_{i+1}^2| \geq m/100$ and such that $G[V_i^1, V_{i+1}^2]$ is $(\epsilon, d', \zeta d', td'/d)$ -superregular. Suppose that F is a 1-regular digraph with $V_1 \cup \dots \cup V_k \subset V(F)$ such that the following properties hold:

- (i) For each edge $V_i V_{i+1} \in J$ 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 $V_i V_{i+1} \in J$ such that D contains a vertex in V_i^1 .
- (iii) Whenever $V_i V_{i+1}, V_j V_{j+1} \in J$ are such that J avoids all edges in the segment $V_{i+1} C V_j$ 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 .

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 $V_i V_{i+1} \in J$. Moreover, if $J = E(C)$ then the last condition can be replaced by: $V_i^1 \cap V_i^2 \neq \emptyset$ for all $i = 1, \dots, k$.

The point here is that F is a vertex-disjoint union of cycles, and so the goal is to replace some edges of F with edges from G to turn these cycles into bigger cycles. Call the resulting 1-factor F' . Doing this repeatedly would hopefully yield a full Hamilton cycle. So the question is: What do we do? A reasonable idea would be to ensure that all vertices in G_i lie in the same cycle in F' obtained from F , since they may have all been in different cycles before. We want to exploit the superregularity of G_i to achieve this goal, and we do this by showing constructing an auxiliary digraph A that identifies each $u \in V_{i+1}^2$ with $f(u) \in V_i^1$ which is the first time the cycle C_u starting from u meets V_i^1 , which is guaranteed by (i). Moreover, this vertex is unique. The superregularity and the discussion surrounding Lemma 4.3 give us a robust expander to which we apply Theorem 3.1 which corresponds to a perfect matching, with which we replace $F[V_i^1, V_{i+1}^2]$, so that all vertices in V_i^1 and V_{i+1}^2 lie on the same cycle. We then iterate this and then justify that the net result is a Hamilton cycle.

Proof of Lemma 4.4. Take $V_i V_{i+1} \in J$, let Old_i be the perfect matching in $F[V_i, V_{i+1}]$. We first prove the following:

Claim. For any edge $V_i V_{i+1} \in J$, we can find a perfect matching New_i in 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 will also be in F' .

We begin by picking ν, τ such that $1/m \ll \nu \ll \tau \ll d'$. For every $u \in V_{i+1}^2$, it belongs to some cycle C_u , so we let $f(u)$ be the first vertex in V_i^1 (after starting the cycle from u). This

exists by (i). $f(u) \neq f(v)$ for $u \neq v$ since all edges are oriented from V_i^1 to V_{i+1}^2 and each vertex in V_i^1 has in- and outdegree equal to 1. We now define an auxiliary digraph A on $V_{i+1}^2 \times V_i^1$ by adding an edge from $(u, f(u))$ to $(v, f(v))$ if $f(u)v \in E(G)$. We thus see that $|N_A^+(u, f(u))| = |N_{G_i}(f(u))|$ and that $|N_A^-(u, f(u))| = |N_{G_i}^-(u)|$, with the last equality coming from the fact that each $w \in V_i^1$ has a unique $v \in V_{i+1}^2$ such that $w = f(v)$ and so $(u, f(u))$ in A will receive the same amount of edges as u in G_i . The first conclusion is that $\delta^0(A) = \delta^0(G_i) \geq \zeta d' |V_{i+1}^2| = \zeta d' |A|$. Moreover, by the discussion after Lemma 4.3, A is a robust (ν, τ) -outexpander, and so, with the previous observation and Theorem 3.1, A has a Hamilton cycle $K = (u_1, f(u_1))(u_2, f(u_2)) \dots (u_m, f(u_m))$. So we let our new matching New_i have edges $f(u_i)u_{i+1}$ with $u_{m+1} = u_1$. Replacing the matching, the vertices of G_i are now in the cycle $u_1 C_{u_1} f(u_1) u_2 C_{u_2} f(u_2) \dots u_m C_{u_m} f(u_m)$. The final part of the claim is obvious from this point: we only modify edges between V_i^1 and V_{i+1}^2 , all of which are on the same cycle in F' .

We now apply the claim to every $V_i V_{i+1} \in J$ sequentially; call the resulting 1-factor F'' . Firstly, note that by (ii) every cycle of F'' contains a vertex in V_i^1 for some $V_i V_{i+1} \in J$. Secondly, all the vertices of $V_i^1 \cup V_{i+1}^2$ lie on some common cycle of F'' , call it C_i . These two facts imply that $F'' = \cup_i C_i$. Consider any two edges $V_i V_{i+1}, V_j V_{j+1} \in J$ such that J avoids all edges in $V_{i+1} C V_j$, and let $P_{ij} = u_{i+1} \dots u'_j$ be the path in F guaranteed by (iii). Since F is a vertex-disjoint union of cycles, it must be that u_{i+1} and u'_j lie on the same cycle in F . Thus, by the last part of the claim, u_{i+1} and u'_j lie on a common cycle in F'' , but $u_{i+1} \in C_i$ and $u'_j \in C_j$, so $C_i = C_j$. Therefore, all the C_i are identical, and F'' is a cycle. (Note that for each edge of J , the vertices in the corresponding 1- and 2-subsets of the clusters belong to the same cycle, so for any sequence of adjacent vertices they are all in the same cycle, and so we only needed to consider the segment gaps.)

Finally, if $J = E(C)$ and $V_i^1 \cap V_i^2 \neq \emptyset$ for all $i = 1, \dots, k$, then little work needs to be done. With the above condition and since all of $V_{i-1}^1 \cap V_i^2$ lies on C_{i-1} and $V_i^1 \cap V_{i+1}^2$ on C_i , we thus conclude that $C_i = C_{i-1}$. This is true for each i and so we only have one cycle. \square

In [2] they tried to justify u_{i+1} and u'_j being on the same cycle by stating that since P_{ij} wound around C it must have been that $P_{ij} \subset F''$ too. The problem with this is that the path P_{ij} may wind around C more than once. If this were to happen, it is possible that P_{ij} intersects V_l^1 or V_{l+1}^2 for some l with $V_l V_{l+1} \in J$. In this case, we may have replaced an edge of P_{ij} with a one of G_i upon exchanging the perfect matchings, and so the path P_{ij} is definitely not guaranteed to be in F'' .

The way I have written the proof above is now correct. However, to make their proof correct, we could replace (iii) with:

- (iii)* Whenever $V_i V_{i+1}, V_j V_{j+1} \in J$ are such that J avoids all edges in the segment $V_{i+1} C V_j$ 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 such that P_{ij} winds around C less than once.

The reason why this now works is because we are forcing the path P_{ij} to be of the form $P_{ij} = u_{i+1} u_{i+2} \dots u_{j-1} u_j$ where $u_l \in V_l$. Since J has no edges in the segment $V_{i+1} C V_j$, we know that $F \cap (V_{i+1}^2 \cup V_{i+2} \cup \dots \cup V_{j-1} \cup V_j^1) = F'' \cap (V_{i+1}^2 \cup V_{i+2} \cup \dots \cup V_{j-1} \cup V_j^1)$, i.e. no edges are changed in this part of the cycle C , thus $P_{ij} \subset F''$.

The following lemma has a slightly more relaxed version of regularity. The key difference is that since we have weakened the regularity condition, we now require the entirety of the adjacent clusters to be superregular, and our subsets V_i^1, V_{i+1}^2 have to be much larger (over half size). The hypotheses (i)-(iii) are the same (so, again, if we want to follow their proof we would need to replace (iii) by (iii)*).

Lemma 4.5. Suppose that $0 < 1/m \ll \epsilon \ll d < 1$. 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 G be a digraph on $V_1 \cup \dots \cup V_k$ such that $G[V_i^1, V_{i+1}^2]$ is (ϵ, d) -superregular for every $V_i V_{i+1} \in J$. For each edge $V_i V_{i+1} \in J$, let $V_i^1 \subset V_i$ and $V_{i+1}^2 \subset V_{i+1}$ be such that $|V_i^1| = |V_{i+1}^2| \geq (1 - d/2)m$. Suppose that is a 1-regular digraph with $V_1 \cup \dots \cup V_k \subset V(F)$ such that the following properties hold:

- (i) For each edge $V_i V_{i+1} \in J$ 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 $V_i V_{i+1} \in J$ such that D contains a vertex in V_i^1 .
- (iii) Whenever $V_i V_{i+1}, V_j V_{j+1} \in J$ are such that J avoids all edges in the segment $V_{i+1} C V_j$ 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 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 $V_i V_{i+1} \in J$. Moreover, if $J = E(C)$ then the last condition can be omitted.

The proof of this is omitted since it is virtually the same, differing only in the justification for why A is a robustly expanding graph, but this follows easily from the regularity.

5. ADDITIONAL ROBUST EXPANSION LEMMAS

In this section we collect some additional results that will be used in the following two sections.

Lemma 5.1. Let $q \in \mathbb{N}$. Suppose that $0 < 1/n \ll \epsilon \ll \nu \leq \tau \ll \alpha < 1$ and that $1/n \ll \xi \leq q\nu^3/3$. 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 multigraph 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 $(1 - \epsilon)\xi n \leq n_x^+, n_x^- \leq (1 + \epsilon)\xi n$ 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 $x \in V(G) = V(Q)$.

We will not prove this fully here for the reason that its proof is not particularly enlightening; it is nothing than a min-cut-max-flow argument.

Sketch of Lemma 5.1. Create a bipartite graph $H = (A, B)$ with $A = B = V(G)$, joining $a \in A$ to $b \in B$ once for every occurrence of ab in Q . Let each such edge have capacity 1. Add a source s^* , edges s^*a of capacity n_a^+ for $a \in A$, a sink t^* and edges t^*b with capacity n_b^- for $b \in B$. Writing $r = \sum_{x \in V(G)} n_x^+/n$, we go through cases to show the a minimum cut must have capacity at least rn . Suppose we have a minimal cut \mathcal{C} , we let $S := \{a \in A : s^*a \notin \mathcal{C}\}$ and $T := \{b \in B : bt^* \notin \mathcal{C}\}$, then the capacity of the cut is $c = \sum_{s \in A \setminus S} n_s^+ + e_H(S, T) + \sum_{t \in B \setminus T} n_t^-$.

We then just work through four cases, which arise by considering the relative sizes of $B \setminus T$ and S and the absolute size of S ; the cases are (writing $T' := B \setminus T$): (i) $|T'| \geq |S| + \nu n/2$; (ii) $|T'| \leq |S| + \nu n/2$ and $\tau n \leq |S| \leq (1 - \tau)n$; (iii) $|T'| \leq |S| + \nu n/2$ and $|S| \leq \tau n/2$; and (iv) $|S| \geq (1 - \tau)n$. Working through these is tedious, so we will only discuss the second, as it is the only one that relies on the robust outexpansion.

We first note that, writing $X = T \cap RN_{\nu, G}^+(S)$, $|X| \geq \nu n/2$, and each $x \in X$ receives at least νn edges from S , so

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

where in the penultimate inequality we used that $(1 + \epsilon)\xi \leq (1 + \epsilon)q\nu^2/3 \leq q\nu^2/2$ (for $\epsilon \leq 1/2$).

The remaining cases are similar and rely on a more mundane property: namely, the minimum semidegree αn of G . The idea is to note that c is at least the size of any (combination of) the three terms, and a different one is considered for each. \square

The following lemma tells us that robust expansion properties are also inherited by blow-ups of graphs. This fact is indispensable in the proofs of Theorems 1.3 and 1.4. We will outline the main argument and omit the calculations.

Lemma 5.2. 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.

Sketch of Lemma 5.2. We first say two vertices in G' are *friends* if they correspond to the same vertex in G (with every vertex being a friend of itself). The idea essentially is to take an $S' \subset V(G')$ with $2r\tau n \leq |S'| \leq (1-2\tau)rn$ and then call $x' \in S'$ *bad* if S' contains at most $\nu^2 r$ friends of x' . We then upper bound the number of bad vertices in this set. Let $S'' \subset S'$ be the vertices which are not bad, and let $S \subset V(G)$ be those which have a copy (i.e. correspond to a vertex) in S'' . The bound gives that $|S| \geq \tau n$, which presents us with two cases: $|RN_{\nu,G}(S)| \geq |S| + \nu n$, or $|S| \geq (1-\tau)n$ and $|RN_{\nu,G}(S)| \geq (1-\tau+\nu)n$ (by restricting to an $F \subset S$ with $|F| = (1-\tau)n$). If $x \in RN_{\nu,G}(S)$, then a copy x' of x has at least $\nu^2 r \cdot \nu n = \nu^3 |G|$ inneighbours in S'' (recalling no vertex in S'' is bad), so $x' \in RN_{\nu^3,G'}^+(S')$, thus $|RN_{\nu^3,G'}^+(S')| \geq r|RN_{\nu,G}(S)|$. We then perform some calculations to show that in either of the cases mentioned above $|RN_{\nu^3,G'}^+(S')| \geq |S'| + \nu^3 rn$. \square

The following result is similar to Lemma 3.2. Note that the expansion parameters are slightly worse. We will not prove this here, since the proof is not too dissimilar.

Lemma 5.3. Let M', n_0 be positive integers and let ϵ, d, ν, τ be positive constants such that $1/n_0 \ll 1/M' \ll \epsilon \ll d \leq d' \leq \nu \leq \tau < 1$ and $d' \leq \nu/20$. Let G a digraph on $n \geq n_0$ vertices such that G is a robust (ν, τ) -outexpander. Let R be the reduced digraph of G with parameters ϵ, d and M' with clusters of size m and let R' be the subdigraph of R whose edges correspond to pairs of density of at least d' . Then R' is a robust $(\nu/4, 3\tau)$ -outexpander.

Admittedly, this result has nothing to do with robust expansion, however, I think this is the best place for it.

Proposition 5.4. Any r -regular multidigraph G contains r edge-disjoint 1-factors.

Sketch of Lemma 5.4. We first create an undirected bipartite graph $H = (A, B)$ where $A = B = V(G)$ and add an edge from a to b for each occurrence of $ab \in E(G)$. Hall's theorem gives us a perfect matching, corresponding to a 1-factor, which we then remove and repeat $r-1$ times. \square

6. AN APPROXIMATE HAMILTON DECOMPOSITION

This section will be concerned with the proof of Theorem 1.4, which formed part of Staden's PhD thesis and was published [3].

At a glance, one can instantly comprehend its utility in the proof of Theorem 1.3; it guarantees edge-disjoint Hamilton cycles covering nearly all of the graph. Kühn's proof, in a nutshell, finds a *very* clever way of applying this: she first removes a *very carefully chosen* subdigraph H from G , then she applies the theorem and removes the Hamilton cycles to be left with a sparse regular subdigraph G' and she finally covers all edges of $G' \cup H$ with Hamilton cycles.

The proof is quite lengthy - about 44 pages - and the use of robust expansion in the proof is somewhat isolated, hence the structure of this section will be a rough sketch of the proof, highlighting the uses of robust expansion. The sketch will be in enough depth so that one has an overarching grasp of the general strategy, but many details and nuances will be omitted, since these detract from the goal of the essay.

6.1. Additional definitions. We introduce a few definitions that are specific to this chapter.

Definition 6.1. Let $G = (A, B)$ be an undirected bipartite graph. Given $\epsilon, d \in (0, 1)$, we say that G is (ϵ, d) -regular if it is ϵ -regular and $d_G(A, B) = d \pm \epsilon$. We say that G is $[\epsilon, d]$ -superregular if it is (ϵ, d) -regular and $d(a) = (d \pm \epsilon)|B|$, $d(b) = (d \pm \epsilon)|A|$ for all $a \in A$ and $b \in B$.

Definition 6.2. Suppose $0 < 1/M' \ll \epsilon \ll \beta \ll d \leq 1$. The *reduced multidigraph* $R(\beta)$ with parameters ϵ, β, d, M' is obtained from the reduced digraph R (with parameters ϵ, d, M') by setting $V(R(\beta)) := V(R)$ and adding $\lfloor d_{G'}(V_i, V_j)/\beta \rfloor$ directed edges from V_i to V_j if $V_i V_j \in E(R)$.

Definition 6.3. Suppose that $D = p \otimes C_n$ is a p -fold blow-up of a cycle C_n of length n . A p -unwinding of D is any edge-disjoint collection $C^1, \dots, C^{p'}$ of p' Hamilton cycles of D .

6.2. Rough sketch. In order to aid clarity, we will split up the sketch into subsections according to the order in which they appear in the proof. This should hopefully aid digestion and let one focus on a certain part of the (sketch) proof if desired.

At the start of the proof, 18 additional constants are chosen, each with its own ' \ll ' relation, and even more arise throughout the proof. I will try to avoid writing the specific constants most of the time, but sometimes it will be unavoidable, as, for example, I do not want to repeatedly remark that something is superregular without specifying which type of superregularity. Furthermore, it may sometimes be desirable to emphasise that the parameters have changed as a result of our latest step.

6.2.1. Applying the diregularity lemma. This first step is setting up the partition of $V(G)$ in the best possible way for the rest of the proof.

We first apply the diregularity lemma with parameters $\tilde{\epsilon}^{12}, d, M'$ to obtain clusters $\tilde{V}_1, \dots, \tilde{V}_{\tilde{L}}$ of size \tilde{m} , an exceptional set V_0 , a pure digraph G' and a reduced digraph \tilde{R} . Call this partition $\tilde{\mathcal{P}}$. Let \tilde{R}' be the spanning subdigraph whose edges correspond to pairs of densities at least d' . If $\tilde{E} = \tilde{V}_i \tilde{V}_j \in E(\tilde{R})$ write $G'(\tilde{E})$ for the subdigraph $(\tilde{V}_i, \tilde{V}_j)_{G'}$ and d_{ij} for the density; so $G'(\tilde{E})$ is $(\tilde{\epsilon}^{12}, d)$ -regular. Let $\tilde{R}(\beta)'$ be the multigraph formed from $\tilde{R}(\beta)$ by only including edges which also correspond to an edge of \tilde{R}' . Write $K = \lfloor d_{ij}/\beta \rfloor$.

The rest of the details of this section are somewhat technical, so we will summarise them. We firstly associate each edge \tilde{E} of \tilde{R} with K edge-disjoint regular subgraphs, each of which is associated with a unique edge from \tilde{V}_i to \tilde{V}_j . We modify all of the clusters $\tilde{V}_1, \dots, \tilde{V}_{\tilde{L}}$ to ensure that the subdigraphs to ensure that each said subgraphs mentioned above satisfy strong conditions on its in- and outdegrees; we do this by removing vertices and adding them to V_0 , henceforth referred to as the *core exceptional set*. In doing so, we retain the regularity of the subgraphs (with slightly worse parameters), and we obtain several bounds on the minimum and maximum semidegrees of $\tilde{R}', \tilde{R}(\beta), \tilde{R}'(\beta)$. Lemma 5.3 also gives that \tilde{R}' is a robust outexpander, and it is a subdigraph of $\tilde{R}'(\beta) \subset \tilde{R}(\beta)$, so we apply Lemma 5.1 with $n_U^\pm := d_{\tilde{R}(\beta)}^\pm(U) - \tilde{r}$ to $(G, Q) := (\tilde{R}', \tilde{R}'(\beta))$ to obtain a sub-multidigraph W of $\tilde{R}'(\beta)$ and hence of $\tilde{R}(\beta)$. Hence $\tilde{R}(\beta) \setminus W$ is an \tilde{r} -regular sub-multidigraph of $\tilde{R}(\beta)$, which, with the aid of Proposition 5.4, gives a decomposition into \tilde{r} 1-factors $\tilde{F}_1, \dots, \tilde{F}_{\tilde{r}}$ of $\tilde{R}(\beta)$.

6.2.2. Thin auxiliary digraphs. We now want to remove (eventually) some subdigraphs which will act as *reservoirs* of well-distributed edges which will be used at various stages in the proof. These will be: H_0^\pm , which connects blown-up cycles; H_1^\pm , which connects vertices in the *special exceptional sets* $V_{0,i}$; and H_2 , which will construct *balancing edges*.

Each of the graphs defined above will have vertex set $V(G)$. To obtain them, for each edge E of $\tilde{R}(\beta)$, $G'(E)$ is a $(\sqrt{\tilde{\epsilon}}, \beta)$ -regular pair and we apply a certain lemma - which is not terribly relevant so will not be stated - to obtain, for $\gamma_1 := \beta_1 := (1-5\gamma)\beta$ and $\gamma_2 = \dots = \gamma_6 := \gamma\beta$, six edge-disjoint

pairs J_1, \dots, J_6 where J_k is $(\tilde{\epsilon}^{1/24}, \gamma_k)$ -regular, which we call $G^*(E), H_0^+(E), H_0^-(E), H_1^+(E), H_1^-(E)$ and $H_2(E)$ respectively. For each such $H(E)$, let $H = \cup_{E \in \tilde{R}(\beta)} H(E)$. For the sake of the proof, $G^*(E)$ can be weakened to being $(\epsilon/8, \beta_1)$ -regular, and the others to $(\epsilon, \gamma\beta)$ -regular.

For H_0^\pm , if $\tilde{A}\tilde{B}$ is an edge of \tilde{R} then for at least $(1 - \epsilon')|\tilde{A}|$ of $x \in \tilde{A}$ and $(1 - \epsilon')|\tilde{B}|$ of $y \in \tilde{B}$, we have $|N_{H_0^+}^+(x) \cap \tilde{B}|, |N_{H_0^-}^-(x) \cap \tilde{A}| \geq \gamma d \tilde{m}/2$. For all $x \in V(G) \setminus V_0$, we have $\gamma \tilde{\alpha} n/3 \leq d_{H_1^\pm}^\pm(x) \leq 2\gamma \tilde{\alpha} n$. It must be noted that H_0^\pm, H_1^\pm, H_2 are absolutely no different, and that we just pick out these properties because they are for what we will use the graphs.

6.2.3. Unwinding cycles. The first step of this section is to remove vertices from each cluster in each cycle in \tilde{F}_t so that each edge E of now corresponds to an superregular graph $G^*(E)$; we have exchanged regularity for superregularity at the expense of a small number of vertices. These are now called *adapted primary (t)-clusters*. Let $\tilde{V}_{0,t}^{\text{spec}}$ denote the set of vertices in G which were removed from the clusters in this step and let $\tilde{V}_{0,t} = V_0 \cup \tilde{V}_{0,t}^{\text{spec}}$. We denote the collection of the adapted primary (t)-clusters together with $\tilde{V}_{0,t}$ by $\mathcal{P}(t)$.

We begin by finding two nested sequences of uniform refinements \mathcal{P}_j and $\mathcal{P}_j(t)$ of $\tilde{\mathcal{P}}$ and $\mathcal{P}(t)$, respectively, for $j = s, p, 2p$. \mathcal{P}_s is a uniform s -refinement of $\tilde{\mathcal{P}}$, \mathcal{P}_p a uniform p -refinement of \mathcal{P}_s and \mathcal{P}_{2p} a uniform 2 -refinement of \mathcal{P}_p ; so we can also say that \mathcal{P}_{2p} is a uniform $2sp$ -refinement of $\tilde{\mathcal{P}}$. Note that we omit the uniformity parameter ϵ , since it adds unnecessary clutter. The next step is to define some blow-up graphs: $R_s := s \otimes \tilde{R}$, such that for $W \in V(\tilde{R})$ the corresponding vertices are the subclusters of W in $\mathcal{P}_s(t)$; $R_p := p \otimes R_s$ where for $U \in V(R_s)$ the vertices in R_p are the subclusters of U in $\mathcal{P}_p(t)$; and $R_s(\beta)$ and $R_p(\beta)$ are defined analogously. If $\tilde{E} = \tilde{U}\tilde{W} \in E(\tilde{R}(\beta))$ and U and W are s -clusters which are subclusters of \tilde{U}, \tilde{W} respectively, then there is a unique edge E in $R_s(\beta)$ from U to W corresponding to \tilde{E} , to which we associate the subgraph $G^*(E) := G^*(\tilde{E})[U, W]$; we can do the same thing for each edge F of $R_p(\beta)$.

From this point we then unwind each cycle C (of length K , say) of \tilde{F}_t in $R_s(\beta)$, obtaining $s - 1$ edge-disjoint 1-factors F'_j (which we say has *original factor type t*). So, in total we have $r_s := (s - 1)\tilde{r}$ edge-disjoint 1-factors F'_1, \dots, F'_{r_s} of $R_s(\beta)$. For every $V \in \mathcal{P}_s(t)$, we let V^1, \dots, V^p be the p -clusters contained in V . We now do the same, unwinding each cycle D of each F'_j in $R_p(\beta)$, but we can do this in a very special way, which will turn out to be useful later. We have that the p -clusters V_l^k for fixed k and $l \leq Ks - 2$ have pairwise distance at least p on this cycle D_d in $R_p(\beta)$. Moreover, for each $1 \leq l \leq Ks$, V_l^1, \dots, V_l^p lie on the same cycle D_d . Now, F'_j corresponds to a set of $p - 1$ edge-disjoint 1-factors F_i , and for each of which we say that it has *intermediate factor type j* and *original factor type t*. For each i , write $V_{0,i}^{\text{spec}} := \tilde{V}_{0,i}^{\text{spec}}$ and $r_p := (p - 1)r_s$. So, we now have r_p edge-disjoint 1-factors F_1, \dots, F_{r_p} of $R_p(\beta)$.

Finally, for each edge $E \in E(F_i)$ we let $G_i(E) := G^*(E)$ and $G_i := \cup_{E \in E(F_i)} G_i(E)$, which we call the i th *slice*. Clearly G_1, \dots, G_{r_p} are edge-disjoint. Since each $E \in E(F_i)$ corresponds to a unique $\tilde{E} \in E(\tilde{F}_t)$, it follows that $G_i(E)$ is $[\epsilon', \beta_1]$ -superregular.

6.2.4. Red clusters and edges. This step exists primarily to lay some groundwork for future sections; it specifies the properties that edges between the exceptional vertices and the rest of $V(G)$ need to satisfy. In the following section we will remove some *bridge vertices* $V_{0,i}^{\text{bridge}}$ from each $G_i \setminus (V_0 \cup V_{0,i}^{\text{spec}})$ and change their neighbourhoods in such a way that the blown-up cycles in G_i are connected via *bridge vertices*. Write $V_{0,i} := V_0 \cup V_{0,i}^{\text{spec}} \cup V_{0,i}^{\text{bridge}}$.

An edge incident to an exceptional vertex and any vertex in a cluster of G_i will be called *i-red* (or just *red*). The goal will be to add edges to G_i such that it is a spanning almost-regular digraph, that no non-exceptional vertex has large i -red degree and that the set of red vertices is well-distributed.

For each i , we will only add red edges incident to carefully chosen $2p$ -clusters. Letting $j = j(i)$ and $t = t(i)$, U_1, \dots, U_p be the p -subclusters of each s -cluster U of $\mathcal{P}_s(t)$, and $U(l)$ and $U(l+p)$ be the $2p$ -clusters contained in U_l , we will add red edges between $V_{0,i}$ and $U(k)$ only if: (a) $t \equiv k \pmod{2p}$; and (b) U is not a *clean* cluster in F'_j . Here *clean for F'_j* means that U belongs to the last K clusters of the cycle of F'_j . One can show that for each adapted primary cluster, exactly one subcluster in $\mathcal{P}_s(t)$ is clean. From now on, when we say (a) or (b) with no further description this is what we are referring to.

The red edges will need to satisfy many properties for our purposes, which we will summarise. Firstly, any 1-factor f of G_i has a path between any pair of successive cycles D_j of F_i . We do this in the sense that we can find a sequence $D_1 x_{i,1} D_2 x_{i,2} \dots x_{i,l} D_l$ where each D_j is a cycle of F_i and each such cycle appears at least once in the sequence, and $V_{0,i}^{\text{bridge}} := \{x_{i,1}, \dots, x_{i,l}\}$ with each such $x_{i,j}$ having exactly κ inneighbours in D_j and κ outneighbours in D_{j+1} . This property will be very useful when we want to apply Lemma 4.4. We also want the red edges to be well-distributed in the sense that: no one vertex has too many red edges; if two red p -clusters are on a cycle of F_i then they are at least p apart; and if a p -cluster contains the final vertex of a red edge in G_i then it contains no initial edge of one, and vice versa. Finally, the slices are all edge-disjoint and $G_i(E)$ is $[2\epsilon', \beta_1]$ -superregular for all $E \in E(F_i)$.

Let $F := r_p/4p$. For each $1 \leq k \leq 2p$, the number of digraphs G_i whose original type t satisfies (a) is $2F$, and we choose an ordering of these. We now decide which of the red $2p$ -clusters send and receive red edges. Let $t = t(i)$ and k satisfy (a). Suppose G_i is the f th graph with original type t (with respect to the ordering chosen above). Recalling that for each adapted primary cluster $W \in \mathcal{P}(t)$ only one s -subcluster is clean. We can pick two lots of $s/2 - 1$ s -clusters of W from each half (after picking an ordering) which avoid this clean cluster. For each s -cluster W_l in W , we call the k th $2p$ -cluster $W_l(k)$ *in-* or *out-red* depending on whether f and l are both in same halves (of their index sets) or both different, respectively. Naturally, a p -cluster V containing an in-red $2p$ -cluster is *in-red*, so the number of in-red p -clusters in each adapted primary cluster W is exactly $s/2 - 1$.

6.2.5. Connecting blown-up cycles. The point of this step is to modify each G_i so that we can ensure that each 1-factor f of G_i has a path connecting any pair of consecutive cycles of F_i . This will be done through choosing *bridge vertices* $x_{i,j}$ in $V(G) \setminus (V_0 \cup V_{0,i}^{\text{spec}})$, whose neighbourhoods we will choose from H_0^\pm .

There are two claims that start this section. The first of which, that uses robust expansion to guarantee a Hamilton cycle in \tilde{R} via Theorem 3.1, says that we can find a sequence $A_1 B_1 A_2 B_2 \dots A_{\bar{L}} B_{\bar{L}} A_1$ of p -clusters in R_p such that: the adapted primary clusters containing A_j and B_j have an edge between them in \tilde{R} ; the predecessor of A_j in F_i is in-red and B_j is out-red; B_j and A_{j+1} lie in the same adapted primary cluster; and every adapted primary cluster contains exactly one A_j and one B_j . The bridge vertices will be chosen in the sets A_j . The subsequent claim essentially says that there will be plenty of candidates for such vertices, i.e. for any of the $2p$ -subclusters of the predecessor of A_j and of B_j we can find plenty of vertices in H_0^- and H_0^+ (respectively) so that their in- and outneighbourhoods (respectively) intersect these two sets in at least κ elements.

The remainder of the section is concerned with finding these bridge vertices $x_{i,j} \in A_j$ and for each of which add κ in- and outedges from H_0^- and to H_0^+ , respectively, in G_i , and then showing that they satisfy all the required properties for being red edges.

The bridge vertices are removed, and in order to ensure that all the clusters have the same size extra vertices are removed. The adapted (t) -clusters become *adapted primary $[i]$ -clusters*.

6.2.6. *Incorporating the core exceptional set V_0 .* So far we have ignored $V_0 \cup V_0^{\text{spec}}$, as both currently have no edges in each G_i . We need to ensure that every vertex in $V_{0,i}$ has outdegree κ . We will focus on V_0 to begin with.

The section begins by proving that for each base $2p$ -cluster which is a subcluster of a base primary cluster W , in G we can find F edge-disjoint bipartite graphs $E_1^+(X), \dots, E_F^+(X)$, with all edges oriented from V_0 to X such that the outdegree of each vertex is sufficiently large, and the indegree of each element of X is not too large; and similarly with edges from X to V_0 . The proof is probabilistic and makes use of the uniform $2sp$ -refinement \mathcal{P}'_{2p} of $\tilde{\mathcal{P}}$.

After this, it is shown that we can add edges to each V_0 so that $d_{G_i}^\pm(x) = \kappa$. We will give a brief overview. Letting G_i be the i th digraph with original type t satisfying (a), for each base s -cluster W_l , we let $W_l(k)$ be the k th base $2p$ -cluster contained in W_l , and we apply the claim to said $2F$ bipartite graphs. We let W_l' be the s -[i]-cluster associated with W_l , we similarly define $W_l(k)'$, and we let $E_f^+(W_l(k)')$ be the subdigraph of $E_f^+(W_l(k))$ and define $E_f^-(W_l(k)')$ similarly. If $1 \leq f \leq F$ is in the first half, then we add edges in $E_f^+(W_l(k)')$ if $l \in I^+(W)$ or $E_f^-(W_l(k)')$ if $l \in I_W^-$, and the reverse if $F < f \leq 2F$ (with the subscript replaced by $f - F$). This means all edges from $G_i \setminus V_0$ to V_0 have initial vertex in an out-red cluster and similarly for the in-red clusters. Clearly, the edges assigned to G_i and $G_{i'}$ for $i \neq i'$ are disjoint. One shows, after calculations, that $d_{G_i}^\pm(x) \geq \kappa$, so edges can be deleted from each $x \in V_0$ such that $d_{G_i}^\pm(x) = \kappa$.

6.2.7. *Incorporating the special exceptional set $V_{0,i}^{\text{spec}}$.* The point of this section is to ensure that all vertices in $V_{0,i}^{\text{spec}}$ have in- and outdegree equal to κ using edges from H_1^\pm .

The first thing to do is to obtain a further property of H_1^\pm . Write S_i^+ and S_i^- for the vertices of the out- and in-red $2p$ -[i]-clusters, respectively, $U(k)$ for the k th base $2p$ -cluster of $U \in \mathcal{P}'_s$, $H_{1,k}^+$ for the spanning subdigraph of H_1^+ consisting of all edges whose final vertex is in $\cup_{U \in \mathcal{P}'_s} U(k)$ and $H_{1,k}^-$ defined similarly. We have that for all $x \in V(G) \setminus V_0$, $\gamma \tilde{\alpha} n / (20p) \leq |N_{H_{1,k}^+}^+(x) \cap S_i^-|, |N_{H_{1,k}^-}^-(x) \cap S_i^+| \leq \gamma \tilde{\alpha} n / p$. The proof comes from considering the partitions and how close they are.

After this, there is a claim saying that there are subdigraphs $Q_i^+ \subset H_1^+$ and $Q_i^- \subset H_1^-$ with the following properties: they are all pairwise edge-disjoint; for each $x \in V_{0,i}^{\text{spec}}$ we have that $|N_{Q_i^+}^+(x) \cap S_i^-|, |N_{Q_i^-}^-(x) \cap S_i^+| \geq \kappa$; and for all $y \in V(G) \setminus V_{0,i}$, the in- and outdegrees are small. The proof is probabilistic and will not be discussed.

We then apply the claim. So, for each $x \in V_{0,i}$ add exactly κ edges in Q_i^+ from x to S_i^- and exactly in Q_i^- to x from S_i^+ . The remainder of the section just checks that the desired properties of red edges hold.

6.2.8. *Finding shadow balancing sequences and adding balancing sequences.* At this point, we have now incorporated all exceptional vertices to form r_p edge-disjoint slices G_i of G . Moreover, each slice is a spanning almost-regular subdigraph of G . What we aim to do at this point is add further red edges, but this time between non-exceptional vertices, in such a way that the number of red edges sent out by vertices in each cluster V equals to the number received by its successor V^+ on the cycle of F_i containing V . We call this the *balancing property* and we can briefly justify why it is needed. Suppose that V is out-red and there is a 1-factor f containing a red edge from $x \in V$ to $V_{0,i}$. Recall that by construction of G_i , the only non-exceptional edges entering V^+ are from V . So, in the case that V^+ is not red, it must be that $f[V, V^+]$ is a perfect matching (we know it is perfect due to the fact $|V| = |V^+|$). However, this would be a contradiction, since there can be no edge in f from x to a vertex in V^+ (since the edge containing x goes to $V_{0,i}$). Naturally, there is a similar problem for U, U^- when U is in-red. We conclude that this property is necessary, and we will later see that it is actually sufficient to just add edges to either predecessor or successor of a red cluster.

Given G_i , let the set of red p -clusters be denoted by $T := T_{\text{in}} \cup T_{\text{out}}$, with $T_{\text{in}}, T_{\text{out}}$ defined in the obvious way. We let $s_i^\pm(V) := \sum_{y \in V} |N_{G_i}^\pm(y) \cap V_{0,i}|$ be the number of red edges leaving entering V , respectively. We define a *balancing sequence* B_i with respect to G_i to be a spanning subdigraph of H_2 such that the in- and outdegrees of any $y \notin V_{0,i}$ is small (bounded by b) and with the following degree conditions: $d_{B_i}^+(V) = s_i^-(V^+) + c$ if $V \in T_{\text{in}}^-$, c if $V \in T_{\text{out}}$ and 0 otherwise; and $d_{B_i}^-(V) = c$ if $V \in T_{\text{in}}$, $s_i^+(V^-) + c$ if $V \in T_{\text{out}}^+$ and 0 otherwise. Here c is defined in terms of several parameters. We see that by adding the edges from each B_i , we indeed have the desired balancing property mentioned.

Before obtaining the balancing sequences, we first obtain *shadow balancing sequences* B'_i . B'_i is a multigraph with vertex set $V(R^*)$, where R^* is an auxiliary digraph on T which we will not define here due to its rather technical and notation-heavy definition, with degree sequences corresponding to the degree conditions mentioned above, i.e. $n_V^+ = s_i^-(V) + c$ for $V \in T_{\text{in}}$ etc. It is clear that the method to attack this is to first show R^* is a robust outexpander and then apply Lemma 5.1. In order to do so, the idea of the proof is to first use the fact that $R_p^{\text{in}} := R_p[T_{\text{in}}]$ and $R_p^{\text{out}} := R_p[T_{\text{out}}]$ are robust outexpanders by Lemma 5.2, since they are both $(s/2 - 1)$ -fold blow-ups of \tilde{R} , which is a robust outexpander. Afterwards, one uses these to split up $RN_{\nu^+, R^*}^+(S)$ and eventually obtain the correct lower bound; we omit the fiddly details here.

Writing R_i^* to make the i -dependence explicit, for each edge $E' \in E(R_i^*)$ there is a unique edge $E \in E(R_p)$ corresponding to it (by the construction of R_i^*). Let $c_i(E)$ be the number of times edge E of R_p is chosen due to B'_i . Replacing the chosen edges $E \in E(R_p)$ with $c_i(E)$ edges in $H_2(E)$, gives the desired balancing sequence B_i . It is not immediately clear whether we can do this for all $1 \leq i \leq r_p$ such that the edges are disjoint, nor whether we can maintain certain properties of red edges. So, the remainder of this section deals with these worries by considering the clusters and regularity. This is somewhat fiddly (parameter-wise), and not particularly interesting, so we will omit it. (This is mainly due to the regularity of H_2 .)

Concluding, for each $1 \leq i \leq r_p$, we have added the edges of B_i to G_i so that now $E(G_i)$ consists of edges from each cluster to its unique successor on F_i together with the i -red edges incident to $V_{0,i}$ and balancing edges B_i (which are also red).

6.2.9. Almost decomposing into 1-factors. We want to be able to decompose each G_i into 1-factors, and to do so we will find a κ -regular spanning subdigraph of each G_i . We will first remove a subdigraph $H_{3,i}$ of $G_i \setminus E(B_i)$ which we will use in the following section. (Note that we obviously do not want to remove any of the balancing edges we just added.) We do this by exploiting the fact that $G_i(E) \setminus E(B_i)$ is $(2\epsilon', \beta_1)$ superregular for each edge E of F_i : with $\beta_2 := (1 - \gamma^2)\beta_1$, we can obtain two edge-disjoint subdigraphs of G_i , which are a $[2\epsilon'^{1/12}, \gamma^2\beta_1]$ -superregular digraph $H_{3,i}$ and another $[2\epsilon'^{1/12}, \beta_2]$ subdigraph which we will continue calling $G_i(E) \setminus E(B_i)$.

Let \mathcal{T}_i be the collection of i -red edges incident to $V_{0,i}$. So, the collection of i -red edges is $\mathcal{T}_i \cup B_i$. As shorthand, write $d_i^\pm(x) := d_{\mathcal{T}_i}^\pm(x) + d_{B_i}^\pm(x)$ for $x \in V(G_i)$ for the number of red out- and inedges (respectively). From our work in the previous section, we have that for each $V \in V(F_i)$, $d_i^+(V) = d_i^-(V^+)$.

Constructing our regular digraph G_i^* is straightforward. For each edge E from V to V^+ in F_i , we seek a subdigraph $G_i(E)^*$ of $G_i(E)$ such that together with the red edges from $B_i \cup \mathcal{T}_i$, every vertex in V has outdegree κ and every vertex in V^+ has indegree κ . This way, $G_i^* := \bigcup_{E \in E(F_i)} G_i(E)^* \cup B_i \cup \mathcal{T}_i$ is a κ -regular spanning subdigraph of G_i . We will briefly explain how we achieve this. For $x \in V$, we let $m_x^+ = d_i^+(x)$ and $m_y^- = d_i^-(y)$ for $y \in V^+$, noting $\sum_{x \in V} m_x^+ = \sum_{y \in V^+} m_y^-$, and then apply a superregularity (rather than robustly expanding) version of Lemma 5.1 to obtain the desired spanning subdigraph $G_i(E)^*$ of $G_i(E) \setminus E(B_i)$.

It is clear that $G_1^*, \dots, G_{r_p}^*$ are edge-disjoint, and we apply Proposition 5.4 to each G_i^* to obtain κ edge-disjoint 1-factors $f_{i,1}, \dots, f_{i,\kappa}$ of G_i .

In light of the work we did in the previous section, the red edges now satisfy slightly different properties than what we mentioned several sections ago. One property is worth mentioning, since it will be important soon. We now have that any out-red cluster p -cluster V is preceded by $p - 3$ p -clusters which are not red, and is succeeded by an in-red p -cluster. We have the same statement if we interchange out and in. This is since previously the only red clusters were in $T = T_{\text{in}} \cup T_{\text{out}}$ and were separated by $p - 1$ non-red clusters, and now the only other red clusters are precisely $T_{\text{in}}^- \cup T_{\text{out}}^+$.

6.2.10. Merging 1-factors into Hamilton Cycles. We now finally want to make the 1-factors from each G_i into Hamilton cycles using Lemma 4.4, using the edges from our pre-reserved digraph $H_{3,i}$. Since the argument will be identical for each i , we will instead label the factors as f_1, \dots, f_κ (dropping the i dependence).

This part is somewhat technical, so we will try to get an overview without getting too bogged down in the details. We begin by calling a non-red cluster *black* and an edge in F_i *black* if its initial and final clusters are black. We see that for all black edges VV^+ in F_i , $f_j[V, V^+]$ is a perfect matching, since every vertex in G_i goes from V to V^+ by construction. We have that for every pair $U_{\text{out}}U_{\text{in}}$ of consecutive red clusters on a cycle of F_i is followed by exactly $p - 3$ black clusters; denote the path of length $p - 4$ of these black clusters be I_U . This means that we can choose $p - 4$ disjoint sets of edges J_1, \dots, J_{p-4} such that for each pair of consecutive red clusters $U_{\text{out}}U_{\text{in}}$, J_q contains exactly one edge of I_U .

We will first describe the procedure for f_1 . Let the cycles of F_i be D_1, \dots, D_l . Let K_1 be the 1-regular digraph consisting of all cycles in f_1 with a vertex in a cluster of D_1 . Apply Lemma 4.4 with D_1 , $J_1 \cap E(D_1)$, K and $H_{3,i}(J_1)$ playing the roles of C , J , F and G respectively. From this we obtain a matching M_1 in $H_{3,i}(J_1)$ and a cycle C_1 with $V(C_1) = V(K_1)$ such that $E(C_1) \subset K_1 \cup M_1$. We replace K_1 with C_1 and call the resulting 1-factor $f_1(1)$. By construction, all cycles of f_1 which contained a vertex in D_1 are now on the same cycle. We let $H_{3,i}^2 := H_{3,i} \setminus M_1$. We then repeat this $l - 1$ more times (once for each D_k) to obtain $f_1' := f_1(l)$. We then carry out a similar procedure for f_2, \dots, f_κ to obtain f_2', \dots, f_κ' , using J_q for $f_{(q-1)\kappa'+1}, \dots, f_{q\kappa'}$ where $\kappa' := \kappa/(p - 4)$.

There is the slight worry that this may fail due to us repeatedly removing edges from $H_{3,i}$ so that $H_{3,i}^t$ (after removing the t th matching) may no longer satisfy the hypotheses of Lemma 4.4. However, this is taken care of in the parameter selection: it can be shown that $H_{3,i}^{\kappa'}(E)$ is superregular (with sufficient parameters) for any edge E in J_q and for any $1 \leq q \leq p - 4$.

Having obtained the edge-disjoint 1-factors f_1', \dots, f_κ' , it remains to show that they are indeed Hamilton cycles. This is actually very straightforward. We take two cycles C and C' of f_1' where C contains a vertex x in some cycle D of F_i and C' contains a vertex x' in some cycle D' of F_i . The properties of the red edges tell us that we can find an interval $D_g x_g D_{g+1} x_{g+1} \dots x_{g'-1} D_{g'} x_{g'}$ with $D_g = D$ and $D_{g'} = D'$. Since the inneighbour of x_g in f_1' is in D (recalling it is a bridge vertex), we have by construction that $x_g \in V(C)$. From this it is easy to see that all vertices of D_{g+1} are also in $V(C)$ and so (after continuing along the subsequence) all vertices in D' are in C , so $C = C'$. We conclude that f_1' is a Hamilton cycle, and we can do the same for f_2', \dots, f_κ' .

We therefore have found κr_p edge-disjoint Hamilton cycles. By parameter selection, $r_p \kappa \geq (1 - \eta)r$, completing the proof.

7. THE FULL DECOMPOSITION

7.1. Preliminary lemmas. As will be discussed in the sketch, most of our work will be for finding Hamilton cycles in the graph $G - V_0$ which wind around a cycle V_1, \dots, V_k in R , which does not include the exceptional set. After having applied Theorem 1.4 it seems plausible that we can cover the edges between the exceptional set and the remaining vertices, but being able to

cover the edges between vertices in the exceptional set itself seems a far more difficult task. We essentially have no idea how the edges in $G[V_0]$ will look. It would be so much more logical if we were to first remove all of the edges connecting the vertices in this set (whilst keeping the graph regular) and then apply Theorem 1.4. The following lemma lets us do this.

Lemma 7.1. Suppose that $0 < 1/n \ll \epsilon \ll \nu \leq \tau \ll \alpha \leq 1$. Let G be a robust (ν, τ) -outexpander with $\delta^0(G) \geq \alpha n$ and let V_0 be a set of vertices in G with $|V_0| \leq \epsilon n$. Then there is a set of ϵn edge-disjoint Hamilton cycles in G which contain all edges of $G[V_0]$.

The strategy of this proof has some interesting ideas. The first step is to split up the edges of $G[V_0]$ into $t := \epsilon n$ edge-disjoint matchings M_1, \dots, M_t . The goal is then to inductively find t edge-disjoint Hamilton cycles C_1, \dots, C_t such that $M_i \subset E(C_i)$; this way we will cover all of $E(G[V_0])$. Now, after removing C_1, \dots, C_{i-1} , the clever part of the proof is that we contract $ab \in M_i$ to form a new vertex e such that its inneighbourhood is that of a and its outneighbourhood is the same as that of b , then what we are essentially doing is forcing the Hamilton cycle guaranteed by Theorem 3.1 to contain each edge of M_i (since it must pass through each e , which in the original graph corresponds to each edge ab).

Proof of Lemma 7.1. If we remove the orientation of each edge, then $G[V_0]$ has maximum degree less than ϵn . By Vizing's theorem, we can obtain an ϵn -edge-colouring of $G[V_0]$, corresponding to $t := \epsilon n$ matchings M_1, \dots, M_t .

Suppose we have found C_1, \dots, C_{i-1} . Let G_i be formed from G by removing $E(C_1 \cup \dots \cup C_{i-1})$. This is still a robust $(\nu/2, \tau)$ -outexpander with $\delta^0(G_i) \geq \alpha n - (i-1) \geq \alpha n/2$. We can see this from recalling that $\epsilon \ll \nu \ll \alpha$. Explicitly, take some $S \subset V(G)$, $\tau n \leq |S| \leq (1-\tau)n$, and some $y \in RN_{\nu, G}^+(S)$. Then y loses at most t inneighbours in S , and so if $\epsilon \leq \nu/2$ then y still has at least $n\nu/2$ inneighbours in S . Thus $y \in RN_{\nu/2, G_i}^+(S)$ and $|RN_{\nu/2, G_i}^+(S)| \geq |RN_{\nu, G}^+(S)| \geq |S| + \nu n$. Moreover, in this case, we also have that $\epsilon \leq \alpha/2$ and so $\delta^0(G_i) \geq \alpha n/2$.

Now, form G'_i from G_i by replacing each $ab \in M_i$ with an edge e such that $N_{G'_i}^+(e) = N_{G_i}^+(b)$ and $N_{G'_i}^-(e) = N_{G_i}^-(a)$. Then G'_i is still a robust $(\nu/3, 2\tau)$ -outexpander with $\delta^0(G'_i) \geq \alpha n/3$. Take some $S \subset V(G)$ with $2\tau n \leq |S| \leq (1-2\tau)n$ and let $S' := S \setminus V_0$. First note that $|S'| \geq \tau n$, so $|RN_{\nu/2, G_i}^+(S')| \geq |S'| + \nu/2$. Now, if $ab \in M_i$ is such that $a, b \in RN_{G_i}^+(S)$, then of course $e \in RN_{\nu/3, G'_i}^+(S)$ (since a is), but we effectively 'lose' one fewer vertex in this set. On the other hand, if $a, b \in N_{G_i}^-(y) \cap S$ for some $y \in RN_{\nu/2, G_i}^+(S)$, then $N_{G'_i}^-(y)$ 'loses' one vertex, thus $|N_{G'_i}^-(y)| \geq |N_{G_i}^-(y)| - |M_i| \geq |N_{G_i}^-(y)| - \epsilon n \geq (\nu/2 - \epsilon)n \geq n\nu/3$, if $\epsilon \leq \nu/6$, and so $y \in RN_{G', \nu/3}^+(S')$. Combining these two observations,

$$|RN_{\nu/3, G'_i}^+(S)| \geq |RN_{\nu/2, G_i}^+(S')| - \epsilon n \geq |S'| + (\nu/2 - \epsilon)n \geq |S| + (\nu/2 - 2\epsilon)n \geq |S| + \nu/3,$$

if we take $\epsilon \leq \nu/12$, confirming that G'_i is a robust $(\nu/3, 2\tau)$ -outexpander.

Thus G'_i has a Hamilton cycle C'_i by Theorem 3.1, but this corresponds to a Hamilton cycle in G_i (and so in G) containing M_i . We repeat this until we have found C_1, \dots, C_t as required. \square

7.2. General strategy and motivation. Before giving a more detailed sketch we will give an overview of the proof, omitting almost all details.

The central idea to this proof is the removal of a *robustly decomposable* subdigraph H , which means that if we add an edge-disjoint sparse regular digraph to it, then we have a Hamilton decomposition of their union. Recalling Theorem 1.4, which was the subject of the previous section, we know that we are guaranteed a decomposition of almost all of the edges of G , which motivates this definition. If we could find such a graph and remove it, then we would apply Theorem 1.4 to find many edge-disjoint Hamilton cycles and remove them, leaving us with a sparse

regular digraph H_0 , then, by the definition of H , we would obtain a Hamilton decomposition of $H \cup H_0$, and hence of G .

As one may expect, it is not clear at all whether such a graph actually exists. The method for constructing such an H is very involved and constitutes the bulk of the paper. We actually let H be the union of three graphs: the *preprocessing graph* PG , the *chord absorber* CA and the *parity extended cycle absorber* PCA . The general idea is that the PG takes care of the leftover edges from the application of Theorem 1.4, and then the remaining two graphs CA and PCA form the robustly decomposable graph H so that we can decompose the leftover of PG with this to get our Hamilton decomposition of G .

7.3. Additional definitions. Inevitably, there is a multitude of definitions in such a long paper, but many are necessary to really get to the heart of the argument. We will list a couple here, with the goal of making our sketch a bit more streamlined. Some of the definitions are particularly hefty with many parts (up to 9), and few of them concern us, so these will be summarised in the sketch.

Definition 7.2. A *shifted* walk from vertex A to vertex B in R is a walk $SW(A, B)$ of the form

$$SW(A, B) = V_1 C V_1^- V_2 C V_2^- \dots V_t C V_t^- V_{t+1},$$

where $V_1 = A, V_{t+1} = B$ and the edge $V_i^- V_{i+1}$ belongs to R for each $i = 1, \dots, t$. We call the edges $V_i^- V_{i+1}$ the *chord edges* of $SW(A, B)$.

Definition 7.3. Given a shifted walk

$$SW(A, B) = V_1 C V_1^- V_2 C V_2^- \dots V_t C V_t^- V_{t+1},$$

the corresponding *chord sequence* $CS(A, B)$ from A to B consists of all the chord edges in $SW(A, B)$ in the same order as they appear. We say that V lies in the *interior* of $CS(A, B)$ if $V \in \{V_2, V_2^-, \dots, V_t, V_t^-\}$.

7.4. Sketch. This paper is sufficiently well divided into sections that the sketch here will have a different flavour. In Section 6, the sketch was somewhat chronological. However, the complexity of the argument here is much greater here, rendering that approach nearly impossible (to do in a meaningful and clear way). Instead, we will sketch the main ideas of the sections leading up to the final proof, and then give a detailed sketch of the proof itself.

7.4.1. Schemes, consistent systems, chord sequences and exceptional factors. This section is a very definition-heavy chapter with some very useful lemmas that will be used later. We will gloss over many of these definitions and try to summarise them, since, for example, a *consistent system* has 9 parameters, 7 items (a mixture of partitions, reduced digraphs, cycles and a digraph) and 9 properties, which would take at least half a page to just define. The summaries should hopefully be enough depth that we can glean the important information for future chapters whilst ignoring the precise details that contribute little to the overarching idea of the proof.

A (k, m, ϵ, d) -*scheme* is a four-tuple (G, \mathcal{P}, R, C) with the following properties. G is a digraph, \mathcal{P} is a partition of $V(G)$ into an exceptional set V_0 of size at most $\epsilon|G|$ and into k clusters of size m , and R is the vertex set of these clusters, i.e. the reduced digraph. Edges of R correspond to $(\epsilon, \geq d)$ -regular pairs. C is a Hamilton cycle in R , and edges in C correspond to $[\epsilon, \geq d]$ -superregular pairs. V_0 is independent in G .

$(G, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ is a *consistent* $(l^*, k, m, \epsilon, d, \nu, \tau, \alpha, \theta)$ -*system* if the following (summarised) properties hold. All graphs are robust (ν, τ) -outexpanders with minimum degree at least α of its size. \mathcal{P} is an l^* refinement of \mathcal{P}_0 , with reduced digraphs R and R_0 respectively. Edges of R correspond to $(\epsilon, \geq d)$ -regular pairs. C_0 and C are Hamilton cycles of R_0 and R respectively, and edges of C correspond to $[\epsilon, \geq d]$ -superregular pairs. In fact, C can be viewed as

winding around C_0 l^* times. We have an edge in R if and only if its parent clusters form an edge in R_0 . If a given vertex has a sufficiently large portion of a cluster of \mathcal{P}_0 in its outneighbourhood (inneighbourhood), then there is a lower bound on the size of its outneighbourhood (inneighbourhood) in each subcluster of this in \mathcal{P} . V_0 is independent in G . These properties imply that R is an l^* -fold blow-up of R_0 , which will allow us later on to apply Lemma 5.2. There is then a short lemma proving that if we delete just a few edges at each vertex, then we still have a consistent (albeit with worse parameters).

We now move on to the chord sequence subsection. We will first motivate this. The vast majority of Hamilton cycles found will wind around a blown-up $C = V_1 \dots V_k$, as we saw in Section 6, and we obviously need to incorporate the vertices of the exceptional set V_0 into the cycle. It is clear that winding around C is not necessary, but all of the framework leading up to this point makes it the best option. If we suppose that $V_0 = \{x\}$, x^+ and x^- are out- and inneighbours of x , then if x^+ and x^- are not in successive clusters then it is impossible to extend the path x^-x^+ into a Hamilton cycle in which the other edges wind around C . So, what we will aim to do is add edges which ‘balance out’ x^-x and xx^+ . This will be done with the aid of *shifted walks* and *chord sequences*.

There is a lemma and a proposition that follow. The lemma says that in a robust expander we can guarantee a short chord sequence which can completely avoid a small set. The proposition very important indeed, whose statement is as follows. Let $U(y)$ be the cluster containing y . Writing $CS := CS(U(x^+), U(x^-)^+)$, if we form CS' from CS by replacing each edge UW with an edge of $G[U, W]$, and supposing this forms a matching which avoids both x^- and x^+ , we form CS^* from CS' by adding the edge x^-x^+ . Taking $U \in \mathcal{P}$, we let U^1 and U^2 be the vertices in U which are not the initial or final vertices, receptively, of an edge in CS^* . Then, for each edge UW of C , $|U^1| = |W^2|$. This essentially means that it will be possible to find a perfect matching in $G[U^1, W^2]$ with additional regularity assumptions. Repeating this for all successive clusters, then the union of all these matchings and the edges in CS^* forms a 1-regular digraph F . Then, using, say, Lemma 4.4, we could turn this 1-factor into a Hamilton cycle.

We now move on to defining the concepts of complete exceptional sequences and exceptional factors, which will play a crucial role in the preprocessing step. An *original exceptional edge* in G is an edge with exactly one endvertex in V_0 , and an *exceptional cover* EC consists of precisely one outedge and one inedge to every vertex in V_0 (so $|EC| = 2|V_0|$). We obtain G^{basic} from G by adding all edges from $N^-(x)$ to $N^+(x)$ for every $x \in V_0$ and deleting V_0 ; any such edge is an *exceptional edge*. Both loops and multiple edges are now tolerated. We often write G^{orig} for G to be clear. It is easy to see that a 1-factor F in G^{basic} with exactly one exceptional edge associated with each exceptional vertex corresponds to a 1-factor F^{orig} of G^{orig} , and a 1-factor F of G clearly corresponds to a 1-factor F^{basic} of G^{basic} . A *complete exceptional sequence* CES is a matching in G^{basic} consisting of exactly one exceptional edge associated with each exceptional vertex. Note that CES^{orig} forms an exceptional cover, but the converse may not be true.

The reason for defining the above objects is that it lets us construct Hamilton cycles in G by first finding a Hamilton cycle D in G^{basic} which contains only one exceptional sequence and no other exceptional edges, then D^{orig} is a Hamilton cycle in G .

In the following section we will aim to convert the leftover edges from applying Lemma 1.4 into Hamilton cycles. What we will actually do is first break up the edges into path systems and then extend these into Hamilton cycles. This will be done with the aid of *exceptional factors* and *exceptional path systems* which will be defined below.

We will summarise the definition of a *complete exceptional path system* $CEPS$ (with respect to C) with parameters (K, L) . We first split up C into L canonical intervals, and for each interval $I = U_j \dots U_{j'}$ we have m/K vertex-disjoint paths in G^{basic} such that: each path starts in U_j and ends in $U_{j'}$; it contains a complete exceptional sequence, but no other exceptional edges, which

avoid the end clusters U_j and $U_{j'}$; and it contains precisely m/K vertices from each cluster in the interval and no others.

A *exceptional factor* EF with parameters (K, L) can be thought of as a 1-factor that induces K vertex-disjoint complete exceptional path systems for each of the L intervals with an added technical property that we will not include here.

This section finishes off with two big lemmas that will be used in the following section. Rather than writing out these lemmas in full (which would take the majority of a page) we will summarise them.

The first lemma lays the groundwork for the construction of the desired exceptional factors, and the second one provides a construction. Three assertions are made (under certain conditions). Firstly, there is a ‘localised’ exceptional cover of the exceptional set that is well distributed amongst the clusters of \mathcal{P} , i.e. none have too many. Secondly, we can find the chord sequences mentioned earlier that ‘balance out’ this exceptional cover, and said sequences are not too long and are, again, well distributed. Finally, we can find edges in G which correspond to the chord sequences mentioned above.

The final lemma of the section then tells us that we can find a set \mathcal{EF} of exceptional factors such that their original versions are all pairwise edge-disjoint subdigraphs of G .

7.4.2. The preprocessing step. As mentioned earlier in the general strategy, the point of this step is to find a preprocessing graph PG such that we can cover the edges leftover from an application of Theorem 1.4 with Hamilton cycles. The reason why we do this is that the edges at the exceptional vertices in the leftover may be distributed very badly. So, in order to control this, we remove PG and then (essentially) apply Theorem 1.4 to $G \setminus E(PG)$ and then cover the leftover G' by edge-disjoint Hamilton cycles in $PG \cup G'$. From this point, we replace G' by PG' , and then in the following section we will work on decomposing PG' with the help of extra edges from the chord absorber.

The strategy here is to first decompose G' into 1-factors and then split each 1-factor into small path systems. Then, using the additional edges from PG , extend each path system into a Hamilton cycle. As one would expect, the difficulty lies in finding suitable edges joining V_0 to the other clusters, and we will tackle this using the exceptional sequences covered in the previous section. Once we have these extra edges, we will use a *path system extender* to get Hamilton cycles.

Firstly, when splitting up the 1-factors, we want to break all cycles. The reason for this is obvious: we cannot extend a small cycle to a whole Hamilton cycle with extra edges. Each 1-factor H must have exactly one exceptional cover H^{exc} , as must every Hamilton cycle, however this very well may contain a cycle, so we must ensure that we split it up. This is done by introducing the notion of *styles*, which essentially describe a subset of indices of the subclusters to which a given set of edges is incident; this is a technical detail and not of great importance. One can show that each edge in H^{exc} has a different style, so we partition H^{exc} according to this, and then we further split each H_i into H_i^- and H_i^+ by placing one edge from each cycle into H_i^- and the rest in H_i^+ . So, we have now successfully separated the exceptional cover of H and broken all cycles.

Secondly, we will use an exceptional factor CB to extend each H_i into a (new) path system Q_i which has exactly one exceptional cover; equivalently, we decompose $H \cup CB^{\text{orig}}$ into edge-disjoint path systems Q_1, \dots, Q_s . We can do this in such a way that each Q_i does not have too many edges.

Having obtained these path systems Q_i , we now want to extend them into Hamilton cycles, which is what the *path system extender* PE will be used for, whose definition will be summarised as follows. A *path system extender* is a spanning subdigraph of $G - V_0$ consisting of the edge-disjoint union of two graphs $\mathcal{B}(C)_{PE}$ and $\mathcal{B}(R)_{PE}$ on $V(G) \setminus V_0$, which are blow-ups of C and

R , respectively, such that the respective edges in C or R correspond to superregular pairs (the variant with four parameters). The motivation for this construction, in particular the introduction of $\mathcal{B}(C)_{PE}$, is that we want to convert 1-factors into Hamilton cycles. However, in order to reach the stage of having 1-factors we will need additional edges, which will get from $\mathcal{B}(R)_{PE}$, and then we will use $\mathcal{B}(C)_{PE}$ to be our superregular reservoir of edges. Then we will use Lemma 4.4. It can be shown that such a PE will always exist.

We work with a general Q now and drop the i dependency for notational clarity. The first thing to do is to find a short chord sequence $CS(U(b), U(a)^+)$ for each edge ab of Q^{basic} . We do this iteratively, avoiding certain clusters if they have already been visited by too many chord sequences. Write $S(Q) := \cup_{ab \in E(Q^{\text{basic}})} CS(U(b), U(a)^+)$ as a multiset and s_E for the number of occurrences of $E \in E(Q^{\text{basic}})$. We let B_E be the bipartite subgraph of $\mathcal{B}(R)_{PE}$ corresponding to E and B'_E be the induced bipartite subgraph corresponding to the just under half of remaining *styles* after removing those of Q ; B'_E and Q are disjoint. Then for each edge E of R we pick a matching M_E of size s_E in B'_E avoiding all matchings chosen previously. We now let $Q_* := E(Q^{\text{basic}}) \cup (\cup_{E \in E(R)} M_E)$, and let Q^{in} and Q^{out} be the sets of final and initial vertices, respectively, of each edge in Q_* . For every cluster U in \mathcal{P} , we let $U^2 := U \setminus Q^{\text{in}}$ and $U^1 := U \setminus Q^{\text{out}}$. If $UW \in E(C)$, then $|U^1| = |W^2|$ by the proposition we mentioned in the previous section, and, exploiting the superregularity of $\mathcal{B}(C)_{PE}[U, W]$, we can show that the subgraph induced by U^1 and W^2 is superregular and hence show that it has a perfect matching M_{UW} . Let F be the union of all the M_{UW} and the edges of Q_* , then F is a 1-factor of $Q^{\text{basic}} \cup PE$. We can then check that the hypotheses of Lemma 4.4 are satisfied to apply it 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 to modify F into a Hamilton cycle D' of G^{basic} , corresponding to a Hamilton cycle $D := (D')^{\text{orig}}$ of G .

So, given an r -factor H , after splitting it up into edge-disjoint 1-factors, we can repeatedly do the above to find edge-disjoint Hamilton cycles covering all edges of H (we will obviously need r edge-disjoint exceptional factors for this).

Now, the only ‘problem’ with the above construction of PE is that it may not be regular. For later sections, we want the leftover (after removing the above Hamilton cycles) to be regular. This motivates the construction of a *preprocessing graph* PG that is the edge-disjoint union of PG^* and PG^\diamond , where PG^* is the edge-disjoint union of the path system extender and exceptional factors and PG^\diamond is a spanning subgraph of $G - V_0$ complementing the degree sequence of PG^* . The result is that PG is a spanning regular subgraph of G^{basic} , and that in PG^{orig} all non-exceptional vertices have the same in- and outdegrees, as do the exceptional vertices (but this number is different). Moreover, the maximum degree of PG^* is small.

Having the definition of PG under our belt, we now want to make sure that we can always find one within a consistent system. To do this, we first find (pairwise originally edge-disjoint) exceptional factors and remove their edges from G to get G_1 , which is still a consistent system. We then obtain our path system extender PG_{PE} in this subdigraph, and we let PG^* be the union of PG_{PE} and these exceptional factors. We now let G_2 be obtained from $G - V_0$ by deleting the edges in PG^* , which is still a robust outexpander, and hence, after a couple of further parameter checks, we can apply Lemma 5.1 to obtain PG^\diamond . Then $PG^* \cup PG^\diamond$ is our desired preprocessing graph.

This section terminates with a decomposition of the leftover PG' into small matchings, which will be extended to Hamilton cycles in the following section using the *chord absorber*. This is very technical, and not hugely relevant, so we will omit this.

7.4.3. The chord absorbing step. Recall that we have a leftover digraph G' which is a regular subdigraph of $G - V_0$. (This is a subdigraph of the preprocessing graph.) The goal of this section is to create *chord absorber* CA in G^{basic} such that $G' \cup CA$ contains Hamilton cycles covering all edges of G' and each such cycle contains exactly one complete exceptional sequence (so that

its original version corresponds to a Hamilton cycle in G), but with the added requirement that removing all of the Hamilton cycles from CA leaves a digraph which is a blow-up of $C = V_1 \dots V_k$. The first two requirements should not be surprising - they are exactly what we looked for in the previous section - however, the final one is where the difficulty lies. The motivation for this added condition is that we eventually want to cover all edges with Hamilton cycles, so we cannot repeatedly have some leftover graph; in other words, we aim to give the leftover graph much more structure.

The chord absorber that we will construct will consist of a blow-up of the cycle C , some exceptional factors and some additional edges which will be constructed with the aid of a *universal walk*.

We define a *universal walk for C with parameter l'* is a closed walk U with the following properties. For each $i = 1, \dots, k$ there is a chord sequence $ECS(V_i, V_{i+1})$ such that U contains all edges of these (counted with multiplicities) and all remaining edges of U lie on C . Moreover, each such chord sequence has at most $\sqrt{l'}/2$ edges and U enters and leaves every cluster exactly l' times. The notation of $ECS(\cdot, \cdot)$ differs from $CS(\cdot, \cdot)$ only in the sense that the former always refers to a chord sequence in U from V_i to V_{i+1} , which we call an *elementary chord sequence*. We also call the edges in $ECS(V_1, V_2) \cup ECS(V_2, V_3) \cup \dots \cup ECS(V_k, V_1)$ the *chord edges* of U . It can be helpful to view U as a multidigraph.

There are a couple of things to note before we move on. Firstly, any edge $V_i V_j \in E(R) \setminus E(C)$ (and so $i \neq j$) has the exact same number of occurrences in $ECS(V_1, V_2), \dots, ECS(V_k, V_1)$ as in U . We will hence identify these occurrences (note that the edges of $ECS(V_i, V_{i+1})$ are allowed to appear in a different order than in U).

If F is a chord edge of U and F' the next one, we let $P(F)$ be the subwalk of U from F to F' (without the edges F and F'). By definition, we note that $P(F)$ contains no chord edges. Note that if F_1 and F_2 are different occurrences of the same chord edge on U , then there is no reason why $P(F_1)$ and $P(F_2)$ need be the same. In light of this, we define the *augmented elementary chord sequence* $AECS(V_i, V_{i+1})$ to be the union 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 $AECS(V_i, V_{i+1})$ as follows: take the ordered sequence $ECS(V_i, V_{i+1})$ and after each edge F insert $P(F)$. Note that there is no reason why this should be connected. We see from this that $\cup_{i=1}^k AECS(V_i, V_{i+1})$ partitions the edges into k parts.

The definition of a universal walk seems somewhat abstract, and it is not immediately clear that such a walk exists. So, we prove that a robust (ν, τ) -outexpander has a universal walk U for C with parameter $l' := 36/\nu^2$. The proof is fairly rudimentary. We first choose the elementary chord $ECS(V_1, V_2), \dots, ECS(V_{j-1}, V_j)$ sequences greedily such that each has at most $3/\nu = \sqrt{l'}/2$ edges and avoids clusters visited too many times by previous chord sequences (recall this from an earlier proposition we mentioned). We then let U^* be the union of these chord sequences. The point now is to add edges to satisfy the rest of the properties and then check it indeed is a closed walk. Writing n_j^{out} for the number of edges of U^* leaving V_j and defining n_j^{in} similarly, one can quickly check that $n_j^{\text{out}} = n_{j+1}^{\text{in}}$. One then adds $l' - 1 - n_j^{\text{in}}$ copies of the edge $V_{j-1} V_j$ for each j , calling the resulting digraph U^\diamond , which is $(l' - 1)$ -regular. Finally, add another copy of each edge of C to obtain U which now satisfies all properties, and it is a quick check to confirm that one can order the edges in U to find a closed walk in R . This check is straightforward: by Proposition 5.4 we can decompose U^\diamond into 1-factors, we then traverse all cycles containing V_1 , then traverse $V_1 V_2$ of C , next we traverse all cycles containing V_2 which have not been traversed so far, then we traverse $V_2 V_3$ of C , and we repeat this until we reach V_1 again.

We are almost ready to define a *chord absorber*. However, we have a couple of preliminary definitions first.

$(G, \mathcal{P}, \mathcal{P}', R, C, U, U')$ is an (l', k, m, ϵ, d) -setup if (G, \mathcal{P}, R, C) is a (k, m, ϵ, d) -scheme and the following conditions hold. U is a universal walk for C with parameter l' and \mathcal{P}' is an ϵ -uniform l' -refinement of \mathcal{P} . Let $V_j^1, \dots, V_j^{l'}$ be the subclusters of V_j , then U' is a closed walk on \mathcal{P}' obtained from U by letting U' visit V_j^a the a th time that U visits V_j . Finally, each edge of U' corresponds to an $[\epsilon, \geq d]$ -superregular pair in G . One thing to notice is that U' is a Hamilton cycle since U visits each cluster in \mathcal{P} exactly l' times, and due to this we call it the *universal subcluster walk*.

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 and replacing each edge UV with a bipartite graph $\mathcal{B}(T)[U, W]$ that is ϵ -regular, all edges are oriented towards W , and its undirected version is r -regular. An r -blow-up of T is the same but without the ϵ -regularity condition.

We are now in a position to define the chord absorber. Like the definition of a path system extender, our graph will be the union of two blow-ups which provide the required edges to convert our leftover path systems into Hamilton cycles. A digraph CA on $V_1 \cup \dots \cup V_k$ is a *chord absorber* if it is the union of two digraphs $\mathcal{B}(C)$ and $\mathcal{B}(U')$ satisfying the following (summarised) properties. Firstly, $\mathcal{B}(C)$ is the union of $\mathcal{B}(C)^*$ and CA^{exc} where $\mathcal{B}(C)^*$ is an (ϵ, r) -blow-up of C and CA^{exc} is a collection of exceptional factors. Secondly, $\mathcal{B}(U')$ is an r' -blow-up of U' such that for every cluster A of \mathcal{P}' there is a partition A_1, \dots, A_4 into sets of equal size such that for all edges AB of U' and each j there is the same number (r' , to be specific) of edge-disjoint perfect matchings between A_j and B_j and $\mathcal{B}(U')[A, B]$ is the union of these. Finally, $\mathcal{B}(C)^*$, $\mathcal{B}(U')$ and $(CA^{\text{exc}})^{\text{orig}}$ are edge-disjoint. It follows that CA is a regular subdigraph of G^{basic} , however $CA^{\text{orig}} = \mathcal{B}(C)^* \cup \mathcal{B}(U') \cup (CA^{\text{exc}})^{\text{orig}}$ is such that all non-exceptional vertices have the same degree, as do all exceptional.

Having defined this digraph, one then shows that we can always find one. This will not be done here, since it is tedious and does not use anything relying on robust expansion.

We now move on to the key result of the section, which is arguably the key of the entire proof. As mentioned towards the end of the previous section, we are given a 1-factor H of $G - V_0$, which is split into small matchings H_i . Now, suppose that we are given a set of complete exceptional path systems H_i'' which are vertex-disjoint from the H_i , then this lemma states that we can extend the path system $H_i' = H_i \cup H_i''$ into a Hamilton cycle. The key thing here though is that the edges that one uses for C_i from $\mathcal{B}(U')$ is ‘predetermined’ and independent of H . To be more precise, for each 1-factor H , we split off a regular digraph $\mathcal{B}'(U')$, and the edges from the Hamilton cycles we generate *completely* cover $H \cup \mathcal{B}'(U')$, and use *none* of the other edges in $\mathcal{B}(U')$. Furthermore, a later lemma will in fact show that we can use all edges of $\mathcal{B}(U')$ in the process, this way the edges leftover from this section will form a subdigraph of $\mathcal{B}(C)$.

Due to the gravity of this result, we will briefly explain the main ideas of the proof. However, it is very long (about 10 pages) and there is little use of robust expansion, so we will not delve too deep.

The general strategy is quite similar to our main lemma from the previous section. We first *balance out* the edges of each H_i by adding edges corresponding to augmented chord sequences to form path systems W_i'' containing each H_i such that all edges of $\mathcal{B}'(U')$ are used. This will be done in such a way that we take the same number of edges from each bipartite graph corresponding to an edge of U' . Then each such W_i'' is extended into a 1-factor F_i , which is transformed into a Hamilton cycle using Lemma 4.5 and the edges of $\mathcal{B}^*(C)$.

The proof is roughly divided into five parts, so we will summarise the goal of each. The first step is to let W_i' be obtained from H_i by adding $ACS(U(y), U(x)^+)$ for each edge $xy \in H_i$, where $ACS(V_j, V_{j'}) := AECS(V_j, V_{j+1}) \cup \dots \cup AECS(V_{j'-1}, V_{j'})$. (Note that this means W_i' contains both edges of H_i and R , and we will eventually replace an edge from R with an edge in the corresponding bipartite subgraph.) We next show that in doing so, the sequences are *locally balanced*, that is: for each cluster V_j in \mathcal{P} , the number of edges of W_i' leaving V_j and the amount

entering V_{j+1} are the same (accounting for multiplicity). We view W'_i as a multiset consisting of edges in $E(U) \cup E(H)$ (or equivalently edges in $E(U') \cup E(H)$). We now show something more global: letting $W := \cup_i W'_i$, each edge of U occurs in W exactly the same number of times. Having these balancing properties is nice, but we will end up replacing edges of U with edges in G , and we do not want to overuse edges AB of U , so we then claim that any edge of U (or U') does not occur too frequently. We now add edges to each W'_i to obtain W''_i which satisfies the following properties: each W''_i is locally balanced and if we let $s_i(AB)$ be the number of times W''_i uses the edge AB , this number cannot be too large or too small - it in fact belongs to a small interval. Now let $W' = \cup_i W''_i$. Due to the clever construction, we then have that $S(AB) := \mathcal{B}'(U')[A, B]$ has the same number of edges as the number of occurrences of the edge AB , and so the idea at this point is replace each edge AB of U' in $W''_i \setminus E(H_i)$ by an edge of $S(AB)$ to obtain yet another digraph W'''_i with $V(W'''_i) = V(H) = V(G) \setminus V_0$ with some useful properties, which we briefly list. Firstly, W'''_i contains all edges of H_i and $W'''_i \cup H_i$ is a path system. For each edge AB of U , $W'''_i[A, B]$ has exactly $s_i(AB)$ edges of $S(AB)$. For each cluster V and its successor V^+ , the sum of the outdegrees in V is equal to the sum of the indegrees of V^+ . Finally, we of course want the W'''_i to be pairwise edge-disjoint. The construction of this is particularly fiddly, and commenting on it would achieve little. We now add all edges in H_i to W'''_i to form W^*_i , which is a graph on $V(H) = V(W'''_i)$. So, W^*_i is a path-system containing that still satisfies the above successor relation, and they are pairwise edge-disjoint. We now move onto the final step of the proof, in which we construct edge-disjoint Hamilton cycles C_i , containing all edges of $(W^*_i)^{\text{orig}}$, such that all edges in $E(C_i) \setminus (W^*_i)^{\text{orig}}$ lie in $\mathcal{B}(C)^*$. As mentioned earlier, the natural way to do this given our machinery is to first extend each such path-system into a 1-factor, and then apply Lemma 4.5, exploiting the robust outexpansion property of our graph and the superregularity of $\mathcal{B}(C)^*$. The first step is to, for each cluster $V_j \in \mathcal{P}$, let V_j^1 be the vertices in V_j with no outneighbours and V_j^2 with no inneighbours. We can now see the importance of the balancing conditions were desired earlier: they give that $|V_j^1| = |V_{j+1}^2|$. This not only is part of a hypothesis of Lemma 4.5, but also guarantees a perfect matching M_j in $\mathcal{B}(C)^*[V_j^1, V_{j+1}^2]$. Taking the union over the M_j with W^*_i gives a 1-factor F_i of $V(G) \setminus V_0$. We now check the remaining hypotheses of Lemma 4.5 and apply it to obtain the desired Hamilton cycles. The fact that they are disjoint comes from the construction: it is done inductively, removing previous cycles, and we still retain the superregularity property (albeit with worse parameters, but they are nevertheless sufficient).

There are two results left from this section; both are important for the rest of the paper. The first is nothing more than a tool for the second - it is not used elsewhere - thus we will just explain second. Suppose we are given a r -factor H of $G - V_0$. We then split this up into 1-factors F_i , and these 1-factors will be split into small matchings H_j^i . To this, we will associate a complete exceptional factor $CEPS_j^i$, and then extend $CEPS_j^i \cup H_j^i$ to a Hamilton cycle using edges of CA . As we remarked earlier, we do this in such a way that all of the edges of $\mathcal{B}(U')$ are used up, so that the leftover of the entire process is a blow-up of C .

7.4.4. Absorbing a blown-up cycle via switches. We now move onto the final preparatory chapter before the main proof. This stage has absolutely nothing to do with robust expansion, and is also the most fiddly, so this we will just give a very short and top-down overview. The reason why robust expansion here is not needed in this chapter is due to the fact that our leftover G' after applying the results from the previous chapters now has a great deal of structure involved. Recall that initially we knew absolutely nothing about the leftover after applying Theorem 1.4, but now we know that G' is a blow-up of C , i.e. all edges wind around C , which is a huge improvement. So, when trying to extend a 1-factor to a Hamilton cycle, we no longer need to exploit robust expansion and borrow edges from a superregular graph, we already have most of the structure we need.

The strategy of the section is to find a *cycle absorber* CyA which will be removed from the original digraph G at the beginning of the proof. This naturally will also be a blow-up of C . What this essentially does is allow us to find a *special* 1-factorisation of $H \cup CyA$ in which either half or all edges of each 1-factor will belong to CyA . Armed with this *special* 1-factorisation, we can then swap pairs of edges (of CyA) between pairs of 1-factors in $H \cup CyA$. The only caveat is that the parity of the total of number of cycles is preserved, so the final step is to extend CyA to a *parity extended cycle absorber* PCA to overcome this problem by instead considering triples of 1-factors to overcome this problem.

7.4.5. Proof of Theorem 1.3. The proof begins by proving a preliminary lemma. Recall the description of a *robustly decomposable graph* at the start of this chapter. Said lemma combines the chord and cycle absorbing steps into such a graph, G^{rob} , with some additional exceptional factors. This has the property that for any r -factor H of $G - V_0$, $G^{\text{rob}} \cup H$ has a (full) Hamilton decomposition. The reason for adding the exceptional factors is clear: CA and PCA are both digraphs on $G - V_0$, and we obviously need to include the exceptional vertices.

We now finally give a sketch of the proof. We begin by choosing a plethora of constants, which we will not list.

The first step is to apply Szemerédi's regularity lemma to G to obtain a partition \mathcal{P}_0 . We then add clusters to the exceptional set, if necessary, in order to ensure certain divisibility conditions between the parameters. Call the reduced digraph R_0 . Next we apply Theorem 3.1 to obtain a Hamilton cycle C_0 of R_0 .

We obtain a uniform refinement $\mathcal{P} = \{V_0, V_1, \dots, V_k\}$ of \mathcal{P}_0 , and let R be the digraph obtained from R_0 by replacing each cluster by its subclusters and each edge of R_0 with a complete bipartite graph. Now, Lemma 5.2 ensures that this digraph R is still a robust outexpander. We also let C be a Hamilton cycle in R obtained from C_0 by winding around it.

The next step is to find a universal walk U of C with parameter l' ; recall that this will be used to find G^{rob} . After adding a small amount of vertices from each cluster of \mathcal{P} to V_0 in order to ensure more divisibility conditions, we then let G_0 denote the digraph obtained from G by deleting all edges between vertices in V_0 . We then have that $(G_0, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ is a consistent system (with certain parameters). After obtaining a further uniform refinement \mathcal{P}'' of \mathcal{P} , we can now find a preprocessing graph PG in G_0^{basic} . Recall that we created this graph for later use, so we obtain G_1 from G_0 by deleting the edges in PG^{orig} , and we are still left with a consistent system (albeit with slightly worse parameters).

We obtain yet another uniform refinement \mathcal{P}''' of \mathcal{P} , and we also find exceptional factors EF_i in the $(G_1, \mathcal{P}_0, R_0, C_0, \mathcal{P}, R, C)$ -consistent system with respect to C and \mathcal{P}''' , such that the $(EF_i)^{\text{orig}}$ are pairwise edge-disjoint. We let $\mathcal{EF} = \cup_i EF_i$. Using these exceptional factors and the lemma mentioned at the start of this section, we then find a spanning subdigraph CA^\diamond of $G - V_0$, and we delete the edges of $CA^\diamond \cup \mathcal{EF}^{\text{orig}}$ from G_1 to obtain G_2 . We use said lemma yet again to find a spanning subdigraph PCA^\diamond of $G - V_0$, and then we define $G^{\text{rob}} := CA^\diamond \cup PCA^\diamond \cup (\mathcal{EF} \cup \mathcal{EF}')^{\text{orig}}$ and $G^{\text{absorb}} := PG^{\text{orig}} \cup G^{\text{rob}}$. One can show that by fiddling around with parameters, G^{absorb} is regular. So, we let G^Δ be obtained from G by removing all edges of G^{absorb} , which is also regular. We now apply Lemma 7.1 to obtain Hamilton cycles \mathcal{C}^Δ covering all edges in $G^{\text{absorb}}[V_0]$ (so that it now becomes an independent set), and we let $G^\#$ be the digraph obtained after deleting the edges in these cycles. Now, harnessing the properties of robust expanders, since we have only removed a few edges from each vertex, we still have a robustly outexpanding digraph, and so we can apply Theorem 1.4 to obtain a set $\mathcal{C}^\#$ of Hamilton cycles, covering almost all the edges of $G^\#$. Let H_0 be the leftover, after removing these edges. Now, since V_0 is an independent set in H_0 , we now have a regular subdigraph of G_0 , which we defined at the start of the sketch. With the aid of the preprocessing graph PG , we then find a set \mathcal{C}_0 of Hamilton cycles in $H_0 \cup PG^{\text{orig}}$ covering all of H_0 . Now, letting PG' be the leftover from PG^{orig} , we have every exceptional vertex

isolated and all non-exceptional vertices with the same degree, so $H_1 := PG' - V_0$ is a regular subdigraph of $G_0 - V_0$, and so $H_1 \cup G^{\text{rob}}$ has a Hamilton decomposition \mathcal{C}_1 .

We therefore have a full Hamilton decomposition of G : $\mathcal{C}^\Delta \cup \mathcal{C}^\# \cup \mathcal{C}_0 \cup \mathcal{C}_1$.

8. APPLICATION TO KELLY'S CONJECTURE

The subject of this section will be proving Theorem 1.5 from which Kelly's conjecture will be a simple corollary for large n .

We will begin by proving that if an oriented graph with the sum of its minimum in- and outdegrees and minimum degree is greater than $3n/2 + \epsilon$ then it is a robust outexpander. From this, it is easy to deduce that an oriented graph with minimum semidegree a tiny bit larger than $3n/8$ is a robust outexpander.

The proof of the theorem is somewhat tedious, and there is not much to it beyond counting, but I felt that it should be included for the sake of completeness, especially since it offers a sufficient condition for a oriented graph to be a robust expander. We begin by assuming for a contradiction that the graph is not a robust (ν, τ) -expander, so there is some $X \subset V(G)$ with $\tau n \leq |X| \leq (1 - \tau)n$ and $|RN_{G,\nu}^+(X)| < |X| + \nu n$, and (writing $RN_{G,\nu}^+(X) = RN$) we split up the vertex set $V(G)$ into $X \cap RN$, $RN \setminus X$, $X \setminus RN$ and what is left over. We essentially just find upper bounds on each of the different degrees mentioned in the previous paragraphs in terms of the sizes of these sets and then find a contradiction from this.

Lemma 8.1. Let $0 < 1/n \ll \nu \ll \tau \leq \epsilon/2 \leq 1$ and suppose that G is an oriented graph on n vertices with $\delta^+(G) + \delta^-(G) + \delta(G) \geq 3n/2 + \epsilon n$. Then G is a robust (ν, τ) -expander.

Proof. We suppose not. Take $X \subset V(G)$ with $\tau n \leq |X| \leq (1 - \tau)n$ such that $|RN_{G,\nu}^+(X)| < |X| + \nu n$. Write $RN := RN_{G,\nu}^+(X)$. We decompose $V(G)$ as $A := X \cap RN$, $B := RN \setminus X$, $C := X \setminus RN$ and $D := V(G) \setminus (A \cup B \cup C)$. So, we note that $RN = A \cup B$ and that $X = A \cup C$. The former implies that $|A| + |B| < |A| + |C| + \nu n$, and so

$$|B| < |C| + \nu n. \quad (8.1)$$

We now split the rest of the proof into three parts, getting an upper bound on each of $\delta^+(G)$, $\delta^-(G)$ and $\delta(G)$ in terms of these sets (and the parameters ν, τ).

Claim 1. $|A| + |B| + |C| \geq 2\delta^+(G) - 2\tau n$.

We split this into two cases depending on the size of A . Let us first suppose that $|A| \geq \tau n/2$. By the definition of C and D , they contain at most νn neighbours in A , so $e(A, C \cup D) \leq \nu n(|C| + |D|) \leq 2\nu|A|(|C| + |D|)/\tau \leq \tau n|A|/2$ (since $\nu \ll \tau$), thus

$$\begin{aligned} e(A, A \cup C \cup D) &= e(A, A) + e(A, C \cup D) \\ &\leq |A|^2/2 + 2\nu n(|C| + |D|)/\tau \\ &\leq |A|(|A| + \tau n)/2. \end{aligned}$$

We deduce that there must be some $x \in A$ with $|N^+(x) \cap (A \cup C \cup D)| \leq |A|/2 + \tau n/2$, and so $\delta^+(G) \leq |A|/2 + |B| + \tau n/2$. Applying (8.1) finishes the claim in this case: $2\delta^+(G) \leq |A| + 2|B| + \tau n < |A| + |B| + |C| + (\tau + \nu)n$.

On the other hand, suppose that $|A| \leq \tau n/2$ (so $|C| \geq \tau n/2$). We also trivially have that $e(C, C \cup D) \leq \nu n(|C| + |D|) \leq \tau n|C|/4$ (using $\nu \ll \tau$). Then $|C| \geq \tau n/2$ and

$$\begin{aligned} e(C, A \cup C \cup D) &= e(C, A) + e(C, C \cup D) \\ &\leq |C|\tau n/2 + \nu n(|C| + |D|) \\ &\leq 3\tau n|C|/4. \end{aligned}$$

We see that there is an $x \in D$ such that $|N^+(x) \cap (A \cup C \cup D)| \leq 3\tau n/4$, thus $\delta^+(G) \leq |B| + 3\tau n/4$, and we now apply (8.1).

Claim 2. $|B| + |C| + |D| \geq 2\delta^-(G) - 3\nu n$.

Consider first the case $D \neq \emptyset$. Since G is oriented, it must be that $\sum_{x \in D} |N^-(x) \cap D| \leq |D|(|D| - 1)/2$, and hence there is some $x \in D$ such that $|N^-(x) \cap D| \leq |D|/2$. By definition, $|N^-(x) \cap X| \leq \nu n$, thus $\delta^-(G) \leq |B| + |D|/2 + \nu n$. Apply (8.1).

If we have that $D = \emptyset$ then, since $|A \cup B| < (1 - \tau)n + \nu n < n$ by assumption, it must be that $C \neq \emptyset$. However, by definition, for each $x \in C$ $|N^-(x) \cap X| \leq \nu n$, so $\delta^-(G) \leq \nu n + |B|$. Apply (8.1).

Claim 3. $|A| + |B| + |D| \geq \delta(G) - 2\nu n$.

This is trivial if $C = \emptyset$, so we suppose it is not. Then $e(C, C) \leq e(X, C) \leq \nu n|C|$, so there is some $x \in C$ with $|N^+(x) \cap C| \leq \nu n$. But, by definition, $C \cap RN = \emptyset$, so $|N^-(x) \cap C| \leq 2\nu n$. Therefore $d(x) \leq |A| + |B| + |D| + 2\nu n$.

We can now finally arrive at a contradiction. We have that

$$\begin{aligned} 3n &= 3(|A| + |B| + |C| + |D|) \\ &\stackrel{(8.1)}{\geq} 3|A| + 4|B| + 2|C| + 3|D| - \nu n \\ &\geq 2\delta^+(G) + 2\delta^-(G) + \delta(G) + |A| + |B| + |D| - 6\nu n - 2\tau n \\ &\geq 2(\delta^+(G) + \delta^-(G) + \delta(G)) - 3\tau n \\ &\geq (3 + 2\epsilon - 3\tau)n \\ &> 3n, \end{aligned}$$

as desired. \square

We are now in a position to complete the proof. We note that if G is a robust (ν, τ) -outexpander, then it is also a robust (ν, τ') -outexpander for all $\tau' \geq \tau$.

Proof of Theorem 1.5. Let $\alpha = 3/8$ in Theorem 1.3 and let $\tau^* := \tau(3/8)$ be the guaranteed parameter. Now, choose additional constants n_0, ν and τ such that $1/n_0 \ll \nu \ll \tau \leq \epsilon/2, \tau^*$. By the previous lemma, we know that G is a robust (ν, τ) -outexpander, and therefore a robust (ν, τ^*) -outexpander. We thus apply Theorem 1.3 with $\alpha = 3/8$ to find a Hamilton decomposition of G . \square

Proof of Kelly's conjecture. Take $n \geq n_0$. Then, since $r = n/2$, G has a Hamilton decomposition into $(n - 1)/2$ edge-disjoint cycles. \square

9. CONCLUDING REMARKS

The primary focus of this essay has been to explore the recent concept of robust expansion and demonstrate its strong ties to Hamiltonicity, highlighting its fundamental use in proving that sufficiently large regular robustly expanding digraphs have a Hamilton decomposition, settling Kelly's conjecture from 1968. Due to its length, spread across two papers summing to 140 pages, much of this had to be sketched. However, the results which require robust expansion have been covered in more depth, which was either a full proof or a more detailed sketch including motivation of why this property is necessary.

Given its age and the fact its 'creators', Daniela Kühn and Deryk Osthus, have only published results which serve only as tools to prove results on Hamiltonicity, I have been able to prove my own result: asymptotically almost every graph is a robust expander. Robustly expanding digraphs

are still young and have only been explored in the context of Hamiltonicity. This therefore poses the question of whether they will have far-reaching consequences like expanders graphs.

REFERENCES

- [1] Daniela Kühn, Deryk Osthus and Andrew Treglown, *Hamilton degree sequences in digraphs*, add publisher
- [2] Daniela Kühn and Deryk Osthus, *Hamilton decompositions of regular expanders: A proof of Kelly's conjecture for large tournaments*, *Advances in Mathematics* 237 (2013) 62-146.
- [3] Daniela Kühn, Deryk Osthus and Katherine Staden, *Approximate Hamilton decompositions of robustly expanding digraphs*
- [4] W.T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, *Geometric and Functional Analysis* Vol. 7 (1997) 322-327
- [5] Shlomo Hoory, Nathan Linial and Avi Wigderson, *Expander graphs and their applications*, *American Mathematical Society*, Volume 43, Number 4, 2006, 439-561
- [6] N. Alon and J. H. Spencer, *The Probabilistic Method*, Second edition (2000)
- [7] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, Wiley-Interscience (2000)
- [8] Michael Mitzenmacher and Eli Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press (2005).
- [9] Daniela Kühn and Deryk Osthus, *Embedding large subgraphs into dense graphs*, *Surveys in Combinatorics*, London Math. Soc. Lecture Notes, 137-167
- [10] 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.